



A systematic literature review on educational recommender systems for teaching and learning: research trends, limitations and opportunities

Felipe Leite da Silva¹ · Bruna Kin Slodkowski¹ ·
Ketia Kellen Araújo da Silva¹ · Sílvia César Cazella^{1,2}

Received: 5 November 2021 / Accepted: 5 September 2022 / Published online: 14 September 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Recommender systems have become one of the main tools for personalized content filtering in the educational domain. Those who support teaching and learning activities, particularly, have gained increasing attention in the past years. This growing interest has motivated the emergence of new approaches and models in the field, in spite of it, there is a gap in literature about the current trends on how recommendations have been produced, how recommenders have been evaluated as well as what are the research limitations and opportunities for advancement in the field. In this regard, this paper reports the main findings of a systematic literature review covering these four dimensions. The study is based on the analysis of a set of primary studies ($N=16$ out of 756, published from 2015 to 2020) included according to defined criteria. Results indicate that the hybrid approach has been the leading strategy for recommendation production. Concerning the purpose of the evaluation, the recommenders were evaluated mainly regarding the quality of accuracy and a reduced number of studies were found that investigated their pedagogical effectiveness. This evidence points to a potential research opportunity for the development of multidimensional evaluation frameworks that effectively support the verification of the impact of recommendations on the teaching and learning process. Also, we identify and discuss main limitations to clarify current difficulties that demand attention for future research.

Keywords Educational recommender systems · E-learning recommendation · Systematic literature review · Computer-mediated teaching and learning

✉ Felipe Leite da Silva
fsliva@ufrgs.br

¹ Centro de Estudos Interdisciplinares em Novas Tecnologias da Educação, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

² Departamento de Ciências Exatas e Sociais Aplicadas, Universidade Federal de Ciências da Saúde de Porto Alegre, Porto Alegre, Rio Grande do Sul, Brazil

1 Introduction

Digital technologies are increasingly integrated into different application domains. Particularly in education, there is a vast interest in using them as mediators of the teaching and learning process. In such a task, the computational apparatus serves as an instrument to support human knowledge acquisition from different educational methodologies and pedagogical practices (Becker, 1993).

In this sense, Educational Recommender Systems (ERS) play an important role for both educators and students (Maria et al., 2019). For instructors, these systems can contribute to their pedagogical practices through recommendations that improve their planning and assist in educational resources filtering. As for the learners, through preferences and educational constraints recognition, recommenders can contribute for their academic performance and motivation by indicating personalized learning content (Garcia-Martinez & Hamou-Lhadj, 2013).

Despite the benefits, there are known issues upon the usage of the recommender system in the educational domain. One of the main challenges is to find an appropriate correspondence between the expectations of users and the recommendations (Cazella et al., 2014). Difficulties arise from differences in learner's educational interests and needs (Verbert et al., 2012). The variety of student's individual factors that can influence the learning process (Buder & Schwind, 2012) is one of the challenging matters that makes it complex to be overcome. On a recommender standpoint, this reflects an input diversity with potential to tune recommendations for users.

In another perspective, from a technological and artificial intelligence standpoint, the ERS are likely to suffer from already known issues noted on the general-purpose ones, such as the cold start and data sparsity problems (Garcia-Martinez & Hamou-Lhadj, 2013). Furthermore, problems are related to the approach used to generate recommendations. For instance, the overspecialization is inherently associated with the way that content-based recommender systems handle data (Iaquinta et al., 2008; Khushro et al., 2016). These issues pose difficulties to design recommenders that best suit the user's learning needs and that distance themselves from user's dissatisfaction in the short and long term.

From an educational point of view, issues emerge on how to evaluate ERS effectiveness. A usual strategy to measure the quality of educational recommenders is to apply the traditional recommender's evaluation methods (Erdt et al., 2015). This approach determines system quality based on performance properties, such as its precision and prediction accuracy. Nevertheless, in the educational domain, system effectiveness needs to take into account the students' learning performance. This dimension brings new complexities on how to successfully evaluate ERS.

As ERS topic has gradually increased in attraction for scientific community (Zhong et al., 2019), extensive research have been carried out in recent years to address these issues (Manouselis et al. 2010; Manouselis et al., 2014; Tarus et al., 2018; George & Lal, 2019). ERS has become a field of application and combination of different computational techniques, such as data mining, information filtering

and machine learning, among others (Tarus et al., 2018). This scenario indicates a diversity in the design and evaluation of recommender systems that support teaching and learning activities. Nonetheless, research is dispersed in literature and there is no recent study that encompasses the current scientific efforts in the field that reveals how such issues are addressed in current research. Reviewing evidence, and synthesizing findings of current approaches in how ERS produce recommendations, how ERS are evaluated and what are research limitations and opportunities can provide a panoramic perspective of the research topic and support practitioners and researchers for implementation and future research directions.

From the aforementioned perspective, this work aims to investigate and summarize the main trends and research opportunities on ERS topic through a Systematic Literature Review (SLR). The study was conducted based on the last six years publications, particularly, regarding to recommenders that support teaching and learning process.

Main trends refer to recent research direction on the ERS field. They are analyzed in regard to how recommender systems produce recommendations and how they are evaluated. As mentioned above, these are significant dimensions related to current issues of the area. Specifically for the recommendation production, this paper provides a three-axis-based analysis centered on systems underlying techniques, input data and results presentation.

Additionally, research opportunities in the field of ERS as well as their main limitations are highlighted. Because current comprehension of these aspects is fragmented in literature, such an analysis can shed light for future studies.

The SLR was carried out using Kitchenham and Charters (2007) guidelines. The SLR is the main method for summarizing evidence related to a topic or a research question (Kitchenham et al., 2009). Kitchenham and Charters (2007) guidelines, in turn, are one of the leading orientations for reviews on information technology in education (Dermeval et al., 2020).

The remainder of this paper is structured as follows. In Section 2, the related works are presented. Section 3 details the methodology used in carrying out the SLR. Section 4 covers the SLR results and related discussion. Section 5 presents the conclusion.

2 Related works

In the field of education, there is a growing interest in technologies that support teaching and learning activities. For this purpose, ERS are strategic solutions to provide a personalized educational experience. Research in this sense has attracted the attention of the scientific community and there has been an effort to map and summarize different aspects of the field in the last 6 years.

In Drachsler et al. (2015) a comprehensive review of technology enhanced learning recommender systems was carried out. The authors analyzed 82 papers published from 2000 to 2014 and provided an overview of the area. Different aspects were analyzed about recommenders' approach, source of information and

evaluation. Additionally, a categorization framework is presented and the study includes the classification of selected papers according to it.

Klašnja-Milićević et al. (2015) conducted a review on recommendation systems for e-learning environments. The study focuses on requirements, challenges, (dis)advantages of techniques in the design of this type of ERS. An analysis on collaborative tagging systems and their integration in e-learning platform recommenders is also discussed.

Ferreira et al. (2017) investigated particularities of research on ERS in Brazil. Papers published between 2012 and 2016 in three Brazilian scientific vehicles were analyzed. Rivera et al. (2018) presented a big picture of the ERS area through a systematic mapping. The study covered a larger set of papers and aimed to detect global characteristics in ERS research. Aiming at the same focus, however, setting different questions and repositories combination, Pinho, Barwaldt, Espíndola, Torres, Pias, Topin, Borba and Oliveira (2019) performed a systematic review on ERS. In these works, it is observed the common concern of providing insights about the systems evaluation methods and the main techniques adopted in the recommendation process.

Nascimento et al. (2017) carried out a SLR covering learning objects recommender systems based on the user's learning styles. Learning objects metadata standards, learning style theoretical models, e-learning systems used to provide recommendations and the techniques used by the ERS were investigated.

Tarus et al (2018) and George and Lal (2019) concentrated their reviews on ontology-based ERS. Tarus et al. (2018) examined research distribution in a period from 2005 to 2014 according to their years of publication. Furthermore, the authors summarized the techniques, knowledge representation, ontology types and ontology representations covered in the papers. George and Lal (2019), in turn, update the contributions of Tarus et al. (2018), investigating papers published between 2010 and 2019. The authors also discuss how ontology-based ERS can be used to address recommender systems traditional issues, such as cold start problem and rating sparsity.

Ashraf et al. (2021) directed their attention to investigate course recommendation systems. Through a comprehensive review, the study summarized the techniques and parameters used by this type of ERS. Additionally, a taxonomy of the factors taken into account in the course recommendation process was defined. Salazar et al. (2021), on the other hand, conducted a review on affectivity-based ERS. Authors presented a macro analysis, identifying the main authors and research trends, and summarized different recommender systems aspects, such as the techniques used in affectivity analysis, the source of affectivity data collection and how to model emotions.

Khanal et al. (2019) reviewed e-learning recommendation systems based on machine learning algorithms. A total of 10 papers from two scientific vehicles and published between 2016 and 2018 were examined. The study focal point was to investigate four categories of recommenders: those based on collaborative filtering, content-based filtering, knowledge and a hybrid strategy. The dimensions analyzed were the machine learning algorithms used, the recommenders' evaluation process, inputs and outputs characterization and recommenders' challenges addressed.

2.1 Related works gaps and contribution of this study

The studies presented in the previous section have a diversity of scope and dimensions of analysis, however, in general, they can be classified into two distinct groups. The first, focus on specific subjects of ERS field, such as similar methods of recommendations (George & Lal, 2019; Khanal et al., 2019; Salazar et al., 2021; Tarus et al., 2018) and same kind of recommendable resources (Ashraf et al., 2021; Nascimento et al., 2017). This type of research scrutinizes the particularities of the recommenders and highlights aspects that are difficult to be identified in reviews with a broader scope. Despite that, most of the reviews concentrate on analyses of recommenders' operational features and have limited discussion on crosswise issues, such as ERS evaluation and presentation approaches. Khanal et al. (2019), specifically, makes contributions regarding evaluation, but the analysis is limited to four types of recommender systems.

The second group is composed of wider scope reviews and include recommendation models based on a diversity of methods, inputs and outputs strategies (Drachsler et al., 2015; Ferreira et al., 2017; Klačnja-Milićević et al., 2015; Pinho et al., 2019; Rivera et al., 2018). Due to the very nature of systematic mappings, the research conducted by Ferreira et al. (2017) and Rivera et al. (2018) do not reach in depth some topics, for example, the data synthesized on the evaluations of the ERS are delimited to indicate only the methods used. Ferreira et al. (2017), in particular, aims to investigate only Brazilian recommendation systems, offering partial contributions to an understanding of the state of the art of the area. In Pinho et al. (2019) it is noted the same limitation of the systematic mappings. The review was reported with a restricted number of pages, making it difficult to detail the findings. On the other hand, Drachsler et al. (2015) and, Klačnja-Milićević et al. (2015) carried out comprehensive reviews that summarizes specific and macro dimensions of the area. However, the papers included in their reviews were published until 2014 and there is a gap on the *visto que* advances and trends in the field in the last 6 years.

Given the above, as far as the authors are aware, there is no wide scope secondary study that aggregate the research achievements on recommendation systems that support teaching and learning in recent years. Moreover, a review in this sense is necessary since personalization has become an important feature in the teaching and learning context and ERS are one of main tools to deal with different educational needs and preferences that affect individuals' learning process.

In order to widen the frontiers of knowledge in the field of research, this review aims to contribute to the area by presenting a detailed analysis of the following dimensions: how recommendations are produced and presented, how recommender systems are evaluated and what are the studies limitations and research opportunities. Specifically, to summarize the current knowledge, a SLR was conducted based on four research questions (Section 3.1). The review focused on papers published from 2015 to 2020 in scientific journals. A quality assessment was performed to select the most mature systems. The data found on the investigated topics are summarized and discussed in Section 4.

