User Modeling and User-Adapted Interaction (2025) 35:15
https://doi.org/10.1007/s11257-025-09434-4

Learner models: design, components, structure, and modelling

A systematic literature review

Felix Böck[1,2] · Michaela Ochs[1] · Andreas Henrich[1] · Dieter Landes[2] · Jochen L. Leidner[2,3] · Yvonne Sedelmaier[2]

Received: 30 December 2024 / Accepted in revised form: 25 July 2025

Abstract
Learning is at the heart of every progress the human species makes. It is most effective when it considers who we are as individuals, what learning approach we prefer and what we already know to begin with. In the digital age, we strive to capture such information in the form of a digital representation—the so-called learner model—, to tailor learning-related systems to this information and build upon it to create more personalised learning experiences. Over recent years, the proliferation of diverse models across various educational applications and disciplines has made it challenging to access targeted research. In this survey, we aim to address this gap, reviewing the latest advances in learner modelling and conducting a comprehensive analysis of the existing approaches, focusing on developments from 2014 to 2023. With the help of a systematic literature review (SLR), we want to provide designers and developers of learner models with a structured overview and simplified entrance into the topic and the field of learner models. We investigate the question: *What do learner models look like and how are they filled, kept up to date and used?* To this end, we analyse and classify existing approaches. Our findings provide a comprehensive and structured overview of the field of learner modelling, allowing researchers to navigate and understand the diverse approaches more easily and providing developers of learner models or adaptive systems with a practical tool to access relevant information according to their needs.

Keywords Learner model · Digital twin · Student modelling · Systematic literature review · Systematic survey · Systematic review · Higher education · Computer science · Computer-assisted learning

Felix Böck and Michaela Ochs contributed equally to this work. Author names other than the joint first authors are in alphabetical order.

Extended author information available on the last page of the article

Published online: 04 November 2025

1 Introduction

Students at universities are becoming increasingly heterogeneous due to various reasons. Internationalisation, for instance, is leading to a mix of different school systems. Other reasons include the elimination of entrance requirements for admission to a degree programme. This ‘education for all’ approach creates new challenges for lecturers as the educational differences between students are increasing. However, the overall goal remains the same: completing the degree programme as successfully as possible. In order to achieve this, it is becoming increasingly important to address the learner’s individual needs. Due to the large number of students, it is usually impossible for lecturers to do this manually, and they must use computer-aided methods (CAM).

Computer-based training (CBT), software systems to support students’ learning process and even human-led instruction can all benefit from a high degree of individualisation. Besides relying exclusively on relatively static knowledge about the learner, such as their age or educational background, individualised support also means considering dynamic information that changes throughout time. For an individualised system or individualised instruction, this can manifest in considering the specific status of the learner, such as their previous knowledge in a particular area, their current emotional, physical or mental state or the learning context the learners find themselves in, and adapting the specific educational methods or mechanisms to this state.

Looking at the current literature in the field of (automatic) adaptive education, a vast amount of research papers deal with this aspect of individualisation, i.e. the aspect of how to adapt learning-related systems to the user. Two basic strands of research can be distinguished here. On the one hand, approaches dealing with the recommendation mechanisms and, on the other hand, those which generate the basis for this, i.e. the models that take care of the user and the items to be recommended. This large strand of research looks at how to create a digital representation of these models and, thus, how to construct the foundation for the concurrent step of automatic adaption in the form of educational recommender systems or similar systems. These systems, as well as the interactions between the various modules, are examined in more detail in relevant literature on the architecture of Intelligent Tutoring Systems (ITS), such as Böck et al. (2025) or Sottilare and Schwarz (2020), Sottilare et al. (2013) and Pavlik et al. (2013). The topic of architectures for learning systems is out of our scope, as we deal exclusively with learner models as a part of adaptive learning systems. As Dillenbourg and Self (1992) already noted, several different theories and techniques for modelling learners exist in research, but often they are not considered in more detail.

In order to offer adaptive computer-based training (CBT), science differentiates between the learner (student) and the content to be taught (knowledge/skills). Figure 1 shows the relationship between the learner (*learner model*) and the content to be learned (*domain model*). While the learner model contains the various learner’s attributes ($\widehat{=}$ WHO), the domain model contains the concepts or entities and their relations, which can be learned ($\widehat{=}$ WHAT). In our fictitious example, the learner Jane Doe is on one side with her attribute specification for age, gender, language and others, including information such as prior knowledge. On the other side, the content to be learned is depicted, which is recursion in this example—a unit from computer science. While we focus on the learner model in this work, it is important to note that the two

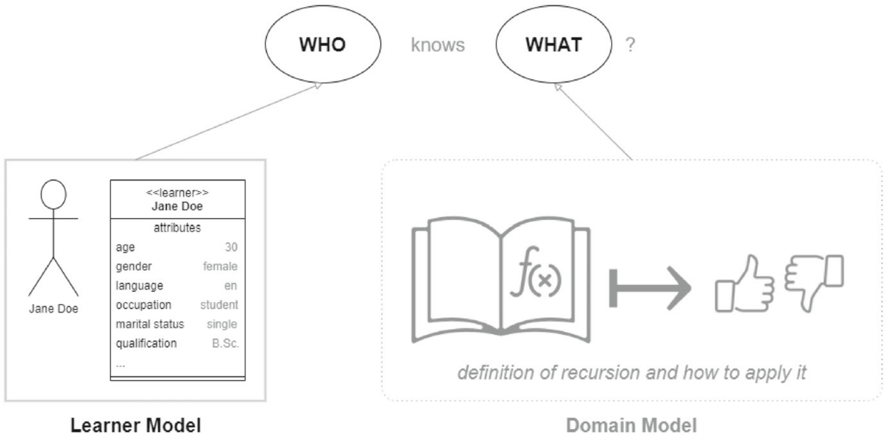

Fig. 1 What to model? Visualisation of the relationship of learner and domain model

models cannot be considered in isolation, as they are related. The scope of the content to be learned, which is in the domain model that the learner has already mastered, must also be incorporated into the learner model.

This systematic literature research builds on this differentiation of the models and focuses exclusively on the digital representation of the learner, the so-called *learner model*. To summarise the definition of the learner model in this light, it is a collection of learner characteristics used to inform the learning process and adapt the content in the domain model. Although such works often include a large variety of learner details, we found that they often remain highly vague about the concrete features that a learner model contains (i.e. the detailed formal representation).

The concept of learner models for individualisation in computer-assisted teaching programmes has been around for over half a century (Self 1974). The challenges in creating a learner model are often still the same, such as the question of which data to integrate and the subsequent representation of the data in the learner model, but also the evolution of the data contained (Self 1979). Several approaches to reviewing the current literature on learner modelling have been undertaken throughout the past years, covering all stages of the process of learner modelling, ranging from the step of data collection and the examination of the types of data contained in the models to describe the nature of the approaches that are commonly used to model this data. We noticed that these works often lack precision in characterising the data contained within the models more systematically and in more detail, i.e. they fail to distinguish between concrete data and derived learner characteristics. Also, they often focus on one or two specific steps of the overall modelling process.

Throughout this work, we will look at the complex environment of (guided) learning and quickly realise that many disciplines work together in this context. In this paper, we explicitly focus only on the sub-disciplines of computer science (technical realisation) and pedagogy (didactics of computer science and computer-supported learning) and deliberately leave out other important areas, such as the humanities. We will look at learner models from the perspective of computer science and pedagogy to answer the major question: *What do learner models look like and how are they filled, kept up to date and used?*, which will be specified further in Sect. 3.1. A criterion for selecting the models is their possible (technical) feasibility, even if they are currently only described theoretically. Purely theoretical thought experiments without reference to practice or the possibility of implementation are not included.

Accordingly, the remainder of this article is structured as follows: To create a common baseline, the most important terms and the scope of this research are defined in Sect. 2. Then, the research methodology is described in detail in Sect. 3, which forms the basis for the systematic literature review in Sect. 4. In the synthesis section of the systematic literature review, research questions that we specify in Sect. 3.1 on the topic of learner models are explored, and a selection of relevant papers is reviewed, with each paper being evaluated concerning the previously established requirements. After shortly describing the application of the SLR method as well as providing an insight into the composition of the results in Sect. 4.1, a general classification of the results that are equally relevant for all research questions is presented in Sect. 4.2. Thereafter, Sect. 4.3 provides a short introduction to how to interpret the results of the following parts, and then we dive into the results for the individual research questions. Section 4.4 thus addresses the structure and components of a learner model. Section 4.5 presents the results of the modelling of the learner model, and Sect. 4.6 deals with the topic of forgetting knowledge. Finally, Sect. 4.7 provides overall results on learner models and their development over the last ten years. After the fine-grained analysis and the results, we dive into further discussion in Sect. 5. On the one hand, we examine the purpose and applications of such learner models in Sect. 5.1, and on the other hand, we show challenges but also opportunities for developing adaptive learning systems with learner models in Sect. 5.2, highlighting outstanding research deficits. After providing a brief outlook on the research of 2024 in Sect. 6, we conclude with a summary and the limitations of this work and give an outlook in Sect. 7.

2 Definition of terms

The term *learner model* is not uniformly defined in the literature. It is, therefore, often freely used and controversially discussed from author to author in (scientific) publications, but what does it mean? To be able to answer this and similar questions specifically, we will reduce the term *learner model* to its core, which is to be considered separately in the first step.

A *model* describes a partial aspect of a real system or subject relevant to the respective purpose as a conscious abstraction of it. The most important characteristics of models are:

- models do not describe the respective system or subject completely, but from a certain point of view;
- several models can exist for a system or subject;
- models exist at different levels of abstraction; from a high-level view to a detailed representation;

- models consist of various components[1];
- relevant properties of the original must be adequately and completely mapped to properties of the models (validation required).

In our field of research, the model serves as a digital representation of the student (learner), aiming for individual learning support in a digital environment. The ambiguous nature of the concept of the model and the diverse shapes a learner can take in digital learning lead to the manifold forms and interpretations of the learner model.

Before we take a closer look at how the term differs from other related concepts in the scientific literature and then refine the definition, we will first examine the learner model in its more general form, the so-called *user model*. According to Johnson and Taatgen (2005), "two things are crucial for an adaptive system to work: the existence of a means to adapt the task and the ability to detect the need for adaptation" (p. 430). User models generally represent this ability to adapt. Typically, the term *user model* (UM) differs widely in the literature (Kay et al. 2022) and depends on the intended use, scope, domain and the way information about the user is collected (Hlioui et al. 2016). In general, one can say that the "user model is a representation of static and dynamic information about an individual that is used throughout the whole interaction process aiming to trigger several adaptation and personalisation effects. [...] [This user model] entails all the information which is considered important to adapt and personalise the user interface (content and navigation) and functionalities to the unique characteristics of a learner. [...] Depending on the domain and goals of the system, user models can include different kinds of characteristics about the learners (e.g. interests, preferences, traits) or data concerning their overall context of use (e.g. environment, time, interaction device type, etc.)" (Germanakos and Belk 2016, p. 79).

Next, it is necessary to highlight the difference between two frequently used terminologies in this context, the so-called *user profile* and the *user model*. The user profile tends to contain personal information about the user without any interpretation (de Koch 2000; Liang et al. 2022). This profile is part of the user model, which is extended by other learner characteristics, such as competencies or other cognitive characteristics, in a modular way through various components. Other definitions and interpretations of the term user profile, such as in the context of the application management process as a pure role concept, may lead to misunderstandings and are not included in this research's understanding of the user model.

As our interest is in tertiary education in the domain of higher education, we concretise the general user model that was just considered to be a more specialised learner model. The *learner model* (LM) is thus an "explicit representation of the individual characteristics and interaction data of a single learner in the form of a machine internal representation" (Böck 2025) and includes many different aspects, which ultimately depend on the purpose of the application. The learner model consists of several subparts (components), whereby each component is usually a characteristic of the learner to be modelled and is managed individually. We understand a learner model exclusively as the digitally processed representation of the learner's characteristics. Inference mechanisms and reasoning are further steps that can be applied to the learner model and that, in our view, are not part of it but must be considered separately. Additionally, it is

[1] The definition of components can be found at the end of this section.

essential to distinguish the learner model from *learning analytics* as the latter involves the analysis of group behaviours to predict individual outcomes, while the former is used as a basis for adapting the learning process for individual learners.