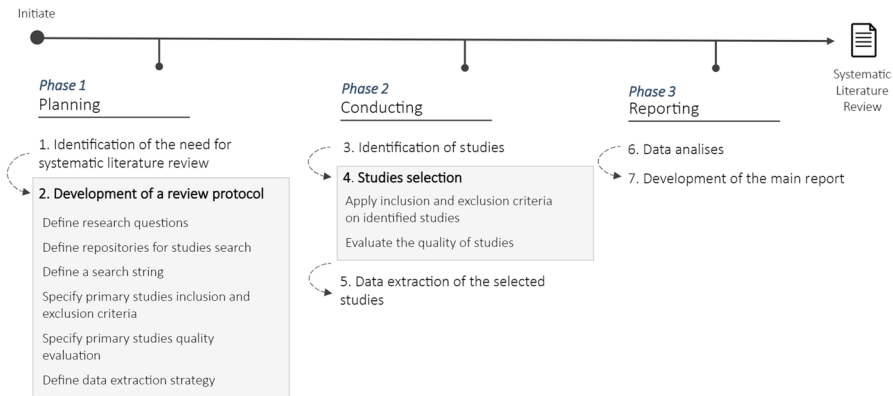


Fig. 1 Systematic literature review phases and activities

3 Methodology

This study is based on the SLR methodology for gathering evidences related to the research topic investigated. As stated by Kitchenham and Charters (2007) and Kitchenham et al. (2009), this method provides the means for aggregate evidences from current research prioritizing the impartiality and reproducibility of the review. Therefore, a SLR is based on a process that entails the development of a review protocol that guides the selection of relevant studies and the subsequent extraction of data for analysis.

Guidelines for SLR are widely described in literature and the method can be applied for gathering evidences in different domains, such as, medicine and social science (Khan et al., 2003; Pai et al., 2004; Petticrew & Roberts, 2006; Moher et al., 2015). Particularly for informatics in education area, Kitchenham and Charters (2007) guidelines have been reported as one of the main orientations (Dermeval et al., 2020). Their approach appears in several studies (Petri & Gresse von Wangenheim, 2017; Medeiros et al., 2019; Herpich et al., 2019) including mappings and reviews on ERS field (Rivera et al., 2018; Tarus et al., 2018).

As mentioned in Section 1, Kitchenham and Charters (2007) guidelines were used in the conducted SLR. They are based on three main stages: the first for planning the review, the second for conducting it and the last for the results report. Following these orientations, the review was structured in three phases with seven main activities distributed among them as depicted in Fig. 1.

The first was the planning phase. The identification of the need for a SLR about teaching and learning support recommenders and the development of the review protocol occurred on this stage. In activity 1, the search for SLR with the intended scope of this study was performed. The result did not return compatible papers with this review scope. Papers identified are described in Section 2. In activity 2, the review process was defined. The protocol was elaborated through rounds of discussion by the authors until consensus was reached. The activity 2 output were the research questions, search strategy, papers selection strategy and the data extraction method.

The next was the conducting phase. At this point, activities for relevant papers identification (activity 3) and selection (activities 4) were executed. In Activity 3, searches were carried out in seven repositories indicated by Dermeval et al. (2020) as relevant to the area of informatics in education. Authors applied the search string into these repositories search engines, however, due to the large number of returned research, the authors established the limit of 600 to 800 papers that would be analyzed. Thus, three repositories whose sum of search results was within the established limits were chosen. The list of potential repositories considered for this review and the selected ones is listed in Section 3.1. The search string used is also shown in Section 3.1.

In activity 4, studies were selected through two steps. In the first, inclusion and exclusion criteria were applied to each identified paper. Accepted papers had they quality assessed in the second step. Parsifal¹ was used to manage planning and conducting phase data. Parsifal is a web system, adhering to Kitchenham and Charters (2007) guidelines, that helps in SLR conduction. At the end of this step, relevant data were extracted (activity 5) and registered in a spreadsheet. Finally, in the reporting phase, the extracted data were analyzed in order to answer the SLR research questions (activity 6) and the results were recorded in this paper (activity 7).

3.1 Research question, search string and repositories

Teaching and learning support recommender systems have particularities of configuration, design and evaluation method. Therefore, the following research questions (Table 1) were elaborated in an effort to synthesize these knowledge as well as the main limitations and research opportunities in the field from the perspective of the most recent studies:

Regarding the search strategy, papers were selected from three digital repositories (Table 2). For the search, “Education” and “Recommender system” were defined as the keywords and synonyms were derived from them as secondary terms (Table 3). From these words, the following search string was elaborated:

("Education" OR "Educational" OR "E-learning" OR "Learning" OR "Learn")
AND ("Recommender system" OR "Recommender systems" OR "Recommendation system" OR "Recommendation systems" OR "Recommending system" OR "Recommending systems")

3.2 Inclusion and exclusion criteria

The first step for the selection of papers was performed through the application of objective criteria, thus a set of inclusion and exclusion criteria was defined. The approved papers formed a group that comprises the primary studies with potential

¹ <http://parsif.al/>

Table 1 SLR research questions

Id	Research question	Rationale
RQ1	How teaching and learning support recommender systems produce recommendations?	There is a variety of input parameters and techniques used for building recommenders for teaching and learning (Drachsler et al., 2015; Manouselis et al., 2014). They have been proposed in an attempt to find the best approach to match users' expectations and recommendations. Also, they include by design intrinsic limitations of each strategy (Garcia-Martinez & Hamou-Lhadj, 2013). Currently, studies are dispersed in literature and, as far as the authors are aware, there are not research synthesizing the knowledge about techniques and inputs used to tackle the field issues. Analyzing the last 6 years trends should clarify the current state of the art on how this kind of recommender has been designed and the latest trends in this sense
RQ2	How teaching and learning support recommender systems present recommendations?	Complementarily to RQ1, this research question leads to a broad analysis of the architecture of teaching and learning support recommender systems proposed by the scientific community. This research question adds to RQ1 and widens the insights of current state of the art on how ERS have been designed
RQ3	How are teaching and learning support recommender systems evaluated?	There are distinct methods aiming on measuring quality dimensions of an educational recommender and a previous study have suggested a growing awareness of the necessity for the elaboration of educational-focused evaluations in the ERS research field (Erdt et al., 2015). Analyzing the last 6 years trends on teaching and learning support recommender system evaluation topic will shed light on the current state of the art and will uncover insights of what evaluation goals have been prioritized as well as how recommenders' pedagogical effectiveness has been measured
RQ4	What are the limitations and research opportunities related to the teaching and learning support recommender systems field?	As the ERS research area have been developing through the past years, research limitations that hinder advancements in the field have been reported or can be observed in current studies. Also, an in-depth investigation can inform of under-explored topics that need further investigation taking into account their potential to contribute to the advancement of area. As long as the authors are aware, literature lacks on identifying current limitations and opportunities in teaching and learning support recommender system research. With this research question it is intended to reveal them from the perspective of the last 6 years scientific production. Answering this question can clarify the needs for future research in this topic

relevance for the scope of the SLR. Table 4 lists the defined criteria. In the description column of Table 4, the criteria are informed and in the id column they are identified with a code. The latter was defined appending an abbreviation of the respective kind of criteria (IC for Inclusion Criteria and EC for Exclusion Criteria) with an index following the sequence of the list. The Id is used for referencing its corresponding criterion in the rest of this document.

Since the focus of this review is on the analysis of recent ERS publications, only studies from the past 6 years (2015–2020) were screened (see IC1). Targeting mature recommender systems, only full papers from scientific journals that present the recommendation system evaluation were considered (see IC2, IC4 and IC7). Also, solely works written in English language were selected, because they are the most expressive in quantity and are within the reading ability of the authors (see IC3). Search string was verified on papers' title, abstract and keywords to ensure only studies related to the ERS field were screened (see IC5). The IC6, specifically, delimited the subject of selected papers and aligned it to the scope of the review. Additionally, it prevented the selection of secondary studies in process (e.g., others reviews or systematic mappings). Conversely, exclusion criteria were defined to clarify that papers contrasting with the inclusion criteria should be excluded from review (see EC1 to EC8). Finally, duplicate searches were marked and, when all criteria were met, only the latest was selected.

3.3 Quality evaluation

The second step in studies selection activity was the quality evaluation of the papers. A set of questions were defined with answers of different weights to estimate the quality of the studies. The objective of this phase was to filter researches with higher: (i) validity; (ii) details of the context and implications of the research; and (iii) description of the proposed recommenders. Research that detailed the configuration of the experiment and carried out an external validation of the ERS obtained higher weight in the quality assessment. Hence, the questions related to recommender evaluation (QA8 and QA9) ranged from 0 to 3, while the others, from 0 to 2. The questions and their respective answers are presented in Table 7 (see Appendix). Each paper evaluated had a total weight calculated according to Formula 1:

$$W_{\text{total}} = \sum_{n=1}^9 \frac{QA_i}{2} \quad (1)$$

Papers total weight range from 0 to 10. Only works that reached the minimum weight of 7 were accepted.

3.4 Screening process

Papers screening process occurred as shown in Fig. 2. Initially, three authors carried out the identification of the studies. In this activity, the search string was applied into search engines of the repositories along with the inclusion and exclusion criteria through filtering settings. Two searches were undertaken on the three repositories at distinct moments,

Table 2 Repositories considered for the SLR

Name	URL	No of papers	Selected?	Rationale
IEEEExplore	http://ieeexplore.ieee.org	310	yes	Within defined threshold
ACM Digital Library	http://dl.acm.org	60	yes	
Science Direct	http://www.sciencedirect.com	386	yes	
Springer Link	https://link.springer.com/	3.613	no	Several studies returned exceeding defined threshold
Engineering Village	https://www.engineeringvillage.com/	1.205	no	
Scopus	https://www.scopus.com/	1.760	no	
Web of Science	https://www.webofscience.com	1018	no	

Table 3 Keywords and their synonyms used in the search string

Keyword	Synonym
Educational	Education, E-learning, Learn, Learning
Recommender system	Recommender systems, Recommendation system, Recommendation systems, Recommending system, Recommending systems

one in November 2020 and another in January 2021. The second one was performed to ensure that all 2020 published papers in the repositories were counted. A number of 756 preliminary primary studies were returned and their metadata were registered in Parsifal.

Following the protocol, the selection activity was initiated. At the start, the duplicity verification feature of Parsifal was used. A total of 5 duplicate papers were returned and the oldest copies were ignored. Afterwards, papers were divided into groups and distributed among the authors. Inclusion and exclusion criteria were applied through titles and abstracts reading. In cases which were not possible to determine the eligibility of the papers based on these two fields, the body of text was read until it was possible to apply all criteria accurately. Finally, 41 studies remained for the next step. Once more, papers were divided into three groups and each set of works was evaluated by one author. Studies were read in full and weighted according to each quality assessment question. At any stage of this process, when questions arose, the authors defined a solution through consensus. As a final result of the selection activity, 16 papers were approved for data extraction.

3.5 Procedure for data analysis

Data from selected papers were extracted in a data collection form that registered general information and specific information. The general information extracted was: reviewer identification, date of data extraction and title, authors and origin of the paper. General information was used to manage the data extraction activity. The

Table 4 Inclusion and exclusion criteria of the SLR

Inclusion criteria		Exclusion criteria	
<i>Id</i>	<i>Description</i>	<i>Id</i>	<i>Description</i>
IC1	Paper published from 2015 to 2020	EC1	Paper published before 2015 and after 2020
IC2	Paper published in scientific journals	EC2	Paper is not published in scientific journals (ex. conference and workshops)
IC3	Paper must be in English language	EC3	Paper is not in English language
IC4	Paper must be a full paper	EC4	Paper is not a full paper
IC5	The search string must be in at least one of the following papers metadata: title, abstract or keywords	EC5	The search string cannot be found in at least one of the following papers metadata: title, abstract or keywords
IC6	Paper should focus on the development of a recommendation system and its application in the educational domain as a tool to support teaching or learning	EC6	Paper does not focus on the development of a recommendation system and its application in the educational domain as a tool to support teaching or learning
IC7	Paper must present the recommendation system evaluation	EC7	Paper does not present the recommendation system evaluation
		EC8	Paper is not a primary study

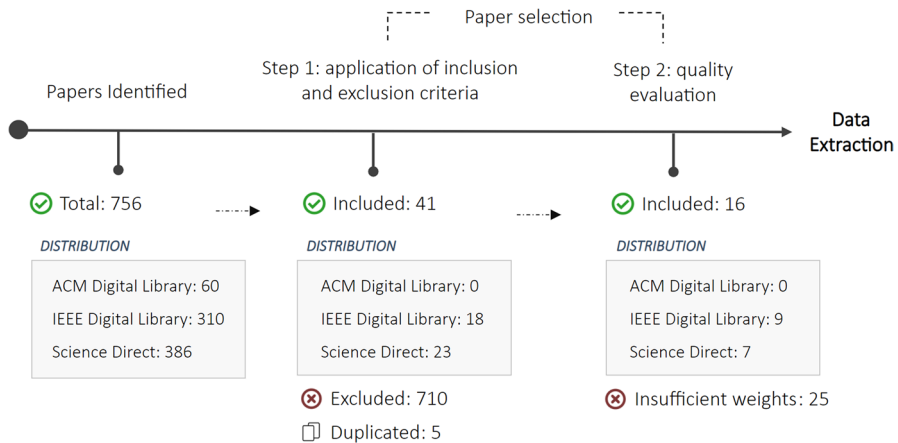


Fig. 2 Flow of papers search and selection

specific information was: recommendation approach, recommendation techniques, input parameters, data collection strategy, method for data collection, evaluation methodology, evaluation settings, evaluation approaches, evaluation metrics. This information was used to answer the research questions. Tabulated records were interpreted and a descriptive summary with the findings was prepared.

4 Results and discussion

In this section, the SLR results are presented. Firstly, an overview of the selected papers is introduced. Next, the finds are analyzed from the perspective of each research question in a respective subsection.

4.1 Selected papers overview

Each selected paper presents a distinct recommendation approach that advances the ERS field. Following, an overview of these studies is provided.

Sergis and Sampson (2016) present a recommendation system that supports educators' teaching practices through the selection of learning objects from educational repositories. It generates recommendations based on the level of instructors' proficiency on ICT Competences. In Tarus et al. (2017), the recommendations are targeted at students. The study proposes an e-learning resource recommender based on both user and item information mapped through ontologies.

Nafea et al. (2019) propose three recommendation approaches. They combine item ratings with student's learning styles for learning objects recommendation. Klačnja-Milićević et al. (2018) present a recommender of learning materials based on tags defined by the learners. The recommender is incorporated in Protus e-learning system.

In Wan and Niu (2016), a recommender based on mixed concept mapping and immunological algorithms is proposed. It produces sequences of learning objects for students. In a different approach, the same authors incorporate the self-organization theory into ERS. Wan and Niu (2018) deals with the notion of self-organizing learning objects. In this research, resources behave as individuals who can move towards learners. This movement results in recommendations and is triggered based on students' learning attributes and actions. Wan and Niu (2020), in turn, self-organization refers to the approach of students motivated by their learning needs. The authors propose an ERS that recommends self-organized cliques of learners and, based on these, recommend learning objects.

Zapata et al. (2015) developed a learning object recommendation strategy for teachers. The study describes a methodology based on collaborative methodology and voting aggregation strategies for the group recommendations. This approach is implemented in the Delphos recommender system. In a similar research line, Rahman and Abdullah (2018) show an ERS that recommends Google results tailored to students' academic profile. The proposed system classifies learners into groups and, according to the similarity of their members, indicates web pages related to shared interests.

Wu et al. (2015) propose a recommendation system for e-learning environments. In this study, complexity and uncertainties related to user profile data and learning activities is modeled through tree structures combined with fuzzy logic. Recommendations are produced from matches of these structures. Ismail et al. (2019) developed a recommender to support informal learning. It suggests Wikipedia content taking into account unstructured textual platform data and user behavior.

Huang et al. (2019) present a system for recommending optional courses. The system indications rely on the student's curriculum time constraints and similarity of academic performance between him and senior students. The time that individuals dedicate for learning is also a relevant factor in Nabizadeh et al. (2020). In this research, a learning path recommender that includes lessons and learning objects is proposed. Such a system estimates the learner's good performance score and, based on that, produces a learning path that satisfies their time constraints. The recommendation approach also provides indication of auxiliary resources for those who do not reach the estimated performance.

Fernandez-Garcia et al. (2020) deals with recommendations of disciplines through a dataset with few instances and sparse. The authors developed a model based on several techniques of data mining and machine learning to support students' decision in choosing subjects. Wu et al. (2020) create a recommender that captures students' mastery of a topic and produces a list of exercises with a level of difficulty adapted to them. Yanes et al. (2020) developed a recommendation system, based on different machine learning algorithms, that provides appropriate actions to assist teachers to improve the quality of teaching strategies.

4.2 How teaching and learning support recommender systems produce recommendations?

The process of generating recommendations is analyzed based on two axes. Underlying techniques of recommender systems are discussed first then input parameters are covered. Studies details are provided in Table 5.

4.2.1 Techniques approaches

Through selected papers analysis is observed that hybrid recommendation systems are predominant in selected papers. Such recommenders are characterized by computing predictions through a set of two or more algorithms in order to mitigate or avoid the limitations of pure recommendation systems (Isinkaye et al., 2015). From sixteen analyzed papers, thirteen ($p = 81,25\%$) are based on hybridization. This tendency seems to be related with the support that hybrid approach provides for development of recommender systems that must meet multiple educational needs of users. For example, Sergis and Sampson (2016) proposed a recommender based on two main techniques: fuzzy set to deal with uncertainty about teacher competence level and Collaborative Filtering (CF) to select learning objects based on neighbors who may have competences similarities. In Tarus et al. (2017) students and learning resources profiles are represented as ontologies. The system calculates predictions based on them and recommends learning items through a mechanism that applies collaborative filtering followed by a sequential pattern mining algorithm.

Moreover, the hybrid approach that combines CF and Content-Based Filtering (CBF), although a traditional technique (Bobadilla, Ortega, Hernando and Gutiérrez, 2013), it seems to be not popular in teaching and learning support recommender systems research. From the selected papers, only Nafea et al. (2019) has a proposal in this regard. Additionally, the extracted data indicates that a significant number of hybrid recommendation systems ($p = 53.85\%$, $n = 7$) have been built based on the combination of methods of treatment or representation of data, such as the use of ontologies and fuzzy sets, with methods to generate recommendation. For example, Wu et al. (2015) structure users profile data and learning activities through fuzzy trees. In such structures the values assigned to their nodes are represented by fuzzy sets. The fuzzy tree data model and users' ratings feed a tree structured data matching method and a CF algorithm for similarities calculation.

Collaborative filtering recommendation paradigm, in turn, plays an important role in research. Nearly a third of the studies ($p = 30.77\%$, $n = 4$) that propose hybrid recommenders includes a CF-based strategy. In fact, this is the most frequent pure technique on the research set. A total of 31.25% ($n = 5$) are based on a CF adapted version or combine it with other approaches. CBF-based recommenders, in contrast, have not shared the same popularity. This technique is an established recommendation approach that produces results based on the similarity between items known to the user and others recommendable items (Bobadilla et al., 2013). Only Nafea et al. (2019) propose a CBF-based recommendation system.

Also, CF user-based variant is widely used in analyzed research. In this version, predictions are calculated by similarity between users, as opposed to the item-based version where predictions are based on item similarities (Isinkaye et al., 2015). All CF-based recommendation systems identified, whether pure or combined with other techniques, use this variant.

The above finds seem to be related to the growing perception, in the education domain, of the relevance of a student-centered teaching and learning process (Krahenbuhl, 2016; Mccombs, 2013). Recommendation approaches that are based on

users' profile, such as interests, needs, and capabilities, naturally fit this notion and are more widely used than those based on other information such as the characteristics of the recommended items.

4.2.2 Input parameters approaches

In regard to the inputs consumed in the recommendation process, collected data shows that the main parameters are attributes related to users' educational profile. Examples are ICT competences (Sergis & Sampson, 2016); learning objectives (Wan & Niu, 2018; Wu et al., 2015), learning styles (Nafea et al., 2019), learning levels (Tarus et al., 2017) and different academic data (Yanes et al., 2020; Fernández-García et al., 2020). Only 25% ($n=4$) of the systems apply item-related information in the recommendation process. Furthermore, with the exception of the Nafea et al. (2019) CBF-based recommendation, the others are based on a combination of items and users' information. A complete list of the identified input parameters is provided in Table 5.

Academic information and learning styles, compared to others parameters, features highly on research. They appear, respectively, in 37.5% ($n=6$) and 31.25% ($n=5$) papers. Student's scores (Huang et al., 2019), academic background (Yanes et al., 2020), learning categories (Wu et al., 2015) and subjects taken (Fernández-García et al., 2020) are some of the academic data used. Learning styles, in turn, are predominantly based on Felder (1988) theory. Wan and Niu (2016), exceptionally, combine Felder (1988), Kolb et al. (2001) and Betoret (2007) to build a specific notion of learning styles. This is also used in two other researchers, carried out by the same authors, and has a questionnaire also developed by them (Wan & Niu, 2018, 2020).

Regarding the way inputs are captured, it was observed that explicit feedback is prioritized over others data collection strategies. In this approach, users have to directly provide the information that will be used in the process of preparing recommendations (Isinkaye et al., 2015). Half of analyzed studies are based only on explicit feedback. The use of graphical interface components (Klašnja-Milićević et al., 2018), questionnaires (Wan & Niu, 2016) and manual entry of datasets (Wu et al., 2020; Yanes et al., 2020) are the main methods identified.

Only 18.75% ($n=3$) ERS rely solely on gathering information through implicit feedback, that is, when inputs are inferred by the system (Isinkaye et al., 2015). This type of data collection appears to be more popular when applied with an explicit feedback method for enhancing the prediction tasks. Recommenders that combine both approaches occur in 31.25% ($n=5$) of the studies. Implicit data collection methods identified are user's data usage tracking, as access, browsing and rating history (Rahman & Abdullah, 2018; Sergis & Sampson, 2016; Wan & Niu, 2018), data extraction from another system (Ismail et al., 2019), users data session monitoring (Rahman & Abdullah, 2018) and data estimation (Nabizadeh et al., 2020).