In this research, the terms *student model* (Greer and McCalla 1994) or *pupil model* are used synonymously with the term *learner model* as our learners are usually university students. On the other hand, the terms *digital DNA model* and *digital sibling model* rarely appear in the literature in this context. Still, they should be seen as synonyms for the term *learner model* in our research. What is often used as a synonym for learner model in the literature is the term *digital twin*, which is intended to symbolise the concept of a digital representation of the learner. The process of creating such a learner model is called *learner modelling* (Piao and Breslin 2018; Khenissi et al. 2015) throughout this research.

3 Research design and methodology

With the increase in digital learning offers, there is also considerable growth in research and practice in the related area of learner models in digital ecosystems, as we have outlined before. To get an overview of related work around the topic of learner models, we decided to conduct a systematic literature review (SLR) following Kitchenham (2004), Kitchenham and Charters (2007). This SLR aims to provide a balanced and comprehensive overview of the current topic of learner models in the literature, which should serve designers and developers as a starting point for developing their individual learner models. The SLR is to be executed in several iterations to be able to report as comprehensively as possible. For this purpose, the first iteration, according to Kitchenham and Charters (2007), was performed, and the gathered literature was analysed. This is followed by at least one further iteration, in which works outside previously defined criteria, such as publisher or year, are also included and are considered and labelled separately. In our work, the second iteration consists of works that are relevant or frequently cited ($\geq$ 20 citations per year) but were not included in the first iteration due to our defined criteria (cf. Sect. 3.4). Following this two-step approach, we explore a comprehensive thematic scope defined by the following research questions.

3.1 Research questions

To answer the complex question *What do learner models look like and how are they filled, kept up to date, and used?*, we divided it further into three sub-areas:

- **RQ1**: What do learner models look like, and what are their structure and components? (Sect. 4.4)
- **RQ2**: Which data sources are used to initialise and update learner models, how is the data collected, and what is the overall modelling approach? (Sect. 4.5)
- **RQ3**: How do learner models evolve during their lifetime concerning the learner's ephemeral knowledge? (Sect. 4.6)

3.2 Identifying relevant sources

Table 1 provides an overview of the sources found to be relevant for this research. As shown, seven electronic database sources [RS1.1–RS1.5] were selected, which are most relevant for the fields of technical aspects of education or (digital) teaching and computer science—in particular, the education and training of computer scientists. An important criterion in selecting data sources from these fields was to obtain published articles—conference proceedings, journal/magazine articles, or technical reports—that meet minimum quality criteria (e.g. peer-reviewed). Senior researchers in the two fields confirmed the relevance and reliability of the selected sources. In addition, we optionally searched preprint servers or repositories [RS2.1–RS2.16] to find scientific publications that were not identified in the primary search. We used these mostly preprint results under consideration of our quality criteria (cf. Sect. 3.4) to refer to the correct publications. If results cannot be assigned to any of the databases mentioned above [RS1.1–RS1.5], they are summarised under *various*, which is particularly relevant for the second iteration. Additional publicly indexed databases [RS3.1–RS3.7] were used for plausibility checks of the results.

In addition to the numerous scientific publications, standards and norms should be examined more closely for a comprehensive overview. We have already done this in a parallel publication (Böck et al. 2024) and include the results in Sect. 4.4 to shed light on as many perspectives on learner models as possible.

3.3 Search strategy

To obtain all the desired results and thus to be able to create as comprehensive an overview of the domain as possible, we have defined two dimensions of the search for this purpose as visualised by Fig. 2. The *learner model* dimension (Fig. 2—top) and its area of application, i.e. the application *context* dimension (Fig. 2—bottom).

For the learner model dimension, on the one hand, we have created a list of synonyms, abbreviations, and alternative spellings with synonym and spelling dictionaries and keywords in relevant publications, as well as feedback from multiple senior researchers. For the context dimension, on the other hand, all possible application contexts for learner models such as 'e-learning' or 'personal learning' that were suitable for our purpose were collected and complemented by possible synonyms and alternative spellings, for instance, 'elearning', 'personalised learning' or 'personalised learning'. We could simplify this basis for our search query by subsuming the variations under the broad terms 'learning' and 'education' (visualised by Fig. 2—bottom), knowing that this would also find syntactically identical but semantically different areas of use. For instance, if the search for 'e-learning', 'adaptive learning' and 'personal learning' is generalised to 'learning', results for 'machine learning' can also be found as it matches 'learning'. These semantically non-matching hits were then filtered out manually. This list does not claim to be complete but covers many relevant search terms.

We linked all these different variations meaningfully using boolean operators as shown in Fig. 3. The syntax of the resulting query in the respective query language

Table 1 Overview of relevant research sources

Relevant research sources		
Reference	Provider	Domain
Electronic database resources		
RS1.1	IEEE Xplore	ieeexplore.ieee.org
RS1.2	ACM Digital Library	dl.acm.org
RS1.3	Elsevier ScienceDirect	www.sciencedirect.com
RS1.4a	SpringerLink	link.springer.com
RS1.4b	SpringerOpen	www.springeropen.com
RS1.4c	Springer Nature	www.springernature.com
RS1.5	pedocs	www.pedocs.de
Additional public alternative preprint servers		
RS2.1	arXiv [Cornell University]	www.arxiv.org
RS2.2	TechRxiv	www.techrxiv.org
RS2.3	ϵasyChair πreprints	www.easychair.org/publications/preprints
RS2.4	zenodo	www.zenodo.org
RS2.5	OSFPreprints	www.osf.io/preprints
RS2.6	SocArXiv	www.socopen.org
RS2.7	figshare	www.figshare.com
RS2.8	HAL	www.hal.science
RS2.9	CERN Document Server	cds.cern.ch/collection/Preprints
RS2.10	Preprints.org	www.preprints.org
RS2.11	viXra.org	www.vixra.org
RS2.12	AUTHOREA	www.authorea.com
RS2.13	Qeios	www.qeios.com
RS2.14	PeerJ	www.peerj.com
RS2.15	F1000Research	www.f1000research.com
RS2.16	Academia	www.academia.edu
Additional publicly indexed databases for plausibility checks		
RS3.1	ResearchGate	www.researchgate.net
RS3.2	Google Scholar	scholar.google.de
RS3.3	Semantic Scholar	www.semanticscholar.org
RS3.4	CiteSeer^{x}	citeseerx.ist.psu.edu
RS3.5a	Elsevier Scopus	www.scopus.com
RS3.5b	Elsevier Engineering Village	www.engineeringvillage.com
RS3.6	dblp computer science bibliography	www.dblp.org
RS3.7	Clarivate—Conference Proceedings Citation Index	www.clarivate.com/products/scientific-and-academic-research

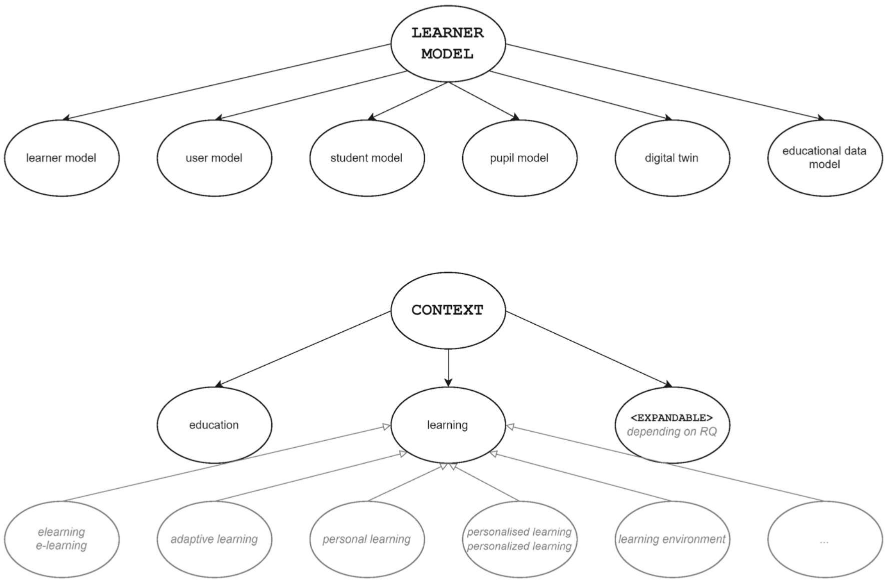

Fig. 2 Overview over search terms: Search terms related to learner model (top) and search terms related to this research's context (bottom), subsumed under umbrella terms.

```
("learner model" OR "user model" OR "student model" OR "pupil
model" OR "digital twin" OR "educational data model")
AND
("learning" OR "education")
IN
(METADATA OR TITLE OR KEYWORDS OR ABSTRACT)
```

Fig. 3 Boolean search terms.

may vary depending on the data source and may, therefore, need minimal modification while the semantics remain unchanged. When searching for German publications, the terms were translated into German and linked with the same operators.

The next step is to define the area to be searched within the relevant selected research sources. The search primarily taps into the fields of metadata, title, keywords and abstract. We do not consider a full-text search appropriate since the results vary greatly and are not pre-sorted and filtered. When looking at the sources that were previously identified to be relevant in Sect. 3.2 in Table 2, it becomes apparent that Springer's services currently only provide the public with full-text search, and they are therefore not included in the search for the moment.

In IEEE Xplore, the *command search* was used, which allows the definition of such search strings. The ACM Digital Library also supports defining the search as a query in the *advanced search*. The logic can also be mapped at Elsevier ScienceDirect in a fixed structure via individual text fields. Most of the publications found are either freely accessible or were obtained via institutional login. Only a few publications could not be found this way, but a private license was required. Several papers were obtained through personal correspondence with the authors.

Table 2 Overview over search options for search providers

Publisher	Search in ... Possible?		
	Data Fields	Full Text	Searched Data Fields
IEEE Xplore	☑	☑	Metadata, Title, Keywords & Abstract
ACM DL	☑	☑	Title, Keywords & Abstract
Elsevier ScienceDirect	☑	☑	Title, Abstract or author-specified Keywords
Springer SpringerLink	—	☑	
Springer SpringerOpen	—	☑	
Springer Springer Nature	—	☑	
pedocs	☑	☑	Title, Keywords & Abstract

3.4 Selection criteria and filters

To limit the number of final results before summarising them, we defined various selection criteria to approach the results in a more targeted manner. Selection criteria include both exclusion criteria and inclusion criteria. Inclusion criteria are specific characteristics a publication must have to be selected, while exclusion criteria are characteristics that disqualify certain publications from being included. In principle, both can be reformulated into each other, and for simplicity, we only use exclusion criteria. Usually, exclusion criteria can be answered in a binary form with true and false—often even automatically (e.g. language is English)—while filters are often based on a more subjective—manual—evaluation (e.g. abstract is relevant). The criteria and filters are listed and briefly explained below.

3.4.1 Exclusion criteria

Period of Selection. Due to rapid technological progress, especially in the field of technology-based education, we attach great importance to the timeliness of the selection period of publications. Therefore, we consider publications from the last decade more closely, i.e. from 01/01/2014 to 31/12/2023.

Language. Only English or German publications were selected from the database resources.[2]

Quality Criteria. From the above database resources, only publications that were subject to quality control before publication were selected. For conference papers, for example, this is usually some kind of peer review. In addition, this criterion is used to weed out (extended) abstracts, workshops, courses, presentations or keynotes, as these generally contain little new and concrete information in sufficient detail.

[2] We know this selection could introduce a geographic bias. Our decision is based on the following considerations: On the one hand, we believe many relevant papers are published in English. On the other hand, interpreting or translating publications in other languages would require additional technical resources beyond the scope of this study. We also found that while relevant English papers often have a technical focus, papers with a more pedagogical outlook are often (only) available in German.

3.4.2 Filter

The filters are run sequentially in the order listed below. Multiple researchers should preferably review these filters and their applications to weed out irrelevant papers.
Title. This filter excludes all papers whose title is not related to the research questions (Sect. 3.1).
Keywords. This filter excludes all papers whose keywords are not related to the research question.
Abstract. This filter excludes all papers whose abstract does not fit the topic of *learner models in adaptive learning systems*.
Conclusion. Since the abstract alone is often not meaningful enough, this filter excludes all papers where the conclusion Brereton et al. (2007) is not related to the research question and the topic described above.
Full-Text. In the last step, after the full-text analysis, a decision is made about the relevance of the present publication.