The aforementioned results indicate that, in the context of the teaching and learning support recommender systems, the implicit collection of data has usually been explored in a complementary way to the explicit one. A possible rationale is

Table 5 Summary of ERS techniques and input parameters used in the selected papers

Research (citation)	Recommendation approach	Main techniques	Main input parameters	Type of data collection strategy	Method for data collection
Sergis and Sampson (2016)	Hybrid (collaborative filtering and fuzzy logic)	Neighbors users based on Euclidean distance and fuzzy sets	(i) ICT Competency; (ii) Rating (users' preferences)	Hybrid	(i) Collection of users' usage data; (ii) User defined
Tarus et al. (2017)	Hybrid (Collaborative Filtering, sequential pattern mining and knowledge representation)	Neighbors users based on cosine similarities, Generalized Sequential Pattern algorithm and student/learning resource domain ontologies	(i) Learning style; (ii) Learning level; (iii) Item attributes (iv) Rating (users' preferences)	Explicit	(i) Questionnaire; (ii) Online test; (iii) N/A; (iv) User defined
Nafea et al. (2019)	Collaborative filtering, content-based filtering and Hybrid (combining the last two approaches)	Neighbors users based on Pearson correlation, Neighbors items based on Pearson correlation and cosine similarity	(i) Learning style; (ii) Item attributes; (iii) Rating (users' preferences)	Explicit	(i) Questionnaire; (ii) Specialist defined
Wan and Niu (2018)	Self-organization based	Self-organization theory	(i) Learning style; (ii) Item attributes; (iii) learning objectives (iv) learners' behaviors	Hybrid	(i) Questionnaire; (ii) Specialist defined / students' feedback; (iii) N/A; (iv) Collection of users' usage data
Rahman and Abdullah (2018)	Group based	Groupuzation algorithm	(i) Academic information; (ii) learners' behaviors (iii) Contextual information	Implicit	(i) Learning management system records; (ii) Collection of users' usage data; (iii) Tracking changes in user academic records and behavior
Zapata et al. (2015)	Hybrid (techniques for group-based recommendation)	Collaborative methodology, voting aggregation strategies and meta-learning techniques	Rating (users' preferences)	Explicit	User defined

Table 5 (continued)

Research (citation)	Recommendation approach	Main techniques	Main input parameters	Type of data collection strategy	Method for data collection
Wan and Niu (2016)	Hybrid (knowledge representation and heuristic methods)	Mixed concept mapping and immune algorithm	(i) Learning styles; (ii) item attributes	Explicit	(i) Questionnaire; (ii) Specialist defined / Students' feedback
Klašnja- Milićević et al. (2018)	Hybrid (Social tagging and sequential pattern mining)	Most popular tags algorithms and weighted hybrid strategy	(i) Tags; (ii) Learners' behaviors	Hybrid	(i) User defined; (ii) Collection of users' usage data
Wu et al. (2015)	Hybrid (Knowledge representation, collaborative filtering and fuzzy logic)	fuzzy tree matching method, neighbors users based on cosine similarity and fuzzy set strategy	(i) Learning activities (ii) Learning objectives; (iii) Academic information; (iv) Rating (users' preferences)	Hybrid	(i) Collection of users' usage data (ii, iii, iv) User defined
Ismail et al. (2019)	Hybrid (Graph based and fuzzy logic)	Structural topical graph analysis algorithms and fuzzy set	(i) Learning interests; (ii) Thesaurus	Implicit	(i) Collection of users' usage data; (ii) Data extraction from another system
Huang et al. (2019)	Cross-user-domain collaborative filtering	Neighbors users based on cosine similarity	Academic information	Explicit	Input file/Dataset
Yanes et al. (2020)	Hybrid (machine learning algorithms)	One-vs-All, Binary Relevance, Classifier Chain, Label Powerset, Multil Label, K Nearest-Neighbors	Academic information	Explicit	Input file/Dataset
Wan and Niu (2020)	Hybrid (fuzzy logic, self-organization and sequential pattern mining)	Intuitionist fuzzy logic, self-organization theory and PrefixSpan algorithm	(i) Learning style (ii) Learning objectives (iii) Tags (iv) Academic information (v) Information from academic social relations	Hybrid	(i, ii) Questionnaire (iii) Extracted from m learners' learning profiles (iv, v) Extracted from e-learning platform records

Table 5 (continued)

Research (citation)	Recommendation approach	Main techniques	Main input parameters	Type of data collection strategy	Method for data collection
Fernandez- Garcia et al. (2020)	Hybrid (data mining and machine learning algorithms)	Encoding, Feature Engineering, Scaling, Resampling, Random Forest, Logistic Regression, Decision Tree, Support Vector Machine, K Nearest Neighbors, Multilayer Perceptron and Gradient Boosting Classifier	Academic information	Explicit	Input file/Dataset
Wu et al. (2020)	Hybrid (neural network techniques)	Recurrent Neural Networks and Deep Knowledge Tracing	Answers records	Explicit	Input file/Dataset
Nabizadeh et al. (2020)	Hybrid (graph based, clustering technique and matrix factorization)	Depth first search, k-means and matrix factorization	(i) Background knowledge; (ii) users' available time; (iii) Learning score	Implicit	(i) Collection of users' usage data; (ii, iii) estimated data

that the inference of information is noisy and less accurate (Isinkaye et al., 2015) and, therefore, the recommendations produced from it involve greater complexity to be adjusted to the users' expectations (Nichols, 1998). This aspect makes it difficult to apply the strategy in isolation and can be a factor that produces greater user dissatisfaction when compared to the disadvantage of the acquisition load of the explicit strategy inputs.

4.3 How teaching and learning support recommender systems present recommendations?

From the analyzed paper, two approaches for presenting recommendations are identified. The majority of the proposed ERS are based on a listing of ranked items according to a per-user prediction calculation ($p=87.5\%$, $n=14$). This strategy is applied in all cases where the supported task is to find good items that assist users in teaching and learning tasks (Ricci et al., 2015; Drachsler et al., 2015). The second one, is based on a learning pathway generation. In this case, recommendations are displayed through a series of linked items tied by some prerequisites. Only 2 recommenders use this approach. In them, the sequence is established by learning objects association attributes (Wan & Niu, 2016) and by a combination of prior knowledge of the user, the time he has available and a learning score (Nabizadeh et al., 2020). These ERS are associated with the item sequence recommendation task and are intended to guide users who wish to achieve a specific knowledge (Drachsler et al., 2015).

In a further examination, it is observed that more than a half (62.5% , $n=10$) do not present details of how recommendations list is presented to the end user. In Huang et al. (2019), for example, there is a vague description of a production of predicted scores for students and a list of the top- n optional courses and it is not specified how this list is displayed. This may be related to the fact that most of these recommenders do not report an integration into another system (e.g., learning management systems) or the purpose of making it available as a standalone tool (e.g., web or mobile recommendation system). The absence of such requirements mitigates the need for the development of a refined presentation interface. Only Tarus et al. (2017), Wan and Niu (2018) and Nafea et al. (2019) propose recommenders incorporated in an e-learning system and do not detail the way in which the results are exhibited. In the six papers that provide insights about recommendation presentation, a few of them (33.33% , $n=2$), have a graphical interface that explicitly seeks to capture the attention of the user who may be performing another task in the system. This approach highlights recommendations and is common in commercial systems (Beel, Langer and Genzmehr, 2013). In Rahman and Abdullah (2018), a panel entitled "recommendations for you" is used. In Ismail et al. (2019), a pop-up box with suggestions is displayed to the user. The other part of the studies exhibits organic recommendations, i.e., naturally arranged items for user interaction (Beel et al., 2013).

In Zapata et al. (2015), after the user defines some parameters, a list of recommended learning objects that are returned similarly to a search engine result. As for the aggregation methods, another item recommended by the system, only the strategy that fits better to the interests of the group is recommended. The result

is visualized through a five-star Likert scale that represents the users' consensus rating. In Klašnja-Milićević et al. (2018) and Wu et al. (2015), the recommenders' results are listed in the main area of the system. In Nabizadeh et al. (2020) the learning path occupies a panel on the screen and the items associated with it are displayed as the user progresses through the steps. The view of the auxiliary learning objects is not described in the paper. These three last recommenders do not include filtering settings and distance themselves from the archetype of a search engine.

Also, a significant number of researches are centralized on learning objects recommendations ($p=56.25\%$, $n=9$). Other researches recommendable items identified are learning activities (Wu et al., 2015), pedagogical actions (Yanes et al., 2020), web pages (Ismail et al., 2019; Rahman & Abdullah, 2018), exercises (Wu et al., 2020), aggregation methods (Zapata et al., 2015), lessons (Nabizadeh et al., 2020) and subjects (Fernández-García et al., 2020). None of the study relates the way of displaying results to the recommended item. This is a topic that needs further investigation to answer whether there are more appropriate ways to present specific types of items to the user.

4.4 How teaching and learning support recommender systems are evaluated?

In ERS, there are three main evaluation methodologies (Manouselis et al., 2013). One of them is the offline experiment, which is based on the use of pre-collected or simulated data to test recommenders' prediction quality (Shani & Gunawardana, 2010). User study is the second approach. It takes place in a controlled environment where information related to real interactions of users are collected (Shani & Gunawardana, 2010). This type of evaluation can be conducted, for example, through a questionnaire and A/B tests (Shani & Gunawardana, 2010). Finally, the online experiment, also called real life testing, is one in which recommenders are used under real conditions by the intended users (Shani & Gunawardana, 2010).

In view of these definitions, the analyzed researches comprise only user studies and offline experiments in reported experiments. Each of these methods were identified in 68.75% ($n=11$) papers respectively. Note that they are not exclusive for all cases and therefore the sum of the percentages is greater than 100%. For example, Klašnja-Milićević et al. (2018) and Nafea et al. (2019) assessed the quality of ERS predictions from datasets analysis and also asked users to use the systems to investigate their attractiveness. Both evaluation methods are carried out jointly in 37.5% ($n=6$) papers. When comparing with methods exclusive usage, each one is conducted at 31.25% ($n=5$). Therefore, the two methods seem to have a balanced popularity. Real-life tests, on the contrary, although they are the ones that best demonstrate the quality of a recommender (Shani & Gunawardana, 2010), are the most avoided, probably due to the high cost and complexity of execution.

An interesting finding concerns user study methods used in research. When associated with offline experiments, the user satisfaction assessment is the most common ($p=80\%$, $n=5$). Of these, only Nabizadeh et al. (2020) performed an in-depth

evaluation combining a satisfaction questionnaire with an experiment to verify the pedagogical effectiveness of their recommender. Wu et al. (2015), in particular, does not include a satisfaction survey. They conducted a qualitative investigation of user interactions and experiences.

Although questionnaires assist in identification of users' valuable information, it is sensitive to respondents' intentions and can be biased with erroneous answers (Shani & Gunawardana, 2010). Papers that present only user studies, in contrast, have a higher rate of experiments that results in direct evidence about the recommender's effectiveness in teaching and learning. All papers in this group have some investigation in this sense. Wan and Niu (2018), for example, verified whether the recommender influenced the academic score of students and their time to reach a learning objective. Rahman and Abdullah (2018) investigated whether the recommender impacted the time students took to complete a task.

Regarding the purpose of the evaluations, ten distinct research goals were identified. Through Fig. 3, it is observed that the occurrence of accuracy investigation excelled the others. Only 1 study did not carry out experiments in this regard. Different traditional metrics were identified for measuring the accuracy of recommenders. The Mean Absolute Error (MAE), in particular, has the higher frequency. Table 6 lists the main metrics identified.

The system attractiveness analysis, through the verification of user satisfaction, has the second highest occurrence. It is present in 62.5% ($n=10$) studies. The pedagogical effectiveness evaluation of the ERS has a reduced participation in the studies and occurs in only 37.5% ($n=6$). Experiments to examine recommendations diversity, user's profile elicitation accuracy, evolution process, user's experience and interactions, entropy, novelty and perceived usefulness and easiness were also identified, albeit to a lesser extent.

Also, 81.25% ($n=13$) papers presented experiments to achieve multiple purposes. For example, in Wan and Niu (2020) an evaluation is carried out to investigate recommenders' pedagogical effectiveness, student satisfaction, accuracy, diversity of recommendations and entropy. Only in Huang et al. (2019), Fernandez-Garcia et al. (2020) and Yanes et al. (2020) evaluated a single recommender system dimension.

The upper evidence suggests an engagement of the scientific community in demonstrating the quality of the recommender systems developed through multidimensional analysis. However, offline experiments and user studies, particularly those based on questionnaires, are mostly adopted and can lead to incomplete or biased interpretations. Thus, such data also signalize the need for a greater effort to conduct real life tests and experiments that lead to an understanding of the real impact of recommenders on the teaching and learning process. Researches that synthesize and discuss the empirical possibilities of evaluating the pedagogical effectiveness of ERS can help to increase the popularity of these experiments.

Through papers analysis is also find that the results of offline experiments are usually based on a greater amount of data compared to user studies. In this group, 63.64% ($n=7$) of evaluation datasets have records of more than 100 users. User studies, on the other hand, predominate sets of up to 100 participants in the experiments (72.72%,

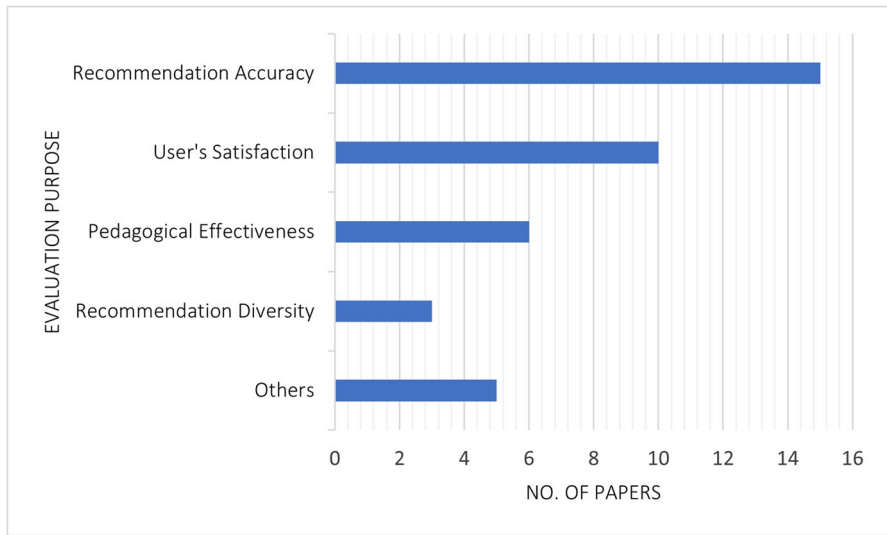


Fig. 3 Evaluation purpose of recommender systems in selected papers

$n=8$). In general, offline assessments that have smaller datasets are those that occur in association with a user study. This is because the data for both experiments usually come from the same subjects (Nafea et al., 2019; Tarus et al., 2017). The cost (e.g., time and money) related to surveying participants for the experiment is possibly a determining factor in defining appropriate samples.

Furthermore, it is also verified that the greater parcel of offline experiments has a 70/30% division approach for training and testing data. Nguyen et al. (2021) give some insights in this sense arguing that this is the most suitable ratio for training and validating machine learning models. Further details on recommendation systems evaluation approaches and metrics are presented in Table 6.

4.5 What are the limitations and research opportunities related to the teaching and learning support recommender systems field?

The main limitations observed in selected papers are presented below. They are based on articles' explicit statements and on authors' formulations. In this section, only those that are transverse to the majority of the studies are listed. Next, a set of research opportunities for future investigations are pointed out.

4.5.1 Research limitations

Research limitations are factors that hinders current progress in the ERS field. Knowing these factors can assist researchers to attempt coping with them on their study and mitigate the possibility of the area stagnation, that is, when new proposed

Table 6 Summary of ERS evaluation settings, approaches and metrics in selected papers

Research (citation)	Evaluation Methodology	Dataset Size / No. of Subjects	No. Recommendable Items	Highlights of Evaluation Approaches	Highlights of Evaluation Metrics
Sergis and Sampson (2016)	Offline	2005	96,196	Layered evaluation, Dataset split (70% training, 30% test) and comparison with collaborative filtering variations	Jaccard coefficient (for user's ICT profile elicitation accuracy) and RMSE
Tarus et al. (2017)	Offline and user study	50	240	Dataset split (70% training, 30% test), comparison with collaborative filtering variations and questionnaire survey	MAE, Precision, Recall and user's satisfaction level
Nafea et al. (2019)	Offline and user study	80	At least 300	Comparison between the proposed algorithms and questionnaire survey	MAE, RMSE and user's satisfaction level
Wan and Niu (2018)	User study	749	3043	A/B test, comparison with Instructors' suggestions and e-learning recommender systems based on genetic algorithm and Markov chain, and questionnaire survey	Average students' score, learning time, learning objects utilization, fitness function, learning objects' hit rate, learning objects' proportions marked with educational meaning tags, non centralized distribution of learning objects, proportion of new recommended items, user's satisfaction level and entropy (time to achieve a stable sequence of recommendations)
Rahman and Abdullah (2018)	User study	60	N/A	A/B test and questionnaire survey	Search time for educational materials, quantity of access to recommended items, user's satisfaction level, level of ease of use and the usefulness of recommender

Table 6 (continued)

Research (citation)	Evaluation Methodology	Dataset Size / No. of Subjects	No. Recommendable Items	Highlights of Evaluation Approaches	Highlights of Evaluation Metrics
Zapata et al. (2015)	Offline and user study	75 for offline experiment and 63 for questionnaire	N/A	Comparison between rating aggregation methods, analysis of appropriate aggregation method selection and questionnaire survey	RMSE, Top 1 Frequency, Average Ranking, Average Recommendation Error, Mean Reciprocal Rank, user's satisfaction level and level of ease of use and the usefulness of recommender
Wan and Niu (2016)	User study	250	235	A/B test, comparison with Instructors' suggestions and e-learning recommender systems based on genetic algorithm, particle swarm optimization and ant colony optimization algorithm, and questionnaire survey	Time spent on learning planning, quantity of recommended learning objects, average score, quantity of students who successfully passed the final exam, average recommendation rounds, Average total recommended learning objects for each learner among all recommendation rounds, average time of learning objects recommendation, average evolution time and user's satisfaction level
Klašnja- Milićević et al. (2018)	Offline and user study	120 for offline experiment and 65 for questionnaire	62	Dataset split (80% training, 20% test), Comparison between tag recommendation methods and questionnaire survey	Precision, Recall, user's satisfaction level and level of ease of use and the usefulness
Wu et al. (2015)	Offline and user study	2213 for offline experiment and 5 for case study	N/A	Dataset split (20%/40%/50% test set), Compared with recommendation approach for e-learning recommender systems proposed by Bobadilla et al. (2009) and case study	MAE

Table 6 (continued)

Research (citation)	Evaluation Methodology	Dataset Size / No. of Subjects	No. Recommendable Items	Highlights of Evaluation Approaches	Highlights of Evaluation Metrics
Ismail et al. (2019)	User study	100 for comparison and 80 for questionnaire	N/A	A/B test, comparison between the proposed recommenders with a control group and a baseline approach and questionnaire survey	Mean Average Precision, knowledge level, visited articles, perceived relevance and user's satisfaction level
Huang et al. (2019)	Offline	1166	782	Dataset split (Training and testing dataset divided according to semesters) and comparison of recommender predicted optional courses and the ground-truth optional courses that student has enrolled on	Average hit rate, average accuracy
Yanes et al. (2020)	Offline	N/A	9	Dataset split (70% training, 30% test) and comparison of different machine learning algorithms	Precision, Recall, F1-measure, Hamming Loss
Wan and Niu (2020)	User study	1119	2386	A/B test, comparison with Instructors' suggestions and with variants of the proposed algorithm, and questionnaire survey	Students' average scores, proportion of students who passed the exam, average learning time, proportion of the resources that learners had visited out of the total number of resources, matching degree between learning objects and learners, diversity of learning objects attributes, proportion of learner's tags, user's satisfaction level, level of usefulness and entropy (time to achieve a stable sequence of recommendations)

Table 6 (continued)

Research (citation)	Evaluation Methodology	Dataset Size / No. of Subjects	No. Recommendable Items	Highlights of Evaluation Approaches	Highlights of Evaluation Metrics
Fernandez-Garcia et al. (2020)	Offline	323	45	Sequential chain of steps with dataset transformations	Accuracy, F1-score
Wu et al. (2020)	Offline	5126	19,136	Dataset split (70% Training, 10% validation, 20% test) and comparison with user-based and item-based collaborative filter, content-based filtering, hybrid recommendation model based on deep collaborative filtering and a knowledge-graph embedding based collaborative filtering	Accuracy, novelty and diversity
Nabizadeh et al. (2020)	Offline and user study	205 for offline experiment and 32 for user study	90 for offline experiment and 59 for user study	Dataset split (Training and testing dataset divided according to a defined observed and unobserved LO classification), algorithms comparison, A/B test, control and experimental groups	AE, number of correctly completed learning objects/ lessons by users, time that users spend to get their goals and user's satisfaction level

recommenders does not truly generate better outcomes than the baselines (Anelli et al., 2021; Dacrema et al., 2021). As a result of this SLR, research limitations were identified in three strands that are presented below.

Reproducibility restriction The majority of the papers report a specifically collected dataset to evaluate the proposed ERS. The main reason for this is the scarcity of public datasets suited to the research's needs, as highlighted by some authors (Nabizadeh et al., 2020; Tarus et al., 2017; Wan & Niu, 2018; Wu et al., 2015; Yanes et al., 2020). Such approach restricts the feasibility of experiment reproduction and makes it difficult to compare recommenders. In fact, this is an old issue in the ERS field. Verbert et al. (2011) observed, in the beginning of the last decade, the necessity to improve reproducibility and comparison on ERS in order to provide stronger conclusions about their validity and generalizability. Although there was an effort in this direction in the following years based on a broad educational dataset sharing, currently, most of the known ones (Çano & Morisio, 2015; Drachsler et al., 2015) are retired, and the remaining, proved not to be sufficient to meet current research demands. Of the analyzed studies, only Wu et al. (2020) use public educational datasets.

Due to the fact that datasets sharing play an important role for recommenders' model reproduction and comparison in the same conditions, this finding highlights the need of a research community effort for the creation of means to supply this need (e.g., development of public repositories) in order to mitigate current reproducibility limitation.

Dataset size / No of subjects As can be observed on Table 6, a few experimental results are based on a large amount of data. Only five studies have information from 1000 or more users. In particular, the offline evaluation conducted by Wu et al. (2015), despite having an extensive dataset, uses MovieLens records and is not based on real information related to teaching and learning. Another limitation concerns where data comes from, it is usually from a single origin (e.g., class of a college).

Although experiments based on small datasets can reveal the relevance of an ERS, an evaluation based on a large-scale dataset should provide stronger conclusions on recommendation effectiveness (Verbert et al., 2011). Experiments based on larger and more diverse data (e.g., users from different areas and domains) would contribute to most generalizable results. On another hand, scarcity of public dataset may be impairing the quantity and diversity of data used on scientific experiments in the ERS field. As reported by Nabizadeh et al. (2020), the increasement of the size of the experiment is costly in different aspects. If more public dataset were available, researchers would be more likely to find the ones that could be aligned to their needs and, naturally, increasing the size of their experiment. In this sense, they could be favored by reducing data acquisition difficulty and cost. Furthermore, the scientific community would access users' data out of their surrounding context and could base their experiments on diversified data.

Lack of in-depth investigation of the impact of known issues in the recommendation system field Cold start, overspecialization and sparsity are some known challenges in the field of recommender systems (Khusro et al., 2016). They are mainly related to a reduced and unequally distributed number of users' feedback or item description used for generating recommendations (Kunaver & Požrl, 2017). These issues also permeate the ERS Field. For instance, in Cechinel et al. (2011) is reported that on a sample of more than 6000 learning objects from Merlot repository was observed a reduced number of users ratings over items. Cechinel et al. (2013), in turn, observed, in a dataset from the same repository, a pattern of few users rating several resources while the vast number of them rating 5 or less. Since such issues directly impact the quality of recommendations, teaching and learning support recommenders should be evaluated considering such issues to clarify in which extent they can be effective in real life situations. Conversely, in this SLR, we detected an expressive number of papers (43.75%, $n=7$) that do not analyze or discuss how the recommenders behave or handle, at least partially, these issues. Studies that rely on experiments to examine such aspects would elucidate more details of the quality of the proposed systems.

4.5.2 Research opportunities

From the analyzed papers, a set of research opportunities were identified. They are based on gaps related to the subjects explored through the research questions of this SLR. The identified opportunities provide insights of under-explored topics that need further investigation taking into account their potential to contribute to the advancement of the ERS field. Research opportunities were identified in three strands that are presented below.