These selection criteria and filters can also be deliberately overridden individually (except for quality) in exceptional cases in our second iteration, for example, if a model is referenced by numerous papers (thus constituting a base work) but was published before 2013. Furthermore, after the full-text analysis, all references of each relevant article and their authors are checked to see if other important publications outside the search scope should be included in the results. These results of the meta-analysis are then also included in the second iteration. This way, publications from different search providers that can be traced back to the same approach and authorship are summarised and treated as one publication.

3.5 Selection of extracted data

Our selection of extracted data serves as a basis for comparing the different models and their features, and thus, identical data is collected from all models found. We primarily focused on extracting data from the publications that address this article's specific research questions (cf. Sect. 3.1). Details on these questions will be addressed in the respective parts in Sect. 4. In addition to the content specific to the research questions, further data was extracted from the publications. This includes the classification of the publication and its approach into the corresponding domain (e.g. higher education, adult learning), as well as further metadata such as the title of the publication, author, year of publication, the number of references, the number of citations (on 18/06/2024) and whether the publication is a literature research paper that takes up other approaches.

Since numerous literature surveys are among the results, it is necessary to consider a procedure for dealing with the contents of secondary literature. The scientifically cleanest solution would be to extract the relevant sources from the secondary literature and research them individually, turning them into primary sources. However, this is not possible in the time available due to the large amount of data to be processed. In order to acknowledge the work involved in each literature research, we have decided to list the surveys as primary literature sources and to present the sources they cite as direct references, adding the postfixes ‘mentioned by’ or ‘cited by’, to acknowledge the actual primary source.

4 Results

4.1 Composition of results

Table 3 shows the number of hits for the boolean query defined in Fig. 3 for each search provider without any filters. To ensure reproducibility, we repeated the searches at irregular intervals. The result is surprising, as the expected quantitative number of results does not match deterministically and reproducibly[3]. We have decided to use 01/01/2024 as the starting point[4].

As Fig. 4 shows, 3464 (3437 in the first iteration plus 27 in the second iteration) results emerged from the initial search run. After applying the objective exclusion criteria (reproducible process), i.e. time, language and quality, in the first iteration, 2792 publications were still available (= 81% of the total number). Then, the more subjective filters (individual process) were applied, assessing if the title, keywords, abstract and conclusion are relevant to the research questions, after which 315 (288 in the first iteration | 27 in the second iteration) results of the original 3437 (= 8%) of the first iteration were still available for full-text analysis. These remaining publications were analysed in relation to the respective research question. The results of this analysis can be found in Sects. 4.4, 4.5 and 4.6.

Figure 5 shows the distribution of relevant publications of the first iteration by year and type of publication. This graph shows that most of the contributions are conference publications. It can be seen that there was a slight drop in publications in 2015 and 2019. Potential explanations for the increase in 2020 might be the global pandemic with the need for digital learning, which is also substantiated by the publications’ content, or the increased interest in adaptive technologies based on artificial intelligence.

After looking at the distribution of the relevant publications from Fig. 5, we now look at the distribution over time for the publications originally found with the queries in the first iteration (2797). Here, a strong upward trend is visible over the last decade. Figure 6 shows this rapid increase in publications over the last few years, which is likely linked to the switch to digital learning driven by the pandemic.

What becomes evident when we examine both figures is that while Fig. 6 shows a significant increase in the publications retrieved before applying the filters, Fig. 5 demonstrates that the number of relevant papers per year remains roughly constant over the decade examined. This indicates that significantly more papers were excluded in recent years compared to earlier years. Our hypothesis for this effect can be explained

[3] According to logical criteria, the search query with only the year filter between 01/01/2014 and 31/12/2023 should provide a constant number of results after the end of 2023, regardless of the concrete day after 31/12/2023. However, this is not the case for all four providers, and the figures vary slightly. In contrast to the numbers on 01/01/2024, for example, a few titles increased on the very next day, i.e. 02/01/2024. In quantitative terms, this means that publications from previous years are still being published online and backdated or at least indexed.

[4] For the exported results on 01/01/2024, please refer to Böck and Ochs (2025).

Table 3 Number of results for each search provider on 01/01/2024

Overview Provider	Number of Results (on 01/01/2024)
IEEE Xplore	1846
ACM Digital Library	316
Elsevier ScienceDirect	1231
pedocs	44
various	27

by looking at Fig. 4. It shows that only a few papers were excluded when the title filter was applied, while most exclusions occurred after applying the keyword filter. In more recent papers, we observed an increase in the use of typical ‘buzzwords’ in titles, such as ‘e-learning’ or ‘digital learning,’ which were captured by our search query. However, examining the keywords provided a clearer understanding of the paper’s content, leading to the exclusion of many. Furthermore, the concepts of digital learning and e-learning have evolved over the period we studied, and they are far more prevalent than they were a decade ago, which could explain why many recent papers have passed the initial retrieval process. Another possible explanation for the increase may be the common phenomenon known as *test of time effect*. Overall, we see the increase in the number of papers on e-learning in general as a clear consequence of the pandemic.

In addition to examining the distribution of papers throughout the years, we also collected domain information for each paper to represent which learners the respective models are constructed for. Although the information was scattered throughout the publications, we classified the results into six groups as follows. Among the publications, more than 60% and, thus, most papers do not specify a certain domain (not specified). The second most referred to is the domain of higher education (25%), i.e. university or college. The remaining domains are less frequently named, with school being the most frequent among them (approx. 8%) and adult learning (approx. 3%), vocational/professional training (approx. 2%) and lifelong learning (approx. 1%) coming after that. This implies that certain domains are underrepresented in the research. Still, at the same time, it shows the versatile character of learner models, which are often not restricted to a certain domain. An overview of the assignment of domains to papers can be found in the “Appendix A”.

4.2 Classification of results

After the analysis, generally applicable findings that apply to all learner models can be extracted. We have found that the various models can be classified according to different dimensions.

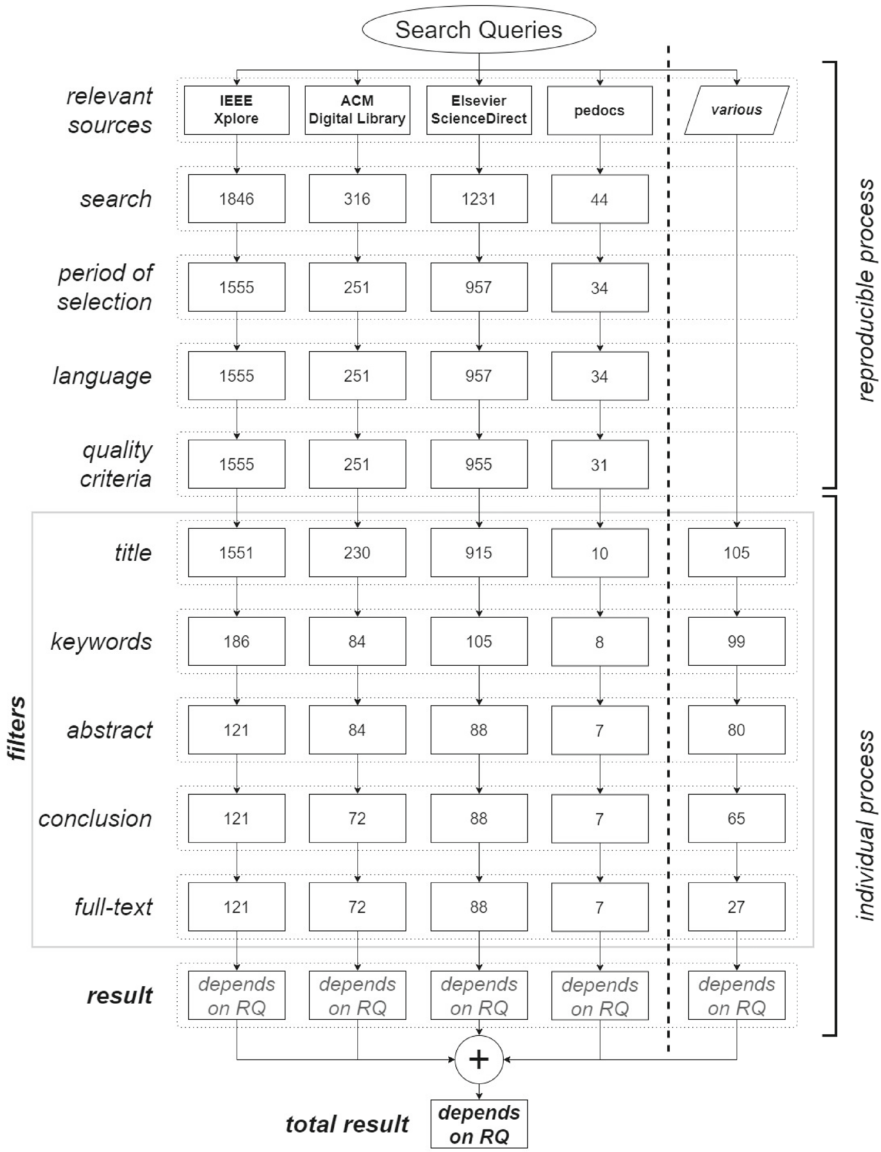

Fig. 4 Search process with filters and number of results for each step and search provider.

4.2.1 Six dimensions

During the data extraction and subsequent analysis and the associated data aggregation, we extracted six dimensions for possible components within a learner model (cf. Fig. 7). Each of these dimensions represents a different way of categorising the possible components within a learner model. All dimensions are equally weighted and equally important. The dimensions can be described simply with six questions:

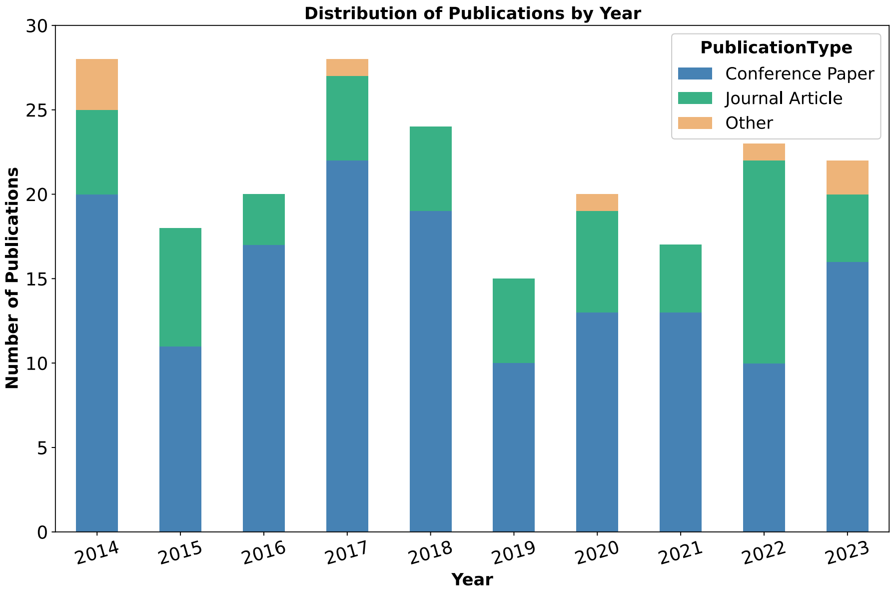

Fig. 5 Relevant publications over time by category—first iteration

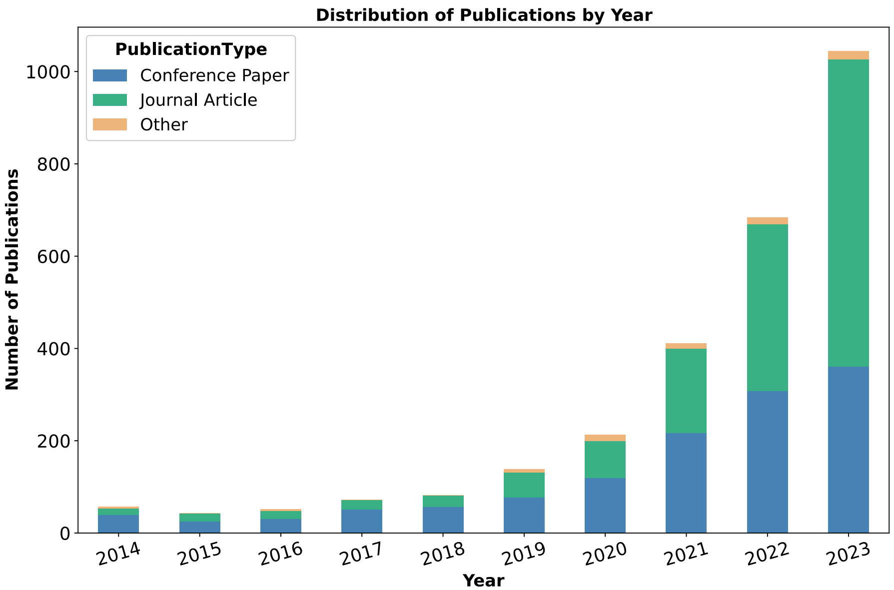

Fig. 6 All publications over time by category—first iteration

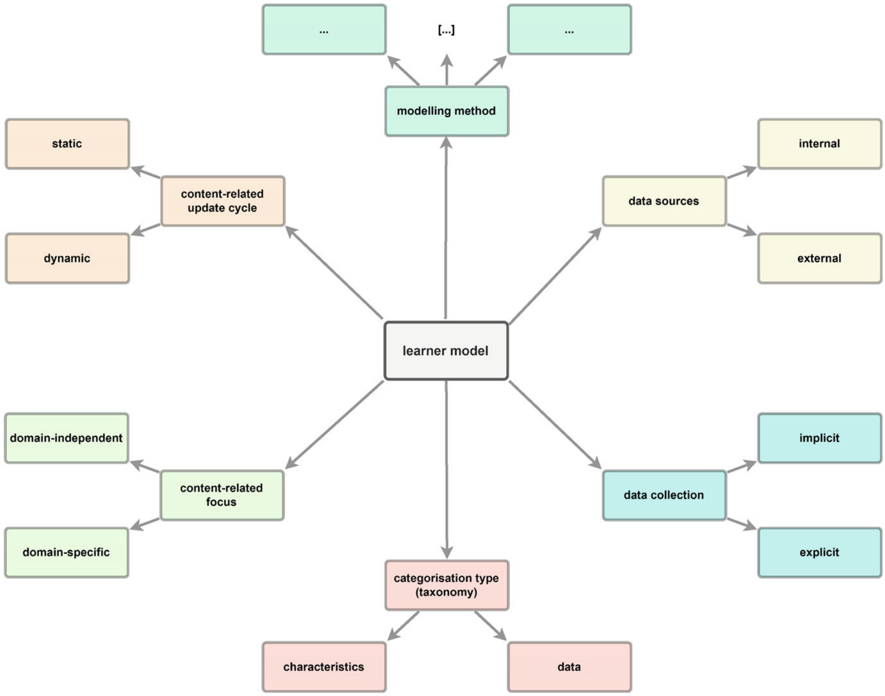

Fig. 7 Overview of the possible dimensions of learner model components.

- Where does the required data come from? (cf. *data sources* in Fig. 7)
- How is the required data collected? (cf. *data collection*)
- How often does the collected data change? (cf. *content-related update cycle*)
- What content does the collected data relate to? (cf. *content-related focus*)
- To which category does the collected data belong? (cf. *categorisation type*)
- How is the collected data represented and modelled in a machine-readable way? (cf. *modelling method*)

In our experience, these six dimensions can be used to describe and assign all the components analysed so far in a learner model. Each of the six dimensions is briefly explained below.

In addition to the six different mentioned dimensions, there can be numerous others, such as categories typically used in the legal domain, i.e. *provided*, *observed*, *derived* and *inferred data* (Abrams 2014), which are commonly referenced by official institutions such as the OECD Working Party on Information Security and Privacy (OECD 2014) or the European Commission (European Commission 2017), but this is a question of perspective, and in this case the legal perspective seemed less suitable to us. Although we did not include the categories in the taxonomy, we show how they relate to our classification in Sect. 4.5.1. In this SLR, we focus on the six categories below that we consider to be relevant and present them briefly before going into detail in Sects. 4.4 and 4.5, respectively.

Data Sources. Data sources describe the location and, thus, the origin of the required or desired data. Here, researchers often distinguish between data that comes from *external* sources outside the in-house system and data that is already available in the system sources (*internal*) and can be processed.
Data Collection. Once the question of where the data comes from has been clarified, data collection focuses on how the data is collected. This can be done either *implicitly*, for example, through interactions in the system (e.g. log files), or *explicitly* by requesting a response from the user. Of course, there are also mixed forms that support both (*hybrid*).
Content-Related Update Cycle. The content-related update cycle dimension describes the change frequency of the attribute values. These can be located on a discretised continuum between *static* and *dynamic* values.
Content-Related Focus. The content focus dimension describes the content-related assignment of the characteristics of the individual attributes. These can be located on a discretised continuum between *domain-independent* (independent of a specific topic to be learned) and *domain-specific* (a specific topic to be learned).
Categorisation Type. The dimension categorisation type defines each component using a specially created taxonomy. This is based on both the data that is collected (*learner data*) and its aggregation (*learner characteristics*).
Modelling Method. The modelling method describes how the individual component is technically modelled and implemented. This will be addressed in Sect. 4.2.3. The questions of where the data comes from (*data sources* dimension) and how it is collected (*data collection* dimension) are explicitly addressed in RQ2 (Sect. 4.5). RQ1 (Sect. 4.4) focuses on the four remaining dimensions that deal with the data.

4.2.2 Exemplary categorisation and use of the taxonomy

Before we take a closer look at the individual dimensions in the following sections, we demonstrate an exemplary categorisation in all six dimensions using one possible component to understand the results better and make it easier to get started. Nakic et al. have explored the ratio of the number of citations of papers on a particular topic and the number of these papers that were published on the topic (2015). Although the ratios show that cognitive skills (24.83%) and personality (22.17%) are increasingly the focus of research, indicating that their potential in adaptive education should be explored further, learning style remains the most frequently cited characteristic (Nakic et al. 2015). In addition, most learner models take a maximum of three learner characteristics into account: knowledge, preferences and learning style (Heidig and Clarebout 2011 mentioned by Alshammari et al. 2014). Therefore, the following paragraphs will use learning style as an example component.

Honey and Mumford define learning styles as "a description of the attitudes and behaviours which determine an individual's preferred way of learning" (1992) as cited by Nakic et al. (2015). In other words, learning styles refer to individual abilities and preferences regarding how a learner perceives, absorbs, and processes learning material (Jonassen and Grabowski 1993). We will now look at the six dimensions separately and assign the learning style to each. If we start with the data source attribute and the question of where the required data to assess the learning style comes from, none of the papers answered this question explicitly, but some information can be deduced from the descriptions. For example, the relevant data for learning usually comes directly from the application and not from external sources (Rahayu et al. 2022). However, the required data will not be obtained exclusively from internal sources; external questionnaires are also used to enrich the internal data, so we categorise the attribute as *hybrid* based on the imprecise description.

The data required to determine a learning style can either be obtained implicitly, for example, by analysing existing learner and behavioural data (Bernard et al. 2022; Aissaoui et al. 2018; Al-Abri et al. 2017; Lwande et al. 2021; Yi et al. 2017b; Mokhtar et al. 2011 cited by Almeida et al. 2018), or explicitly. For this purpose, the learner is interviewed directly, and their learning style is determined based on the results of the questionnaire (Yi et al. 2017a; Gasmi et al. 2021; Almeida et al. 2021; Al-Rajhi et al. 2014; Elghouch et al. 2016; Ishak and Ahmad 2016; Huang et al. 2017; Raj and Renumol 2018; Miranda et al. 2023; Clabaugh et al. 2015 cited by Almeida et al. 2018). Fuzzy logic, decision trees and neural networks are mainly used to automatically identify learning styles from behavioural data (Wang et al. 2025). For this reason, there is no precise categorisation for data collection, and we, therefore, set it to *hybrid* (Ding et al. 2018; Afini Normadhi et al. 2019; Wang et al. 2021; Lin and Wang 2023).

The content-related update cycle of learning styles is very *static* in relation to the current age and learning situation and does not change perceptibly (Heidig and Clarebout 2011 mentioned by Liang et al. 2022; Ayari et al. 2023; Rahman et al. 2023; Chaturvedi and Ezeife 2014; Nguyen 2014b; Abyaa et al. 2019; Cassidy 2004 cited by Aitdaoud et al. 2015; NASSP 1979 cited by Hlioui et al. 2016). Learning style is rarely regarded as a dynamic value in the literature, but there are exceptions (Aissaoui et al. 2018; Abdo and Noureldien 2016). Regarding the content-related focus, the learning style is independent of the content to be learned (*domain-independent*) and can be transferred to all learning situations (Abyaa et al. 2019; Abdennour et al. 2023). The categorisation type in our specially designed taxonomy for possible learner model components is in the learner data area, specifically within the sub-area of *master data* in the *personal traits* category (*learner data > master data > personal traits*)[5]. The modelling method is more complex, as there is no representative in the literature for all learning style modelling. Various approaches model learning styles based on *(dynamic) Bayesian networks*, for example, Al-Abri et al. (2017), Raj and Renumol (2018), Schiaffino et al. 2008 cited by Alshammari et al. (2014) or based on a Markov model (Nguyen 2014b) or as an ontology (Rani et al. 2015; Tarus et al. 2017; Akharraz et al. 2018).

All dimensions are summarised again in Table 4 for a quick overview of this exemplary categorisation. For the learning style component, there are various possibilities of semantic characterisation. Over 70 learning style models have been proposed (Andaloussi et al. 2017). These range from the Felder-Silverman learning style model (FSLSM) (Felder 1988) to David Kolb's learning style model (Yaghmaie and Bahreininejad 2011), to name just a few. The study by Özyurt and Özyurt (2015) showed that Felder-Silverman was the preferred learning style model (42%), followed

[5] The assignment within the taxonomy is based on the many publications analysed and is part of the research and results of this publication (cf. Table 5 for full taxonomy of learner attributes). Whenever we refer to categories of this taxonomy throughout this article, we use italics.

Table 4 Learning styles categorisation in six dimensions

Dimension	Dimension Specification
Data Sources	*hybrid*
Data Collection	*hybrid*
Content-Related Update Cycle	*static*
Content-Related Focus	*domain-independent*
Categorisation Type	*learner data > master data > personal traits*
Modelling Method	*(Dynamic) Bayesian network*

by Kolb's (14.5%), Visual-Auditory-Kinaesthetic/Visual-Auditory-Reading-Writing-Kinaesthetic (10.1%), Mixed (8.7%) and Cognitive (8.7%) styles. Apart from this, the Myers-Briggs-type indicator and models from Honey & Mumford, Gregorc, Keefe and Witkin & Goodenough were used as learning styles in a limited number of studies.

Our aim in demonstrating the classification for the learning style was to enhance the transparency of our approach within a reasonable scope. As part of this work, we extracted and analysed over 180 other possible learner characteristics from the analysed publications. Analysing these individually would exceed the scope of this paper, which is why we have stored the further classifications for other components in an external repository (Böck and Ochs 2025).

4.2.3 Modelling approaches

Numerous works deal with the challenges of modelling learner models, such as the topics of granularity (Pelánek 2025) and the appropriate modelling method for a particular learner characteristic (Pelánek 2017). We summarise the modelling methods we identified in the papers as nine modelling approaches. We have derived them based on two existing classifications of student modelling surveys from different years, i.e. Chrysafiadi and Virvou (2013) and Abyaa et al. (2019) and supplemented them based on the scientific publications analysed by us to enable an up-to-date classification. As Fig. 8 shows, Chrysafiadi and Virvou (2013) distinguish nine categories, and thus, they use a rather fine-grained approach. In contrast, 2019 only distinguish five categories, fusing approaches such as fuzzy modelling and Bayesian networks under uncertainty modelling and refining the broad term ‘machine learning techniques’ that was used by Chrysafiadi and Virvou (2013) into clustering & classification and predictive modelling, reacting to the variety of emerging approaches in this field. As it is still applicable and comprehensive overall, we used Chrysafiadi et al.'s work as a foundation but updated parts with Abyaa et al.'s more recent classification. Also, we specified their broad category of clustering & classification further into two distinct categories to allow distinguishing between unsupervised and supervised approaches, including the traditional method of stereotyping.