Study of the potential of overlooked user's attributes The papers examined present ERS based on a variety of inputs. Preferences, prior knowledge, learning style, and learning objectives are some examples (Table 5 has the complete list). Actually, as reported by Chen and Wang (2021), this is aligned with a current research trend of investigating the relationships between individual differences and personalized learning. Nevertheless, one evidence that rises from this SLR also confirms that "some essential individual differences are neglected in existing works" (Chen & Wang, 2021). The papers sample suggests a lack of studies that incorporate, in recommendation model, others notably relevant information, such as emotional state and cultural context of students (Maravanyika & Dlodlo, 2018; Salazar et al., 2021; Yanes et al., 2020). This indicates that further investigation is needed in order to clarify the true contributions and existing complexities of collect, measure and apply these other parameters. In this sense, an open research opportunity refers to the investigation of these other users' attributes in order to explore the impact of such characteristics on the quality of ERS results.

Increase studies on the application of ERS in informal learning situations Informal learning refers to a type of learning that, typically, occurs out of an education institution (Pöntinen et al., 2017). In it, learners do not follow a structured

curriculum or have a domain expert to guide him (Pöntinen et al., 2017; Santos & Ali, 2012). Such aspects influence how ERS can support users. For instance, in informal settings, content can come from multiple providers, as a consequence, it can be delivered without taking into account a proper pedagogical sequence. ERS targeting this scenario, in turn, should concentrate on organizing and sequencing recommendations guiding users' learning process (Drachsler et al., 2009).

Although literature highlight the existence of significative differences on the design of educational recommenders that involves formal or informal learning circumstance (Drachsler et al., 2009; Okoye et al, 2012; Manouselis et al., 2013; Harrathi & Braham, 2021), through this SLR was observed that current studies tend to not be explicit in reporting this characteristic. This scenario makes it difficult to obtain a clear landscape of the current field situation in this dimension. Nonetheless, through the characteristics of the proposed ERS, it was observed that current research seems to be concentrated on the formal learning context. This is because recommenders from analyzed papers usually use data that are maintained by institutional learning systems. Moreover, recommendations, predominantly, do not provide a pedagogical sequencing to support self-directed and self-paced learning (e.g., recommendations that build a learning path to lead to specific knowledge). Conversely, informal learning has increasingly gained attention of the scientific community with the emergence of the coronavirus pandemic (Watkins & Marsick, 2020).

In view of this, the lack of studies of ERS targeting informal learning settings open a research opportunity. Specifically, further investigation focused on the design and evaluation of recommenders that take into consideration different contexts (ex. location or used device) and that guide users through a learning sequence to achieve a specific knowledge would figure prominently in this context considering the less structured format informal learning circumstances has in terms of learning objectives and learning support.

Studies on the development of multidimensional evaluation frameworks Evidence from this study shows that the main purpose of ERS evaluation has been to assess recommender's accuracy and users' satisfaction (Section 4.4). This result, connected with Erdt et al. (2015) reveals a two decade of evaluation predominantly based on these two goals. Even though others evaluation purposes had a reduced participation in research, they are also critical for measuring the success of ERS. Moubayed et al. (2018), for example, highlights two e-learning systems evaluation aspects, one is concerned with how to properly evaluate the student performance, the other refers to measuring learners' learning gains through systems usage. Tahereh et al. (2013) identifies that stakeholder and indicators associated with technological quality are relevant to consider in educational system assessment. From the perspective of recommender systems field, there are also important aspects to be analyzed in the context of its application in the educational domain such as novelty and diversity (Pu et al., 2011; Cremonesi et al., 2013; Erdt et al., 2015).

Upon this context, it is noted that, although evaluating recommender's accuracy and users' satisfaction give insights about the value of the ERS, they are not sufficient to fully indicate the quality of the system in supporting the learning process. Other different factors reported in literature are relevant to take in consideration. However, to the best of our knowledge, there is no framework that identifies and organizes these factors to be considered in an ERS evaluation, leading to difficulties for the scientific community to be aware of them and incorporate them in studies.

Because the evaluation of ERS needs to be a joint effort between computer scientists and experts from other domains (Erdt et al., 2015), further investigation should be carried out seeking the development of a multidimensional evaluation framework that encompass evaluation requirements based on a multidisciplinary perspective. Such studies would clarify the different dimensions that have the potential to contribute to better ERS evaluation and could even identify which one should be prioritized to truly assess learning impact with reduced cost.

5 Conclusion

In recent years, there has been an extensive scientific effort to develop recommenders that meet different educational needs; however, research is dispersed in literature and there is no recent study that encompasses the current scientific efforts in the field.

Given this context, this paper presents an SLR that aims to analyze and synthesize the main trends, limitations and research opportunities related to the teaching and learning support recommender systems area. Specifically, this study contributes to the field providing a summary and an analysis of the current available information about the teaching and learning support recommender systems topic in four dimensions: (i) how the recommendations are produced (ii) how the recommendations are presented to the users (iii) how the recommender systems are evaluated and (iv) what are the limitations and opportunities for research in the area.

Evidences are based on primary studies published from 2015 to 2020 from three repositories. Through this review, it is provided an overarching perspective of current evidence-based practice in ERS in order to support practitioners and researchers for implementation and future research directions. Also, research limitations and opportunities are summarized in light of current studies.

The findings, in terms of current trends, shows that hybrid techniques are the most used in teaching and learning support recommender systems field. Furthermore, it is noted that approaches that naturally fit a user centered design (e.g., techniques that allow to represent students' educational constraints) have been prioritized over that based on other aspects, like item characteristics (e.g., CBF Technique). Results show that these approaches have been recognized as the main means to support users with recommendations in their teaching and learning process and provide directions for practitioners and researchers who seek to base their activities and investigations on evidence from current studies. On the other

hand, this study also reveals that highly featured techniques in the major topic of general recommender systems, such as the bandit-based and the deep learning ones (Barraza-Urbina & Glowacka, 2020; Zhang et al., 2020), have been under-explored, implying a mismatch between the areas. Therefore, the result of this systematic review indicates that a greater scientific effort should be employed to investigate the potential of these uncovered approaches.

With respect to recommendation presentation, the organic display is the most used strategy. However, most of the researches have the tendency to not show details of the used approach making it difficult to understand the state of the art of this dimension. Furthermore, among other results, it is observed that the majority of the ERS evaluation are based on the accuracy of recommenders and user's satisfaction analysis. Such a find open research opportunity scientific community for the development of multidimensional evaluation frameworks that effectively support the verification of the impact of recommendations on the teaching and learning process.

Lastly, the limitations identified indicate that difficulties related to obtaining data to carry out evaluations of ERS is a reality that extends for more than a decade (Verbert et al., 2011) and call for scientific community attention for the treatment of this situation. Likewise, the lack of in-depth investigation of the impact of known issues in the recommendation system field, another limitation identified, points to the importance of aspects that must be considered in the design and evaluation of these systems in order to provide a better elucidation of their potential application in a real scenario.

With regard to research limitations and opportunities, some of this study findings indicate the need for a greater effort in the conduction of evaluations that provide direct evidence of the systems pedagogical effectiveness and the development of a multidimensional evaluation frameworks for ERS is suggested as a research opportunity. Also, it was observed a scarcity of public dataset usage on current studies that leads to limitation in terms of reproducibility and comparison of recommenders. This seems to be related to a restricted number of public datasets currently available, and such aspect can also be affecting the size of experiments conducted by researchers.

In terms of limitations of this study, the first refers to the number of datasources used for paper selection. Only the repositories mentioned in Section 3.1 were considered. Thus, the scope of this work is restricted to evidence from publications indexed by these platforms. Furthermore, only publications written in English were examined, thus, results of papers written in other languages are beyond the scope of this work. Also, the research limitations and opportunities presented on Section 4.5 were identified based on the extracted data used to answer this SLR research questions, therefore they are limited to their scope. As a consequence, limitations and opportunities of the ERS field that surpass this context were not identified nor discussed in this study. Finally, the SLR was directed to papers published in scientific journals and, due to this, the results obtained do not reflect the state of the area from the perspective of conference publications. In future research, it is intended to address such limitations.

Appendix

Table 7

Table 7 Quality evaluation questions and answers

Id	Questions	Available Answers (Weight)
QA1	Does the paper clearly present the research contributions?	<p>The paper explicitly lists or describes the research contributions. They are clearly connected with the presented results. (2pts)</p> <p>The paper provides a general description of research contributions. They are clearly connected with the presented results. (1pt)</p> <p>The paper provides a general description of research contributions. They are not clearly connected with the presented results. (0.5pts)</p> <p>The paper does not clearly provide a research contribution. If presented, the contributions are not clearly connected with the results of the study. (0)</p>
QA2	Does the paper clearly present how research differs from other related works?	<p>The research is compared in detail with related works. Authors provide the strengths and/or weaknesses of each related work and position their research granularly, stating contributions explicitly. (2pts)</p> <p>The research is compared in detail with related works. Authors provide a general description of the related works and they position their research stating contribution explicitly. (1.0pts)</p> <p>The paper provides a general description of related works. A brief introduction to each study or groups of it is introduced without identifying strengths and/or weaknesses. The authors explain how their research stands out without explicitly comparing it with the related works. (0.5pts)</p> <p>The paper does not position the research in relation to other works in the area. The unique contribution of the study is general or not presented explicitly. (0)</p>
QA3	Does the paper clearly present the limitations of the research?	<p>The paper lists or describes the limitations of the study. If the evaluation produces any results that are difficult to explain, the challenges are detailed presented. (2pts)</p> <p>The paper presents the study limitations with a general description. If the evaluation produces any results that are difficult to explain, the paper describes the challenges in a general way. (1.0pts)</p> <p>The paper does not explicitly present or list the limitations of the study. Nonetheless, the paper presents some research-related challenges when discussing the results of the experiments or in the conclusion. (0.5pts)</p> <p>The paper does not present the limitations of the study. If the evaluation presents any results that are difficult to explain, the paper does not describe the challenges. (0)</p>

Table 7 (continued)

Id	Questions	Available Answers (Weight)
QA4	Does the paper clearly present directions for future work?	<p>The paper explicitly lists or describes directions for future work. These are based on experiment results or limitations explicitly discussed. (2pts)</p> <p>The paper explicitly lists or describes directions for future work. Yet, such directions are not linked with experiment results or the paper does not present the motivations or foundations. (1.0pts)</p> <p>The paper presents directions for future work in general. (0.5pts)</p> <p>The paper does not present directions for future work. (0)</p>
QA5	Does the paper clearly present the inputs for the recommendation generation process?	<p>The paper explicitly presents the recommender system input parameters and the way it is collected. When the recommender does not produce recommendations based on a user profile, the authors describe the input elements used. When such information cannot be understood directly, the authors describe in detail each element that composes it and how these elements are obtained (2pts)</p> <p>The paper presents the recommender input parameter through a general description or it is noticed a partial omission of some information, for example, through the use of “etc.” at the end of a list. (1.0pts)</p> <p>The paper does not present the recommender input parameter. (0)</p>
QA6	Does the paper clearly present the techniques/methods for the recommendation generation process?	<p>The paper describes the techniques and methods used for the recommendation generation process. They are presented in a continuous flow, beginning with an overview followed by specific details of the proposal. Authors may or may not use one or more illustrations to present the iterations and how the proposed ERS functions in detail. (2pts)</p> <p>The paper describes the techniques and methods used for the recommendation generation process. Yet, these elements are not presented in a continuous flow, beginning with an overview followed by specific details of the proposal. Authors may or may not use one or more illustrations to present the iterations and how the proposed ERS functions in detail. (1.0pts)</p> <p>The paper presents the techniques and methods used for the recommendation generation process in general. The presentation does not have a continuous flow that begins with an overview and then the specific details of the proposal. The author does not use illustrations to present the iterations and the functioning of the proposed recommender or uses illustrations that lack important components crucial for their understanding. (0.5pts)</p> <p>The paper does not present the techniques and methods used in the elaboration of the recommender (0)</p>

Table 7 (continued)