Our selection of modelling approaches can be found at the bottom of Fig. 8. While *overlay modelling* defines a subset of the domain or expert model to represent the learner's knowledge (Stansfield et al. 1976), *perturbation modelling* extends it with

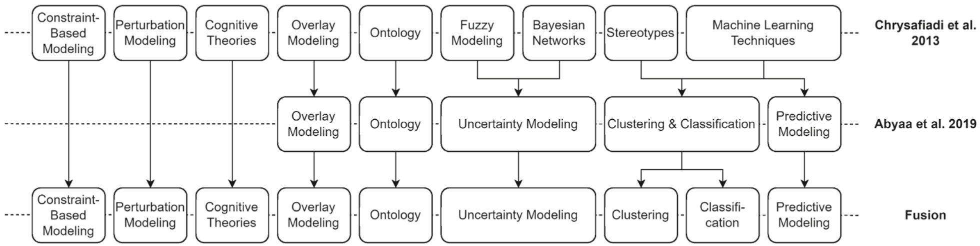

Fig. 8 Overview of the fusion of modelling approaches based on Chrysafiadi and Virvou (2013); Abyaa et al. (2019).

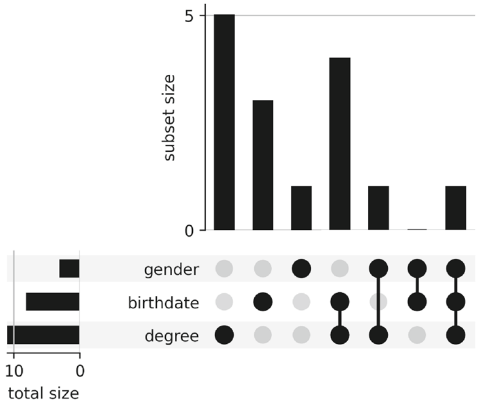

Fig. 9 UpSet plot with example data: Co-occurring learner information in learner models.

the concept of misconceptions (Mayo 2001). *Ontologies* define structured knowledge and are often used for domain modelling; however, we refer to it only as a structured user representation. *Uncertainty modelling* handles uncertainty in knowledge representation, which may, according to Abyaa et al. (2019), include techniques such as fuzzy logic (Zadeh 1965) or Bayesian networks (Conati et al. 2002). *Constraint-based modelling* uses constraints to explain learner behaviour or understanding (Ohlsson 1994). *Classification* approaches group learners based on similarity in a supervised manner, including stereotypes (Rich 1979), while *clustering* approaches do the same in an unsupervised manner. *Predictive modelling* approaches predict possible outcomes based on varying inputs. They may appear in the form of Item Response Theory (IRT) (Hambleton and Swaminathan 1985), or its extensions such as the multidimensional variant (M-IRT) (Reckase 1972) and Diagnostic Classification Models (DCM) (Rupp et al. 2010), machine learning techniques, deep learning (DL) and neural networks such as Recurrent Neural Networks (RNN), Natural Language Processing (NLP) or Bayesian Knowledge Tracing (BKT) (Corbett and Anderson 1994). Last, *cognitive theories* seek to understand the learning process, such as the OCC theory of emotions (Ortony et al. 1988). This also includes computational cognitive models based on cognitive architectures such as ACT-R (Anderson et al. 2004), which is also grounded in cognitive theory. In the coming sections, we will refer back to these approaches.

4.3 Interpretation of results

In this work, we use UpSet plots (Lex et al. 2014) regularly since they are particularly helpful in displaying combinations of data, such as components of the learner models or data sources.

To facilitate the understanding of these plots, we provide a short example in Fig. 9, which describes the information that can co-occur within learner models. The side histogram represents the total frequency of the individual information (e.g. degree) across all possible combinations. In contrast, the upper histogram represents the result frequencies of a specific set of information, e.g. learner models that only hold the

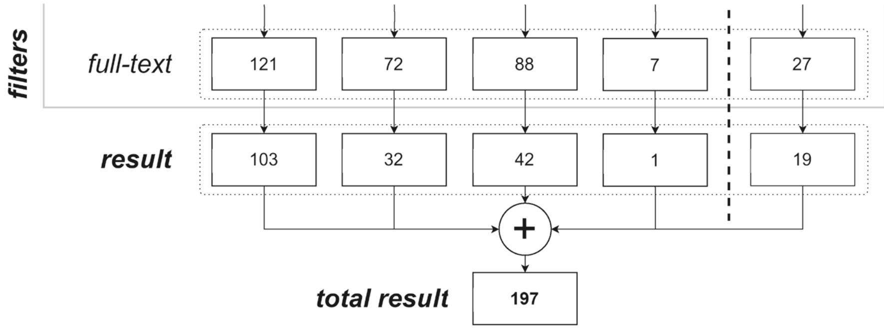

Fig. 10 Search process results for RQ1.

degree (column 1), learner models that only contain the degree and the birthdate but not the gender (column 4), or learner models that contain all elements (column 7). While bullets represent the individual items, lines that connect a set of bullets indicate that several items are combined in the learner model by an n-ary boolean 'AND' operator—and thus are all used to describe the learner by the models. If the readability of the UpSet plot suffers from a large number of combinations, we do not visualise all possibilities but use a threshold to display only the most frequent occurrences. The complete UpSet plot can then be found in "Appendix A". If such a threshold value is used, this is indicated in the respective plot.

4.4 RQ1: Structure and components of learner models

The learning process depends on several individual characteristics, which can also be combined in various ways to influence learning performance (Jonassen and Grabowski 1993). In this section, we are concerned with answering RQ1: *What do learner models look like and what are their structure and components?*[6]. To do this, one must determine, on the one hand, how learner models generally look in terms of structure and composition (*syntax*) and, on the other hand, what the actual meaning of the individual components (*semantics*) is and their effect. In addition to the question of which components are possible, the effect of each component on learning behaviour is also an interesting question but not the subject of this research. We intend to provide the complete spectrum of a maximally developed learner model as a normative component for creating new learner models. As research continues, it is not surprising that even the same authors recommend and develop different sets of their respective learner models over the years (Brusilovsky 1996, 2001; Brusilovsky and Millán 2007; Bull and Pain 1995; Bull et al. 2001; Bull 2004; Kay and Bull 2015; Bull 2020).

As shown in Fig. 10, 197 relevant publications from various sources remained for the answer to RQ1. This is a relative ratio of 62.5% of publications that deal with learner models and address RQ1. The selection of publications concerning the learner

[6] At this point, it should be noted again that in this work, we are not concerned with models in the sense of learning theories or psychological models, but with data models or machine data representations in the sense of the data collected about the learner and the modelling of their specifics.

model's structure, elements and possible modelling approaches was examined in more detail below.

As we see in this subsection, the learner models described are rather superficial and lacking in detail. Many explanations refer to standards that are frequently used and then expand on these with individual aspects (Qiu et al. 2020; Zhao 2021; Wang et al. 2021; Jun-min et al. 2018; Fei Zhou et al. 2017). For this reason, we will first take a closer look at the prevailing standards and then at our overarching findings based on the body of papers guided by the previously introduced classification of results (Sect. 4.2.1) in more detail, including a taxonomy of attributes, for instance.

4.4.1 Standards for learner models

In another study (Böck et al. 2024), we have already dealt with the same research question in the context of existing norms and standards. The aim was to look for a standardised approach to how learner models are created and which sub-components are standardised—regardless of their purpose and use.

We reviewed 868 standards, of which 16 remained for the question of the structure and components of learner models, which were analysed in more detail in the next step. The remaining 16 publications already differ fundamentally in terms of their purpose and the reason for their existence. They range from exchange options for learner data (ISO/IEC 2019, 2016) and their implementation (IEEE 2022) to the specification of learning resources (ISO/IEC 2015), for example for people with cognitive and learning disabilities (W3C 2021), to a complete, comprehensive digital ecosystem of which the learner is a part of (1EdTech 2001; Ed-Fi Alliance 2023). The type of standard also differs, e.g. from generally applicable standards (ISO/IEC 2021) to active standards of consortia (IEEE 2023; JADN 2021) or their pre-development: drafts (IEEE 1997).

Of the remaining standards, three concern themselves more intensively with learner models. In ISO/IEC (2021), a mobile learner model and its specific attributes for learning, such as device, connectivity or location, is described. In 1EdTech (2001), the 1EdTech consortium focuses on the interoperability of internet-based information systems for learners interacting with other systems. The specification of the learner is based on a data model that describes the learner's characteristics required to record and manage the learner's progress, goals, performance and learning experience. In the late 1990s, IEEE P1484.2 already attempted in IEEE (1997) to specify the syntax and semantics of a learner model by centralising public and private learner information (also known as PAPI Learner).

Research on possible norms and standards in the field of learner models has shown that there are few approaches in this area and that these are either specialised (ISO/IEC 2021) or complex and more than just learner models (1EdTech 2001) or have not been pursued further (IEEE 1997). However, no standards have been found that describe a generalised valid learner model that can be extended according to individual needs or that deal with creating such learner models (Böck et al. 2024). The following approaches are among the best-known de facto standards of learner models.

- *IEEE PAPI* (IEEE 1997):
  The draft (IEEE P1484.2, 1997) defines the syntax and semantics of a learner model

by characterising the learners and their knowledge/skills, separating them into public and private information (IEEE Personal and Private Information (PAPI)). The information about the learner is represented in multiple levels of granularity and therefore divided into six categories: (1) PAPI Learning Staff (demographic data), (2) PAPI Learner Relations (relationships with other learners), (3) PAPI Learner Safety (enrolment information), (4) PAPI Leaner Performance (future goals), (5) PAPI Learner Preference (preferences) and (6) PAPI Leaner Portfolio (previous experience). In the early 2000s, the draft was transferred from IEEE SA (IEEE/LTSC) to ISO (ISO/IEC JTC1 SC36) in collaboration with IMS (now 1EdTech), but its presence soon diminished (Böck et al. 2024).

- *IMS LIP* (Smythe et al. 2001):
  This specification addresses the interoperability of Internet-based learner information systems with other systems that support the Internet learning environment. For this purpose, the representation of the learner is specified and stored on a learner information server and made available to other applications. The learner specification is based on a data model that describes the characteristics of a learner that are necessary for recording and managing the learner's progress, goals, achievements and learning experience. The learner characteristics are divided into eleven categories, i.e. (1) identification, (2) goal, (3) qualifications, certifications and licenses (qcl), (4) activity, (5) transcript, (6) interest, (7) competency, (8) affiliation, (9) accessibility, (10) security key and (11) relationship. The standard is based on the notion of a classic CV.
- *CELTS-11* (Li et al. 2002; CELTS 2012):
  The Learner Model Standard (CELTS-11) established by the Educational Informationisation Technology Standard Committee in China covers the basic information of learners, learning information and other aspects of characteristic information (Wang et al. 2012). The Department of Science and Technology in the Ministry of Education has established a new technical Learner Model standard for online education. The later founded Chinese E-Learning Technology Standards (CELTS) committee implemented the technical standards for online education in China in 2002 (Li et al. 2010). The model includes eight categories: (1) individual information (e.g. identification, name, mail, personal information), (2) academic information (e.g. study plan, study completion status), (3) management information (e.g. rewards, penalties, learning events) (4) relationship information (e.g. the relationship between the learner and others), (5) safety information (e.g. security credentials), (6) preference information (e.g. device preference, emotion, interest), (7) performance information (e.g. issue object, learning experience, learning ability, granularity) and (8) portfolio information (e.g. accomplished work and projects of the learner).
- *GB/T 29805-2013* (GB/T 2014):
  The national standard of China known as GB/T 29805-2013 Information Technology—Learning, Education and Training—Learner Model describes a model that divides the subset of learner information into eight categories of information: (1) personal information, (2) academic information, (3) management information, (4) relationship information, (5) security information, (6) preference information, (7) performance information and (8) portfolio information. This stan-

dard applies to building a database of learner information for different learner characteristics, migrating and exchanging learner information between different systems, improving the portability of learner information data and ensuring the security of user data.

- *eduPerson* (Internet2 2003):
  eduPerson is an LDAP schema incorporating commonly used personal and organisational attributes relevant to higher education. In the current version (202208) v4.4.0 (MACE 2022), 16 different attributes can be assigned to a learner. There is no official categorisation of the attributes, but they can be divided into (a) general attributes that contain information about the learner, such as name, and (b) learner-specific attributes, such as the learner's affiliation and relationships.
- *FOAF* (Brickley and Miller 2014):
  The FOAF language is defined as a dictionary of named properties and classes to link people and information using the Web by which users can retain some control over their information in a non-proprietary format. The specification summarises three types of networks: social networks involving human collaboration, friendship and association; representational networks that provide a simplified, factual view of a cartoon universe; and information networks that utilise web-based linking to share independently published descriptions of this interconnected world.

Although these models have a set of common learner characteristics, they differ in their primary purposes and how systems utilise their embedded information. It is common practice to create a learner model for a learning system by integrating various learner model standards to leverage their unique advantages. Ouf et al. (2017) compare the standards mentioned and conclude that the IEEE PAPI and IMS LIP standards focus on the performance of the learner without taking into account the goals and purpose of education, which is important for individualised education. EduPerson is considered the best standard for data exchange between educational institutions, while FOAF represents relationships between learners but does not collect data. Both eduPerson and FOAF neglect the learning style, which is considered essential for personalised e-learning systems. None of the standards mentioned (IEEE PAPI, IMS LIP, FOAF, eduPerson, CELTS) take learning time into account, which is important for accurate performance assessment. IMS LIP integrates interests, while the other standards do not, although they are crucial for choosing appropriate learning activities.

Hlioui et al. (2016) also compare the most common standards and conclude that they are not complete enough to cover all learner data that can be exchanged by e-learning systems (Oubahssi and Grandbastien 2006 mentioned by Hlioui et al. 2016). They note that PAPI is focused on administrative data at the expense of preferences, competencies and academic performance. The IMS LIP standard is more of an extension of the PAPI model (Lazarinis et al. 2009 mentioned by Hlioui et al. 2016). The elements are based on the concept of a classical CV (Panagiotopoulos et al. 2012 mentioned by Hlioui et al. 2016) so that elements such as learning objectives, behaviour, preferences and the collaborative aspect of learners are not considered.

However, numerous in-house developments of learner models build on the standards already mentioned and attempt to eliminate their disadvantages or risks with extensions and adaptations (Ghallabi et al. 2015 and Zghibi et al. 2012 and Panagiotopoulos et al. 2012 and Hadj-M'tir 2010 mentioned by Hlioui et al. 2016). Learner information is stored in a customised, standardised format based on IEEE PAPI and IMS LIP and extended by performance data such as grades and learning progress (Rahayu et al. 2022). Others use specifications that already include more than basic learner data, such CELTS-11 (Zhao 2021; Wang et al. 2021; Jun-min et al. 2018; Fei Zhou et al. 2017) or GB/T 29805-2013 (Qiu et al. 2020), which includes basic personal information about the users, their learning style, level of knowledge and other user characteristics.

In addition to the numerous learner model standards, there are many scientific publications from scientific conferences and journals. The publications counted in Fig. 10 range from ideas to approaches that have already been evaluated. In the following, we compare and categorise the different concepts.

4.4.2 Taxonomy of attributes

From the analysed literature, we have extracted and summarised the most common attributes of a learning model, as far as known, including their modelling approach. The more frequently an attribute is used in the learner models, the greater its impact seems to be, which would explain the frequency. Attributes are, for example, name, age, origin and the current level of knowledge of a domain, as well as previous education and, if already available, behavioural data in a digital system. The first difficulty is that not every attribute is interpreted similarly. We, therefore, catalogue the attributes under different categories. The data can thus be divided into categories such as *master data* (*demographic & biographic data*, *academic/educational data*) and also *cognitive characteristics* or *affective & motivational characteristics*. The aim of the taxonomy is to characterise the various components of a learner model to highlight similarities between the components. Another alternative would be characterising the components based on their impact on learning. However, this requires great effort due to the numerous studies and evidence, so we have chosen the first categorisation. We have organised this categorisation into a specially developed taxonomy to better understand the individual features and make it easier to understand the results and get started. To enhance clarity and maintain a concise presentation, we have intentionally excluded detailed discussions of individual works due to the extensive scope of the analysed literature. Instead, we describe the categories in general terms, referencing commonly cited examples from the reviewed publications. A differentiated list, as well as detailed information on the individual publications analysed, can be found in the Appendix and also in the repository which was created for this specific purpose (Böck and Ochs 2025).

Table 5 extends the categorisation to include all attributes found within a learner model and their assignment to the taxonomy. It shows that many attributes can be found in the area of *cognitive characteristics*, closely followed by attributes from the *personal traits*, *academic/educational data* and *contextual data* categories; while others, such as *physiological data*, occur less frequently, as these are often more difficult to collect and model. The table already provides an initial overview of the variety of possible attributes of a learner model and visualises the distribution of the various attributes in their respective categories.

Table 5 Overview of the various possible attributes within a learner model based on the examined publications. Classification according to the main categories and subcategories of our taxonomy—learner characteristics

Learner characteristics	
cognitive characteristics	cognitive ability, cognitive level, cognitive style, cognitive deficits, attention, concentration, (work) memory, knowledge/prior knowledge/foreknowledge, intelligence level, competence, proficiency, skills, ability level, ability to understand and learn, analysing abilities, comprehension, analysing and reasoning, abilities to solve problems, critical thinking, abilities to make decisions, communication skills, collaborative skills, creation, information processing abilities, phonological processing, orthographic processing, lexical access, perception, perceptual style, beliefs, learner's credibility, objectives
metacognitive characteristics	self-awareness, self-management, self-monitoring, self-assessment, self-regulation, self-explanation
affective & motivational characteristics	mental state, engagement, effort, commitment, volition, motivation, user intentions, expectations, learner opinion, desires, emotions, feelings, sentiments, mood state, self-esteem, satisfaction, flow/immersion, anxiety

Most attributes can be modelled in different ways. We therefore present all the options found in an unprioritised manner. The general categories of the individual features are briefly described below, and the most important findings are summarised. The taxonomy is divided into two main sections: *learner characteristics* and *learner data*.

Learner Characteristics are the descriptive properties derived by analysing learner data and can be divided into *cognitive characteristics*, *metacognitive characteristics* and *affective & motivational characteristics*.[7]

Cognitive Characteristics. The term cognitive refers to learner functions associated with perception, learning, remembering, thinking and knowledge. The three most common categories of cognitive characteristics in the literature are cognitive styles, abilities and deficits. In this context, the topics of attention and concentration or working memory are often mentioned, but the topics of knowledge and competencies are also highly present. Starting with individual skills and their ability level through to pure knowledge, which can be divided into prior knowledge[8] and newly acquired knowledge.

Metacognitive Characteristics. Metacognition is a term from psychology and neuroscience that describes learners' ability to reflect on their thought processes and scrutinise decisions. This includes skills such as self-awareness, self-management with self-monitoring, self-assessment and self-regulation.

[7] These three categories could be subdivided even more finely, but the clear-cut classification of the different attributes is then not always necessarily given, so we deliberately do without the intermediate level in this case.

[8] Some authors also use other synonyms for prior knowledge, such as *foreknowledge*

Table 6 Continuation of Table 5—learner data

Learner data		
master data	personal traits	personality, interests, intentions, ambitions, learning targets, destination, goal setting, learning purpose, learner preferences, learning preferences, learning styles, daily habits, learner's abilities, difficulties, helpfulness, health conditions, disability, prerequisites, user needs, specific needs, mobility needs
	demographic & biographic data	name, birthdate, gender, address, contact, tongue language, ethnicity, nation, origin, cultural, cultural-linguistic heritage, affiliation, belonging, occupation, location information
	identification data	user metadata, credentials, username, password, (u)id, email, enrolment number, user type
	academic/educational data	degree, level of education, certifications of completion, academic achievement, any relevant certifications or training, discipline, proficiency, profession, vocational aptitude, qualifications, expertise, experience, languages, phase of studying/school stage, enrolment, learning plan, previous courses taken, generalisation of biases
(learning) output data	(learning) behaviour data	confidence, activity plan, learning path, progress, activities, user records, consumption level, consumption amount, consumption frequency, task engagement, successes, failed courses, errors, error categories, misconceptions/contradictions
	learning resource/assessment data	biases, assignments, scores, grades, quality of results, problem-solving time, completion rate, click-through rate, stay time
(social) interaction data	social content data	learner's contribution, open data, social ties, projects, groups
	interaction data	interactions with the system, social behaviours, degree of collaboration, communication, relationships, roles of relations, reaction time
(system) environmental data	contextual data	socio-technical context, social context, learning situation, ecosystem, environment, cultural, device, device parameter, mobile sensors, physical location, geographical location, time, screen size, background noise, calendar, appointments
	technical usage data	activity logs, events, browsing history
psychomotor & physiological data	physiological data	eye tracking, heart rate, skin response, brainwave
	psychomotor data	physical movement, physical activities, motor skills, coordination

Affective & Motivational Characteristics. Affectivity summarises the entirety of the emotional and mental life, which comprises mood, emotion and motivation. In the literature, motivation and engagement are often equated and mentioned along with effort, commitment, and volition. However, the expectations and desires of a learner also play an important role, as both topics influence the emotions and mood state of the individual. The spectrum ranges from satisfaction to anxiety.

Learner Data is the actual, specific information collected and analysed to determine these learner characteristics. The learner data category is subdivided into five main sections. Each of these five main categories has at least two subcategories.

Master Data. *Master data* covers basic information about a learner that is necessary for their identification and management in a system. This includes *demographic & biographic data* that describes a learner's structural characteristics and personal life stories, as well as *personal traits* that define character traits and behaviour. This is supplemented by administrative data (*identification data)*, which is used specifically for identification and management in administrative processes, and *academic/educational data*, which relates to a learner's educational background, achievements and learning processes.

Demographic & Biographic Data. Demographic data includes information about a learner's structural characteristics. It is enriched with biographical data, which refers to personal information about a learner's life story or career without their educational background. The learner's name, gender, date of birth, place of origin, address and other contact details are often used as personal information. In addition, this data is optionally expanded to include the learner's mother tongue, cultural aspects, affiliations, occupation, and other information about the learner's professional life.

Personal Traits. Personal traits data refers to a person's character attributes, behaviour and personality traits. Interests and ambitions are often reflected in learning targets and goal setting. Important for subsequent personalisation and, above all, individualisation is information about learners and learning preferences, such as learning styles, so that the learning process can be selected as appropriately as possible. This also requires information about abilities and difficulties.

Identification Data. Identification data, specifically management or administrative data, is specific information used to identify and manage learners within an administrative system. Most of the data used here is metadata about the learner (such as a unique ID), as well as their user data (for example, email, enrolment number, user type) and credentials (username, password).

Academic/Educational Data. Academic/educational data is information related to the learners' educational path, achievements and learning processes. The most recent educational degree and the associated current level of education are important for getting started. Certificates of completion, academic achievements, and relevant training that have led to this help to understand the learner better. The previous learning plan and the courses taken also provide information about the learner.

(Learning) Output Data. *(Learning) output data* covers information that records and analyses the behaviour of learners during the learning process and their interactions with learning materials and assessment tools. This data captures how learners use the resources provided and perform in any assessment form.

Learning Resource/Assessment Data. Learning resource/assessment data is information about learning materials and assessment tools used in the educational process. All learning elements, such as tasks started and their results are stored, regardless of whether they have been successfully completed, along with other event data such as problem-solving time, completion rate or click-through rate.

(Learning) Behaviour Data. (Learning) behaviour data records and analyses learners' behaviour during the entire learning process. This is data that records the overarching behaviour across several learning elements and between different learning elements. The current learning path can be determined from the stored activity plan of the completed activities and their progress. In addition, generic information such as mistakes made and misconceptions are also saved.

(Social) Interaction Data. *(Social) interaction data* covers information about the content that learners create and share on social networks and platforms, as well as the interactions and activities that result from this content. This data captures both the semantic content and the behaviour of learners in social environments, providing a comprehensive picture of their social interactions and engagements.

Social Content Data. Social content data refers to the semantic content created and shared by learners on social networks and platforms. This includes learner contributions—for example, in groups, projects or individual work—and all open data freely accessible to a learner on the internet.

Interaction Data. Interaction data is information that results from learners' interactions, activities and behaviour in social networks and platforms. Interactions with the system and social behaviour within the system are stored. These usually address questions such as: Who is communicating with whom, what are the learner's relationships with others, and what is the level of collaboration?

(System) Environmental Data. *(System) environmental data* covers information that describes the specific context and the technical requirements of the system under which interaction or use of a system occurs. This includes both *contextual data*, which defines the environment and circumstances of a situation, and *technical usage data*, which records the technical behaviour and use of systems, applications or devices independently of the learner.

Contextual Data. Contextual data, also known as context-aware parameters, describes the specific framework or circumstances surrounding a particular situation or interaction. This data provides additional details about the context in which an action occurs or a decision is made. This context can be socio-technical and address the circumstances under which learning occurs: Is it standard online learning or emergency remote teaching (ERT)? Therefore, the ecosystem in which learning takes place plays a major role, as do the learner's end device and its characteristics.