Id	Questions	Available Answers (Weight)
QA7	Does the paper clearly present the target audience of the recommendations?	<p>The paper explicitly presents the recommender target audience, contextualizes how the research addresses or minimizes a specific issue of them and, whenever possible, provides specific characteristics, such as their age range, education level (e.g., students) or teaching level (e.g., professors). (2pts)</p> <p>The paper explicitly presents recommender target audience and contextualizes how their research resolves or minimizes a specific problem of this audience. However, they do not present the specific characteristics of this audience. (1,0pts)</p> <p>The paper presents a general description of the recommender target audience. (0,5pts)</p> <p>The paper does not specify the recommender target audience (e.g., individuals that use the system are identified only as users). (0)</p>
QA8	Does the paper clearly present the setting of the experiment?	<p>The paper explicitly describes the settings of the experiment. All of these main elements are listed: 1) Number of users; 2) Number of items to recommend; 3) Kind of recommended items; 4) Source of the data used. (3pts)</p> <p>The paper explicitly describes the settings of the experiment. Still, one of the following elements is not explained: 1) Number of users; 2) Number of items to be recommended; 3) Kind of recommended items; 4) Source of the data used. (1,5pts)</p> <p>The paper provides a general description of the experiment settings. Yet, it is noted that there is a little detail regarding the experiment configuration and more than one of the following key elements is not explained: 1) Number of users; 2) Number of items to be recommended; 3) Kind of recommended items;</p> <p>4) Source of the data used. (0,75pts)</p> <p>Authors do not describe the experiment settings. (0)</p>

Table 7 (continued)

Id	Questions	Available Answers (Weight)
QA9	Does the paper clearly present how the recommender was evaluated?	<p>The paper describes the evaluation steps and if the experiment is based on methodologies and/or theories. It also explicitly presents how the evaluation carried out validates that the research offers direct contributions to the educational development of the user or those who relate to him. It was conducted an external validation through an online evaluation or an experiment based on control/experimental groups. (3pts)</p> <p>The paper describes the evaluation steps and if the experiment is based on methodologies and/or theories. It also explicitly presents how the evaluation carried out validates that the research offers direct contributions to the educational development of the user or those who relate to him. It was conducted an internal validation through an offline evaluation followed or not by questionnaire-based user study. (2,5pts)</p> <p>The paper describes the evaluation steps, but it does not explicitly justify the used approach. It also explicitly presents how the evaluation carried out validates that the research offers direct contributions to the educational development of the user or those who relate to him. It was conducted either an internal or external validation. (2,0pts)</p> <p>The paper describes the evaluation steps and if the experiment is based on methodologies and/or theories. It does not explicitly present how the evaluation carried out validates that the research offers direct contributions to the educational development of the user or those who relate to him. (1,0pts)</p> <p>The paper describes evaluation steps; however, they do not explicitly justify the experiment approach. It also does not explicitly present how the evaluation carried out validates that the research offers direct contributions to the educational development of the user or those who relate to him. (0,75pts)</p> <p>The paper generally presents the evaluation and does not explicitly justify the experiment approach. It also does not explicitly present how the evaluation carried out validates that the research offers direct contributions to the educational development of the user or those who relate to him. (0,35pts)</p> <p>The paper does not present the recommender evaluation (0)</p>

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10639-022-11341-9>.

Author contribution **Felipe Leite da Silva:** Conceptualization, Methodology approach, Data curation, Writing – original draft. **Bruna Kin Slodkowski:** Data curation, Writing – original draft. **Ketia Kellen Araújo da Silva:** Data curation, Writing – original draft. **Sílvio César Cazella:** Supervision and Monitoring of the research; Writing – review & editing.

Data availability statement The datasets generated during the current study correspond to the papers identified through the systematic literature review and the quality evaluation results (refer to Section 3.4 in paper). They are available from the corresponding author on reasonable request.

Informed consent This research does not involve human participation as research subject, therefore research subject consent does not apply.

Authors consent with the content presented in the submitted manuscript.

Financial and non-financial interests The authors have no relevant financial or non-financial interests to disclose.

Research involving human participants and/or animals This research does not involve an experiment with human or animal participation.

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

- Anelli, V. W., Bellogín, A., Di Noia, T., & Pomo, C. (2021). Revisioning the comparison between neural collaborative filtering and matrix factorization. *Proceedings of the Fifteenth ACM Conference on Recommender Systems*, 521–529. <https://doi.org/10.1145/3460231.3475944>
- Ashraf, E., Manickam, S., & Karuppayah, S. (2021). A comprehensive review of course recommender systems in e-learning. *Journal of Educators Online*, 18, 23–35. https://www.thejeo.com/archive/2021_18_1/ashraf_manickam_karuppayah
- Barraza-Urbina, A., & Glowacka, D. (2020). Introduction to Bandits in Recommender Systems. *Proceedings of the Fourteenth ACM Conference on Recommender Systems*, 748–750. <https://doi.org/10.1145/3383313.3411547>
- Becker, F. (1993). *Teacher epistemology: The daily life of the school* (1st ed.). Editora Vozes.
- Beel, J., Langer, S., & Genzmehr, M. (2013). Sponsored vs. Organic (Research Paper) Recommendations and the Impact of Labeling. In T. Aalberg, C. Papatheodorou, M. Dobrev, G. Tsakonas, & C. J. Farrugia (Eds.), *Research and Advanced Technology for Digital Libraries* (Vol. 8092, pp. 391–395). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40501-3_44
- Betoret, F. (2007). The influence of students' and teachers' thinking styles on student course satisfaction and on their learning process. *Educational Psychology*, 27(2), 219–234. <https://doi.org/10.1080/01443410601066701>
- Bobadilla, J., Serradilla, F., & Hernando, A. (2009). Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems*, 22(4), 261–265. <https://doi.org/10.1016/j.knosys.2009.01.008>
- Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, 46, 109–132. <https://doi.org/10.1016/j.knosys.2013.03.012>
- Buder, J., & Schwind, C. (2012). Learning with personalized recommender systems: A psychological view. *Computers in Human Behavior*, 28(1), 207–216. <https://doi.org/10.1016/j.chb.2011.09.002>

- Canó, E., & Morisio, M. (2015). Characterization of public datasets for Recommender Systems. (2015 IEEE 1st) *International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI)*, 249–257. <https://doi.org/10.1109/RTSI.2015.7325106>
- Cazella, S. C., Behar, P. A., Schneider, D., Silva, K.K., & Freitas, R. (2014). Developing a learning objects recommender system based on competences to education: Experience report. *New Perspectives in Information Systems and Technologies*, 2, 217–226. https://doi.org/10.1007/978-3-319-05948-8_21
- Cechinel, C., Sánchez-Alonso, S., & García-Barriocanal, E. (2011). Statistical profiles of highly-rated learning objects. *Computers & Education*, 57(1), 1255–1269. <https://doi.org/10.1016/j.compedu.2011.01.012>
- Cechinel, C., Sicilia, M. -Á., Sánchez-Alonso, S., & García-Barriocanal, E. (2013). Evaluating collaborative filtering recommendations inside large learning object repositories. *Information Processing & Management*, 49(1), 34–50. <https://doi.org/10.1016/j.ipm.2012.07.004>
- Chen, S. Y., & Wang, J.-H. (2021). Individual differences and personalized learning: A review and appraisal. *Universal Access in the Information Society*, 20(4), 833–849. <https://doi.org/10.1007/s10209-020-00753-4>
- Cremonesi, P., Garzotto, F., & Turrin, R. (2013). User-centric vs. system-centric evaluation of recommender systems. In P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, & M. Winckler (Eds.), *Human-Computer Interaction – INTERACT 2013*, 334–351. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40477-1_21
- Dacrema, M. F., Boglio, S., Cremonesi, P., & Jannach, D. (2021). A troubling analysis of reproducibility and progress in recommender systems research. *ACM Transactions on Information Systems*, 39(2), 1–49. <https://doi.org/10.1145/3434185>
- Dermeval, D., Coelho, J.A.P.d.M., & Bittencourt, I.I. (2020). Mapeamento Sistemático e Revisão Sistemática da Literatura em Informática na Educação. *Metodologia de Pesquisa Científica em Informática na Educação: Abordagem Quantitativa*. Porto Alegre. <https://jodi-ojs-tdl.tdl.org/jodi/article/view/442>
- Drachslar, H., Hummel, H. G. K., & Koper, R. (2009). Identifying the goal, user model and conditions of recommender systems for formal and informal learning. *Journal of Digital Information*, 10(2), 1–17. <https://jodi-ojs-tdl.tdl.org/jodi/article/view/442>
- Drachslar, H., Verbert, K., Santos, O. C., & Manouselis, N. (2015). Panorama of Recommender Systems to Support Learning. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 421–451). Springer. https://doi.org/10.1007/978-1-4899-7637-6_12
- Erdt, M., Fernández, 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>
- Felder, R. (1988). Learning and teaching styles in engineering education. *Journal of Engineering Education*, 78, 674–681. Washington.
- Fernandez-García, A. J., Rodríguez-Echeverría, R., Preciado, J. C., Manzano, J. M. C., & Sanchez-Figueroa, F. (2020). Creating a recommender system to support higher education students in the subject enrollment decision. *IEEE Access*, 8, 189069–189088. <https://doi.org/10.1109/ACCESS.2020.3031572>
- Ferreira, V., Vasconcelos, G., & França, R. (2017). Mapeamento Sistemático sobre Sistemas de Recomendações Educacionais. *Proceedings of the XXVIII Brazilian Symposium on Computers in Education*, 253–262. <https://doi.org/10.5753/cbie.sbie.2017.253>
- García-Martínez, S., & Hamou-Lhadj, A. (2013). Educational recommender systems: A pedagogical-focused perspective. *Multimedia Services in Intelligent Environments. Smart Innovation, Systems and Technologies*, 25, 113–124. https://doi.org/10.1007/978-3-319-00375-7_8
- George, G., & Lal, A. M. (2019). Review of ontology-based recommender systems in e-learning. *Computers & Education*, 142, 103642–103659. <https://doi.org/10.1016/j.compedu.2019.103642>
- Harrathi, M., & Braham, R. (2021). Recommenders in improving students' engagement in large scale open learning. *Procedia Computer Science*, 192, 1121–1131. <https://doi.org/10.1016/j.procs.2021.08.115>
- Herpich, F., Nunes, F., Petri, G., & Tarouco, L. (2019). How Mobile augmented reality is applied in education? A systematic literature review. *Creative Education*, 10, 1589–1627. <https://doi.org/10.4236/ce.2019.107115>
- Huang, L., Wang, C.-D., Chao, H.-Y., Lai, J.-H., & Yu, P. S. (2019). A score prediction approach for optional course recommendation via cross-user-domain collaborative filtering. *IEEE Access*, 7, 19550–19563. <https://doi.org/10.1109/ACCESS.2019.2897979>

- Iaquinta, L., Gemmis, M. de Lops, P., Semeraro, G., Filannino, M. & Molino, P. (2008). Introducing serendipity in a content-based recommender system. *Proceedings of the Eighth International Conference on Hybrid Intelligent Systems*, 168–173. <https://doi.org/10.1109/HIS.2008.25>
- Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 16(3), 261–273. <https://doi.org/10.1016/j.eij.2015.06.005>
- Ismail, H. M., Belkhouche, B., & Harous, S. (2019). Framework for personalized content recommendations to support informal learning in massively diverse information Wikis. *IEEE Access*, 7, 172752–172773. <https://doi.org/10.1109/ACCESS.2019.2956284>
- Khan, K. S., Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine*, 96(3), 118–121. <https://doi.org/10.1258/jrsm.96.3.118>
- Khanal, S. S., Prasad, P. W. C., Alsadoon, A., & Maag, A. (2019). A systematic review: Machine learning based recommendation systems for e-learning. *Education and Information Technologies*, 25(4), 2635–2664. <https://doi.org/10.1007/s10639-019-10063-9>
- Khusro, S., Ali, Z., & Ullah, I. (2016). Recommender Systems: Issues, Challenges, and Research Opportunities. In K. Kim & N. Joukov (Eds.), *Lecture Notes in Electrical Engineering* (Vol. 376, pp. 1179–1189). Springer. https://doi.org/10.1007/978-981-10-0557-2_112
- Kitchenham, B. A., & Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. *Technical Report EBSE 2007–001*. Keele University and Durham University Joint Report. https://www.elsevier.com/data/promis_misc/525444systematicreviewsguide.pdf
- Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, 51(1), 7–15. <https://doi.org/10.1016/j.infsof.2008.09.009>
- Klašnja-Milićević, A., Ivanović, M., & Nanopoulos, A. (2015). Recommender systems in e-learning environments: A survey of the state-of-the-art and possible extensions. *Artificial Intelligence Review*, 44(4), 571–604. <https://doi.org/10.1007/s10462-015-9440-z>
- Klašnja-Milićević, A., Vesin, B., & Ivanović, M. (2018). Social tagging strategy for enhancing e-learning experience. *Computers & Education*, 118, 166–181. <https://doi.org/10.1016/j.compedu.2017.12.002>
- Kolb, D., Boyatzis, R., Mainemelis, C., (2001). *Experiential Learning Theory: Previous Research and New Directions Perspectives on Thinking, Learning and Cognitive Styles*, 227–247.
- Krahenbuhl, K. S. (2016). Student-centered Education and Constructivism: Challenges, Concerns, and Clarity for Teachers. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 89(3), 97–105. <https://doi.org/10.1080/00098655.2016.1191311>
- Kunaver, M., & Požrl, T. (2017). Diversity in recommender systems – A survey. *Knowledge-Based Systems*, 123, 154–162. <https://doi.org/10.1016/j.knosys.2017.02.009>
- Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2010). Recommender systems in technology enhanced learning. In F. Ricci, L. Rokach, B. Shapira, & P. Kantor (Eds.), *Recommender Systems Handbook* (pp. 387–415). Springer. https://doi.org/10.1007/9780-387-85820-3_12
- Manouselis, N., Drachsler, H., Verbert, K., & Santos, O. C. (2014). *Recommender systems for technology enhanced learning*. Springer. <https://doi.org/10.1007/978-1-4939-0530-0>
- Manouselis, N., Drachsler, H., Verbert, K., & Duval, E. (2013). Challenges and Outlook. *Recommender Systems for Learning*, 63–76. <https://doi.org/10.1007/978-1-4614-4361-2>
- Maravanyika, M., & Dlodlo, N. (2018). An adaptive framework for recommender-based learning management systems. *Open Innovations Conference (OI)*, 2018, 203–212. <https://doi.org/10.1109/OI.2018.8535816>
- Maria, S. A. A., Cazella, S. C., & Behar, P. A. (2019). Sistemas de Recomendação: conceitos e técnicas de aplicação. *Recomendação Pedagógica em Educação a Distância*, 19–47. Penso.
- McCombs, B. L. (2013). The Learner-Centered Model: Implications for Research Approaches. In Cornelius-White, J., Motschnig-Pitrik, R. & Lux, M. (eds), *Interdisciplinary Handbook of the Person-Centered Approach*, 335–352. 10.1007/978-1-4614-7141-7_23
- Medeiros, R. P., Ramalho, G. L., & Falcao, T. P. (2019). A systematic literature review on teaching and learning introductory programming in higher education. *IEEE Transactions on Education*, 62(2), 77–90. <https://doi.org/10.1109/te.2018.2864133>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1), 1. <https://doi.org/10.1186/2046-4053-4-1>

- Moubayed, A., Injadat, M., Nassif, A. B., Lutfiyya, H., & Shami, A. (2018). E-Learning: Challenges and research opportunities using machine learning & data analytics. *IEEE Access*, 6, 39117–39138. <https://doi.org/10.1109/access.2018.2851790>
- Nabizadeh, A. H., Gonçalves, D., Gama, S., Jorge, J., & Rafsanjani, H. N. (2020). Adaptive learning path recommender approach using auxiliary learning objects. *Computers & Education*, 147, 103777–103793. <https://doi.org/10.1016/j.compedu.2019.103777>
- Nafea, S. M., Siewe, F., & He, Y. (2019). On Recommendation of learning objects using Felder-Silverman learning style model. *IEEE Access*, 7, 163034–163048. <https://doi.org/10.1109/ACCESS.2019.2935417>
- Nascimento, P. D., Barreto, R., Primo, T., Gusmão, T., & Oliveira, E. (2017). Recomendação de Objetos de Aprendizagem baseada em Modelos de Estilos de Aprendizagem: Uma Revisão Sistemática da Literatura. *Proceedings of XXVIII Brazilian Symposium on Computers in Education- SBIE*, 2017, 213–222. <https://doi.org/10.5753/cbie.sbie.2017.213>
- Nguyen, Q. H., Ly, H.-B., Ho, L. S., Al-Ansari, N., Le, H. V., Tran, V. Q., Prakash, I., & Pham, B. T. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Mathematical Problems in Engineering*, 2021, 1–15. <https://doi.org/10.1155/2021/4832864>
- Nichols, D. M. (1998). Implicit rating and filtering. *Proceedings of the Fifth Delos Workshop: Filtering and Collaborative Filtering*, 31–36.
- Okoye, I., Maull, K., Foster, J., & Sumner, T. (2012). Educational recommendation in an informal intentional learning system. *Educational Recommender Systems and Technologies*, 1–23. <https://doi.org/10.4018/978-1-61350-489-5.ch001>
- Pai, M., McCulloch, M., Gorman, J. D., Pai, N., Enanoria, W., Kennedy, G., Tharyan, P., & Colford, J. M., Jr. (2004). Systematic reviews and meta-analyses: An illustrated, step-by-step guide. *The National Medical Journal of India*, 17(2), 86–95.
- Petri, G., & Gresse von Wangenheim, C. (2017). How games for computing education are evaluated? A systematic literature review. *Computers & Education*, 107, 68–90. <https://doi.org/10.1016/j.compedu.2017.01.00>
- Petticrew, M., & Roberts, H. (2006). Systematic reviews in the social sciences a practical guide. *Blackwell Publishing*. <https://doi.org/10.1002/9780470754887>
- Pinho, P. C. R., Barwaldt, R., Espindola, D., Torres, M., Pias, M., Topin, L., Borba, A., & Oliveira, M. (2019). Developments in educational recommendation systems: a systematic review. *Proceedings of 2019 IEEE Frontiers in Education Conference (FIE)*. <https://doi.org/10.1109/FIE43999.2019.9028466>
- Pöntinen, S., Dillon, P., & Väisänen, P. (2017). Student teachers' discourse about digital technologies and transitions between formal and informal learning contexts. *Education and Information Technologies*, 22(1), 317–335. <https://doi.org/10.1007/s10639-015-9450-0>
- Pu, P., Chen, L., & Hu, R. (2011). A user-centric evaluation framework for recommender systems. *Proceedings of the fifth ACM conference on Recommender systems*, 157–164. <https://doi.org/10.1145/2043932.2043962>
- Rahman, M. M., & Abdullah, N. A. (2018). A personalized group-based recommendation approach for web search in E-Learning. *IEEE Access*, 6, 34166–34178. <https://doi.org/10.1109/ACCESS.2018.2850376>
- Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender Systems: Introduction and Challenges. I Ricci, F., Rokach, L., Shapira, B. (eds), *Recommender Systems Handbook*, 1–34. https://doi.org/10.1007/978-1-4899-7637-6_1
- Rivera, A. C., Tapia-Leon, M., & Lujan-Mora, S. (2018). Recommendation Systems in Education: A Systematic Mapping Study. *Proceedings of the International Conference on Information Technology & Systems (ICITS 2018)*, 937–947. https://doi.org/10.1007/978-3-319-73450-7_89
- Salazar, C., Aguilar, J., Monsalve-Pulido, J., & Montoya, E. (2021). Affective recommender systems in the educational field. A systematic literature review. *Computer Science Review*, 40, 100377. <https://doi.org/10.1016/j.cosrev.2021.100377>
- Santos, I. M., & Ali, N. (2012). Exploring the uses of mobile phones to support informal learning. *Education and Information Technologies*, 17(2), 187–203. <https://doi.org/10.1007/s10639-011-9151-2>
- Sergis, S., & Sampson, D. G. (2016). Learning object recommendations for teachers based on elicited ICT competence profiles. *IEEE Transactions on Learning Technologies*, 9(1), 67–80. <https://doi.org/10.1109/TLT.2015.2434824>

- Shani, G., & Gunawardana, A. (2010). Evaluating recommendation systems. In F. Ricci, L. Rokach, B. Shapira, & P. Kantor (Eds.), *Recommender Systems Handbook* (pp. 257–297). Springer. https://doi.org/10.1007/978-0-387-85820-3_8
- Tahereh, M., Maryam, T. M., Mahdihyeh, M., & Mahmood, K. (2013). Multi dimensional framework for qualitative evaluation in e-learning. *4th International Conference on e-Learning and e-Teaching (ICELET 2013)*, 69–75. <https://doi.org/10.1109/icelet.2013.6681648>
- Tarus, J. K., Niu, Z., & Yousif, A. (2017). A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. *Future Generation Computer Systems*, 72, 37–48. <https://doi.org/10.1016/j.future.2017.02.049>
- 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., Drachsler, 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. <https://doi.org/10.1109/TLT.2012.11>
- Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., & Duval, E. (2011). Dataset-Driven Research for Improving Recommender Systems for Learning. *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, 44–53. <https://doi.org/10.1145/2090116.2090122>
- Wan, S., & Niu, Z. (2016). A learner oriented learning recommendation approach based on mixed concept mapping and immune algorithm. *Knowledge-Based Systems*, 103, 28–40. <https://doi.org/10.1016/j.knsys.2016.03.022>
- Wan, S., & Niu, Z. (2018). An e-learning recommendation approach based on the self-organization of learning resource. *Knowledge-Based Systems*, 160, 71–87. <https://doi.org/10.1016/j.knsys.2018.06.014>
- Wan, S., & Niu, Z. (2020). A hybrid E-Learning recommendation approach based on learners' influence propagation. *IEEE Transactions on Knowledge and Data Engineering*, 32(5), 827–840. <https://doi.org/10.1109/TKDE.2019.2895033>
- Watkins, K. E., & Marsick, V. J. (2020). Informal and incidental learning in the time of COVID-19. *Advances in Developing Human Resources*, 23(1), 88–96. <https://doi.org/10.1177/1523422320973656>
- Wu, D., Lu, J., & Zhang, G. (2015). A Fuzzy Tree Matching-based personalized E-Learning recommender system. *IEEE Transactions on Fuzzy Systems*, 23(6), 2412–2426. <https://doi.org/10.1109/TFUZZ.2015.2426201>
- Wu, Z., Li, M., Tang, Y., & Liang, Q. (2020). Exercise recommendation based on knowledge concept prediction. *Knowledge-Based Systems*, 210, 106481–106492. <https://doi.org/10.1016/j.knsys.2020.106481>
- Yanes, N., Mostafa, A. M., Ezz, M., & Almuayqil, S. N. (2020). A machine learning-based recommender system for improving students learning experiences. *IEEE Access*, 8, 201218–201235. <https://doi.org/10.1109/ACCESS.2020.3036336>
- Zapata, A., Menéndez, V. H., Prieto, M. E., & Romero, C. (2015). Evaluation and selection of group recommendation strategies for collaborative searching of learning objects. *International Journal of Human-Computer Studies*, 76, 22–39. <https://doi.org/10.1016/j.ijhcs.2014.12.002>
- Zhang, S., Yao, L., Sun, A., & Tay, Y. (2020). Deep learning based recommender system. *ACM Computing Surveys*, 52(1), 1–38. <https://doi.org/10.1145/3285029>
- Zhong, J., Xie, H., & Wang, F. L. (2019). The research trends in recommender systems for e-learning: A systematic review of SSCI journal articles from 2014 to 2018. *Asian Association of Open Universities Journal*, 14(1), 12–27. <https://doi.org/10.1108/AAOUJ-03-2019-0015>

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.