Technical Usage Data. Technical usage data comprises information relating to the technical behaviour and use of systems (user-independent), applications or devices. This technical data includes events and activity logs and, for example, the browsing history within the learning platform.

Psychomotor & Physiological Data. *Psychomotor & physiological data* cover information about a learner's physical and biological functions of the body (*physiological data*) and the interaction between mental processes and motor skills (*psychomotor data*).

Physiological Data. Physiological data refers to information that describes a learner's physical and biological functions. This includes data such as heart rate, skin responses, brainwaves and even eye tracking, which is recorded via webcams or other sensors.

Psychomotor Data. Psychomotor data relates to information that describes the interaction between a learner's mental processes and motor skills. This can be, for example, physical movements or activities.

Evaluation of Components and their Dependencies

The mapping of possible attributes of a learner as components of the learner model to a category of our taxonomy often does not allow a clear assignment, as many attributes are context-sensitive. For instance, behavioural data can be found in "logs" in general, but also in the category *(learning) behaviour data* about learning data or *interaction data* in the social environment. If the reference is not directly apparent from the context, we have mapped it for all meaningful possibilities. The complete figure with all possible attributes shows the various combinations and possibilities of individual attributes (Böck and Ochs 2025).

Analysis of the Main Categories A closer look at the distribution of the learner characteristics and learner data used in learner models in the literature is shown in Fig. 11[9],[10]. A total of $n = 129$ out of 197 works dealing with RQ1 provide detailed information about the components of their learner models, which makes a relative share of 65%. Only 14 of 129 focus exclusively on learner characteristics (11%). We will only look at the top categories from Table 5—column 1. Looking at Figure 11, it is immediately noticeable that the combination of *(learning) output data, cognitive characteristics and master data* ($n = 14$) occurs most frequently. This is followed by the two individual categories *cognitive characteristics* ($n = 12$) and *master data* ($n = 9$), as well as their combination ($n = 9$). The combination of *(learning) output data and cognitive characteristics* ($n = 7$) also occurs frequently in relation to the other possible combinations. The relevance of these main categories becomes apparent when looking at the absolute number of occurrences: 70% of the analysed works contain *cognitive characteristics* ($n = 90$), closely followed by *master data* ($n = 84$) with 65% and *(learning) output data* ($n = 76$) with 59%. Therefore, combining these maximum frequent items forms the maximum frequent item set *(learning) output data, cognitive characteristics and master data*. Most learner models (88%) consist of a mix of pure data (such as *master data*) and characteristics derived from it (such as *cognitive characteristics*).

Figure 12 shows the main categories of the components in learner models over time. Looking at the distribution over the last ten years, no clear trend can be identified as to which of the categories has increased or decreased in relevance. What can be said, however, is that the categories of *cognitive characteristics* and *master data* have

[9] To keep the visualisation of item combinations manageable and as clear as possible, only those which appeared at least twice have been included into this plot. A complete version of the plot can be found in "Appendix Fig. 20".

[10] Combinations that appear only once have been excluded from the plot for the sake of readability. Therefore, the frequencies on the left-hand side histogram differ slightly from those presented above. A complete version of the plot can be found in "Appendix Fig. 20".

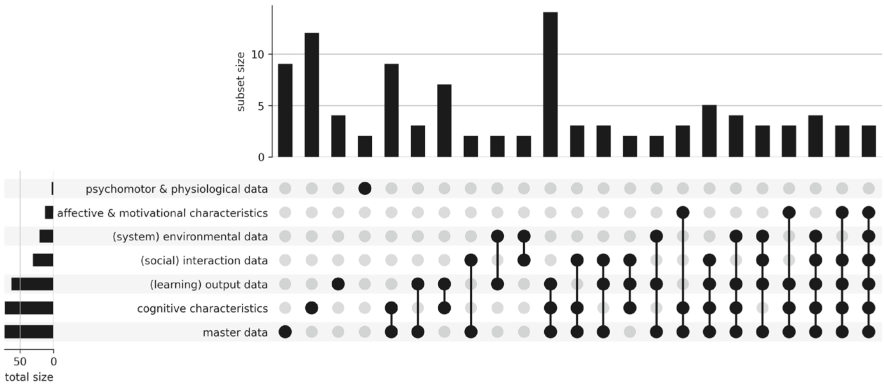

Fig. 11 Frequency of use of (combinations of) main data categories for components of the learner model. Overall source frequency—left histogram; frequency of combination of sources—top histogram.

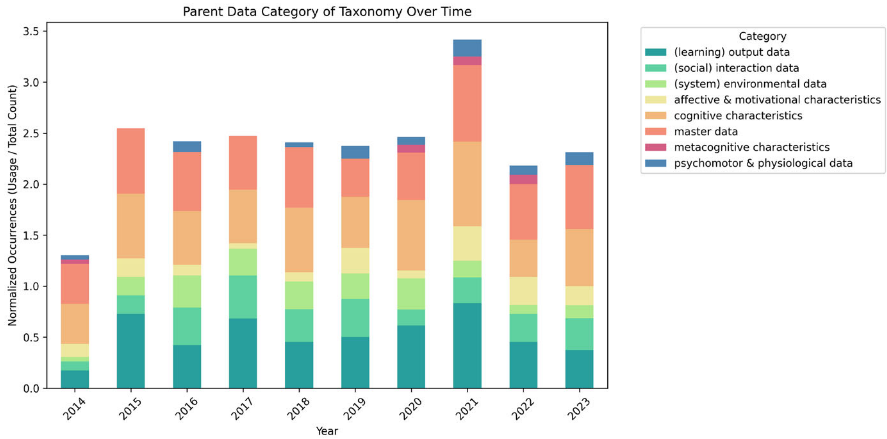

Fig. 12 Main data categories for components of the learner model over time (occurrences have been normalised as the count of papers per year varies.

been consistently important over the years and have accounted for the largest share. *Metacognitive characteristics*, for example, are weakly represented and can only be found in individual years, such as 2014, 2020, 2021 and 2022. Similarly, *psychomotor & physiological data* could also not be found in all years; corresponding publications were missing in 2015 and 2017.

Analysis of the subcategories While the main categories of the taxonomy were analysed in the previous part, this part takes a closer look at the subcategories. It can be seen that *personal traits* ($n = 70$) occurs most frequently with 54%, closely followed by *(learning) behaviour data* ($n = 61$) with 47%, *demographic & biographic data* ($n = 41$) and *interaction data* ($n = 39$) with 32% and 30% respectively. Learner models with purely *identification data* do not occur at all (this is also the reason why it is not shown in the plot) and as a subset in only 6% ($n = 8$) of the analysed publications. The proportion of *learning resource/assessment data*—abbreviated in

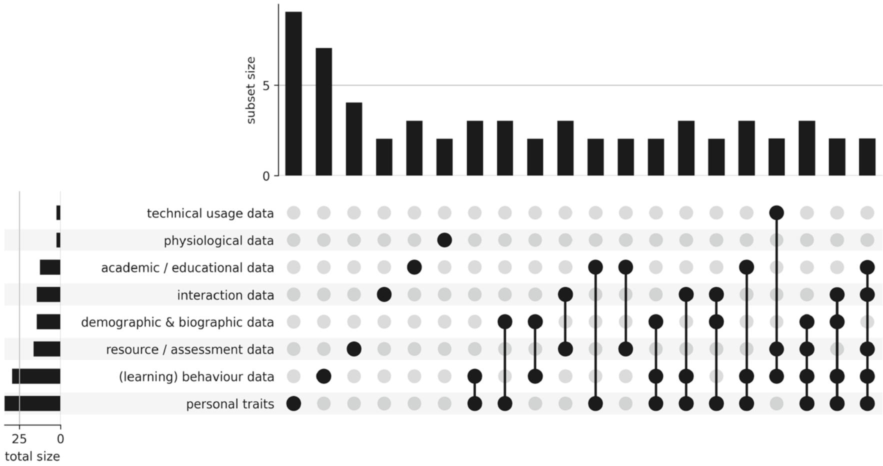

Fig. 13 Frequency of use of (combinations of) data categories for components of the learner model. Overall source frequency—left histogram; frequency of combination of sources—top histogram.

the plot as *resource/assessment data*—is surprisingly tiny at only 26% ($n = 33$). This could be related to the phenomenon that the insights gained so far from the learning data alone are insufficient and, therefore, enriched with additional attributes addressing other categories of learner data and characteristics. The various combinations of data categories occur relatively rarely and amount to an average of 2% ($n = 2$). Figure 13 shows the subcategories. For better readability, the data was standardised and then normalised, so that occurrences below two were removed. This process reduces the content to be displayed in the graph to the essentials.[11]

Looking at the development of the data categories over the last few years, as shown in Fig. 14, no general conclusion can be drawn from the combined total of all data categories over time. However, when the categories are considered individually, it can be seen that *personal traits* and *(learning) behaviour data* make up the most significant proportion of data per year on a relatively constant basis. Yet, not all categories are consistently addressed in the publications. For instance, *identification data*, *physiological data*, and *psychomotor data* are missing in some years.

For the sake of completeness, Fig. 15 shows the distribution for the main categories and subcategories of our taxonomy in relative and absolute numbers.

4.4.3 Content-related update cycle

The *content-related update cycle* addresses how often the content of the individual attributes or components changes and should, therefore, be updated. The change frequency of the content of the individual attribute values is on a discrete continuum between *static* and *dynamic* values. Classification is usually done automatically based

[11] Combinations that appear only once have been excluded from the plot for the sake of readability. Therefore, the frequencies on the left-hand side histogram differ slightly from those presented above. A complete version of the plot can be found in "Appendix Fig. 21".

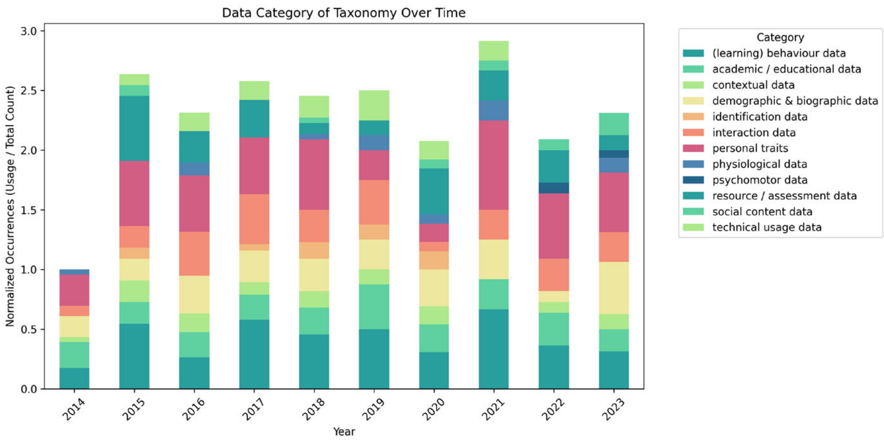

Fig. 14 Data categories for components of the learner model over time (occurrences have been normalised as the count of papers per year varies.

```
learner model attributes
├─learner characteristics (71% | n = 91)
│ ├─cognitive characteristics (70% | n = 90)
│ ├─metacognitive characteristics (3% | n = 4)
│ └─affective & motivational characteristics (18% | n = 23)
└─learner data (67% | n = 89)
  ├─master data (65% | n = 84)
  │ ├─demographic & biographic data (32% | n = 41)
  │ ├─personal traits (54% | n = 70)
  │ ├─identification data (6% | n = 8)
  │ └─academic / educational data (27% | n = 35)
  ├─(learning) output data (59% | n = 76)
  │ ├─(learning) behaviour data (47% | n = 61)
  │ └─learning resource / assessment data (26% | n = 33)
  ├─(social) interaction data (33% | n = 42)
  │ ├─social content data (6% | n = 8)
  │ └─interaction data (30% | n = 39)
  ├─(system) environmental data (24% | n = 31)
  │ ├─contextual data (13% | n = 17)
  │ └─technical usage data (13% | n = 17)
  └─psychomotor & physiological data (9% | n = 11)
    ├─physiological data (8% | n = 10)
    └─psychomotor data (2% | n = 2)
```

Fig. 15 Tree visualisation of this article's taxonomy of learner model attributes, structured as a hierarchy of main categories and subcategories. The distribution is presented in relative and absolute numbers.

on the attribute. Frequent static values are created at the initialisation of a new learner and retain the same value over a long period, such as name, date of birth, gender or mother tongue. Very dynamic attributes are usually related to the learning progress, such as learning targets or progress. However, the change frequency cannot be generalised and transferred to other components; each component must be considered individually. The content of some attributes can also change over time from more static to more dynamic attributes and vice versa. In addition, the attributes can also differ between learner models depending on the interpretation and use case: for this reason, the definition of the time in which an attribute value changes is decisive. Relevant literature on this topic also reveals a distinction in the type of data collected, whether it is predominantly more static (Chignell and Hancock 1988; Jeremić et al. 2012) or more dynamic (Germanakos and Belk 2016; Tadlaoui et al. 2019). However, as these works already imply, a clear assignment on such a continuum is hardly feasible and complex, and any generalised assignment would not meet our quality standards. For this reason, we have not assigned specific categories to publications regarding this aspect.

4.4.4 Content-related focus

The *content-related focus* deals with the semantics of the content of the individual attributes or components and their relationships to specific learning content or general content. Attribute values can be classified on a discretised continuum between *domain-independent* (independent of a specific topic to be learnt) and *domain-specific* (a specific topic is to be learnt). Domain-independent attributes store general information about the learner, independent of the semantic content being learnt. This can include name, contact information and personal data. Domain-specific attributes relate to the content to be learnt and depend on it, such as motivation, interest and goals. Referring back to the example from Fig. 1, recursion as a subject is a specific learning content in some modules and, therefore, represents domain-specific content. However, the content-related focus cannot be generalised and transferred to other components; each component must be considered individually. The content of some attributes can also change over time, namely from more domain-specific to more domain-independent attribute values and vice versa. In addition, the attributes can also differ between the learning models depending on the interpretation and use case. For this reason, the categorisation can only be considered at one point but may change over time. Similar to the content-related update cycle, providing a clear classification in our taxonomy is complex and of limited value, as there are few instances that can be exclusively labelled as either domain-independent or domain-specific. For this reason, we have decided not to assign categories to the examined publications regarding this aspect.

4.4.5 Modelling method

Only 24% of the publications (48 out of 197) that deal with the possible components of a learner model touch on modelling these components or treating them superficially. Often, only the various modelling techniques are listed, without more precise reference to the individual attributes or their differences (Liang et al. 2022; Sani and Aris

2014a, b; Lei and Mendes 2021; Kurup et al. 2016) or roughly characterised according to the types of attributes. For example, structural modelling is attributed to vectors of features or stereotypes and behaviour modelling to classification systems or semantic technologies (Pucheta et al. 2022). There are different modelling approaches for the individual attributes. It often depends on the application and cannot always be generalised. Knowledge can be modelled with the help of overlay models (Kaliwal and Deshpande 2022; Abdi et al. 2021; Beyyoudh et al. 2018; Iqbal and Nurjanah 2016; Al-Rajhi et al. 2014) or with other methods such as fuzzy logic rules to update the knowledge levels (Lei and Mendes 2021; Chrysafiadi et al. 2020; Wang et al. 2021) and Bayesian Knowledge Tracing (Yudelson et al. 2013; Huang et al. 2016; Kurup et al. 2016; Oeda and Kakizaki 2022; Huang et al. 2016; Lei and Mendes 2021; Faucon et al. 2020) or (Dynamic) Bayesian Networks (Raj and Renumol 2018; Almeida et al. 2018; Alday 2018; Kavitha et al. 2018; Ferreira et al. 2017; Ramírez Luelmo et al. 2020). Stereotypes are often used for learners' common attributes (Iqbal and Nurjanah 2016), but ontologies are also becoming increasingly popular for such attributes (Guettat and Farhat 2015; Luna et al. 2015; Qinghong et al. 2015; Rahayu et al. 2022; Nurjanah 2018; Zhao et al. 2017; Wagner et al. 2014). An overview and categorisation of the most widespread modelling approaches can be found in Sect. 4.2.3 on general insights.

4.4.6 Discussion of results

To conclude this subsection, we summarise some of the most striking aspects that emerged from the literature review. The study aimed to find the learner information and characteristics used in adaptive e-learning systems. Based on the papers reviewed, the following gaps were identified and informed the design of the proposed framework through design criteria.

The analysis of the relevant scientific publications has shown many different approaches to how learner models are composed. Only 129 out of 197 (65%) of the publications analysed deal with the question of a learner model's possible components and structure (RQ1). The level of detail ranges from very superficial to enumerations of many different details and possibilities. This abstract view is also reflected in Figures 11 and 13, which show the distribution of the different categories of possible learner attributes. During our large-scale research, we observed that many of the analysed publications incorporate learner characteristics (e.g. learning style) into their models; however, they do not specify why, how or which reproducible influence these may have on learning or whether it makes sense to use them as parameters for adaptation. In addition, we could not find clear guidance or recommendations on which combinations of learner characteristics are reasonable or useful to incorporate into a learner model in the publications. No general recommendation can be derived from the systematic analysis of the statistical data collection either. However, it can be said that the combination of *(learning) output data, cognitive characteristics and master data* occurs most frequently at around 11%, closely followed by the respective individual categories or their combination (cf. Fig. 11). Generally, it can be stated that *learner characteristics* occur in 71% and *learner data* in 67% of the publications analysed. The most frequently mentioned data categories are *personal traits* (54%), closely followed by *(learning) behaviour data* (47%) and *demographic & biographic data* (32%). In the literature, the same learner characteristics are often used and supplemented by a few more exotic ones (sensor measurement values like Vildjiounaite et al. (2023); Muldner and Burleson (2015)): administrative data are usually included—such as name, gender, age—which is enriched with behavioural data (*interaction data* in the system) and learning data (knowledge about a topic). The modelling of this data is also not uniform but highly diverse in the analysed literature. For example, knowledge characteristics can be modelled using overlay models, knowledge tracing or fuzzy logic. This variability presents developers of learner models with major challenges in decision-making options and subsequent feasibility. In addition, there is, unfortunately, a lack of information in the analysed publications about where exactly the data comes from for the individual components, how it is processed and what the concrete implementation of the individual components of a learner model looks like. Often, these topics are only touched on in the publications analysed for the individual components and are considered a marginal issue or even omitted. This lack of presence in the publications is not consistent with the importance of the topic. This makes it all the more difficult to obtain specific, in-depth information for creating an optimised learner model, so it can be recreated as accurately as possible. Few publications deal with the efficiency of individual learner characteristics such as Effenberger and Pelánek (2021). Our analysed publications show a largely representative sample in this context, as missing information is generally rare in the literature. However, we have not yet examined in more detail the interaction of these individual components and the modelling of the complete learner model. This will be discussed in the following section.

4.5 RQ2: Data sources, data collection and modelling approach

After diving into the structure and components of state-of-the-art learner models, a subsequent question arises: *Which data sources are used to initialise and update learner models, how is the data collected, and what is the overall modelling approach?* (RQ2). Following the structure of this question, this subsection will—after a short introduction into the central concepts that underlie this question—first explore the data sources that are commonly used when populating a learner model with data (Sect. 4.5.2) and updating such a model with the collected data (Sect. 4.5.3) and will then examine what overall modelling approach is used (Sect. 4.5.3). It, therefore, refers to the data sources and data collection dimensions from the previously presented classification (cf. Sect. 4.2.1).

4.5.1 Definitions of terms

Before examining the steps of initialisation and updating, the fundamental terms used in this section—i.e. data source, data collection method and modelling approach—will be explained, and examples will be given for each of them. First, speaking of a *data source*, this research refers to the origin of the data that is used to populate the learner model (*Where does the data come from?*). The types of data sources used in this research are:

- **System:** The environment where student interaction occurs. This could be a Learning Management System (LMS), Adaptive Learning System (ALS) or any other educational platform.
- **External source:** A data origin separate from the primary system. This could be an external database, a third party service or any other data repository consulted for additional information.
- **Self-report:** Learners report through surveys, questionnaires, feedback forms, or personal reflections (e.g. e-portfolios, journals).
- **Sensors & wearables:** Devices that capture data about the student's physical state or environment, such as motion or temperature, and heart rate monitors.
- **Assessment:** A tool used to evaluate student performance, such as quizzes, tests or assignments, which may but does not have to be embedded into a system.
- **Observation source:** A data origin characterised by observing students' actions or behaviour, such as cameras.
- **Third party:** A party (person) that is not the student but an instructor, peer, expert or parent.

Second, the *data collection method* represents the way or technique of how the model is populated with data (*How is the data collected?*). The data collection methods this research examines are:

- **Automated collection:** An automated way of gathering data such as tracking, analytics or logging, usually via a system, API or a sensor, without requiring direct user input.
- **Observation:** A way of collecting data by observing learner behaviour in real-time (classroom observation) or via recorded sessions.
- **Direct interaction:** A direct way of engaging with the learners in data elicitation such as interviews or discussions.
- **Assessment-based:** A performance-related way of eliciting learner data such as quizzes, projects, tests, assessments or other tasks.
- **Self-reported:** A data collection approach that refers to learners providing relevant information about their experiences, behaviour or reflections, typically through surveys, feedback forms or diaries.
- **Third party reporting:** A data collection approach that refers to gathering data via other parties than the learner, providing insights from instructors, parents or peers.

These data collection techniques can—in alignment with recent approaches (Pucheta et al. 2022; Andaloussi et al. 2017; Alshammari et al. 2014)—be subsumed under the terms *explicit*, *implicit*, or the combination of both, which is usually referred to as *hybrid* or *mixed*. When it comes to explicit methods, the user is involved in the data collection, and the data collection is, therefore, noticeable. Common examples for this case are any kind of reporting, direct interaction, or assessments, but observation may also be considered an explicit method in some cases (e.g. observer present in the classroom). In contrast, implicit defines the usually unnoticed collection of data, which happens—in most cases—through automated collection via a system, sensor, or device. Hybrid approaches combine both methods, typically using explicit methods for initialising the model and more implicit or hybrid methods for updating the model.

Last, the *modelling approach* refers to the technique of learner modelling (*How is the learner modelling executed?*), which has already been outlined in Sect. 4.2.[12]

Data collected from the above data sources using the mentioned collection techniques can be further categorised into the common categories used in a more legal context, which were introduced in Section 4.2.1. Abrams (2014), founder of the Information Accountability Foundation and a member of the OECD Working Party on Information Security and Privacy, provides a structured taxonomy of these categories, dividing data into *provided* data, *observed* data, *derived* data and *inferred* data[13]. Provided data is given directly by the users (here: the learners), such as when they fill in a registration form, which they are fully aware of (cf. *self-report*). Observed data is collected by observation via facial recognition or sensor technology, for instance (cf. *observation source*, *sensors & wearables*). Derived and inferred data create new data based on existing user data and, therefore, cannot be directly associated with a specific data source in our classification. While derived data is created using rather simple algorithms and mechanical methods to compute new data from existing data, inferred data uses more sophisticated, often probabilistic, analysis to produce new data (cf. *predictive modelling*) (Abrams 2014). Since we observed a wide variety of data, data sources and data collection methods that extended beyond these categories, we decided to take a more inductive approach rather than strictly assigning data to these categories.

As one can see from the structure of this section, the findings for RQ2 relate to two fundamental steps of modelling learners in adaptive systems (Abdennour et al. 2023), i.e. initialising and updating the learner model with data. In the first step of initialisation, the model is filled with initial learner data, typically elicited by questionnaires, forms or similar instruments. If the model is to be kept up to date, new data is usually elicited for the updating step, typically user behaviour or assessments within the adaptive system and less frequently, contextual data or physiological data is included as well. The subsequent sections present the results from the literature review for both of these steps.

4.5.2 Initialising learner models

When it comes to characterising data in learner models, a common practice is to distinguish between *static* and *dynamic* data (Germanakos and Belk 2016). In many cases, the process of initialisation introduces relatively static data such as *demographic & biographic data*, *identification data* (i.e. e-mail) or personal preferences such as learning style. The step of initialising a learner model may traditionally be taken by assigning a learner to a certain known group (cf. stereotyping (Rich 1979)) or via other methods such as assuming details about the learners based on previous learners' data or by having the learners themselves provide initial information. Eliciting initial

[12] While RQ1 included the individual representation (modelling approach) of single components of the learner model (e.g. representing knowledge as an ontology), RQ2 will consider the modelling technique for the learner model as a whole on a more abstract, component-independent level. Examples are approaches like fuzzy logic, probabilistic modelling or others that we have classified in Sect. 4.2.

[13] Abrams taxonomy consists of further subcategories which are omitted for clarity and scope at this point.

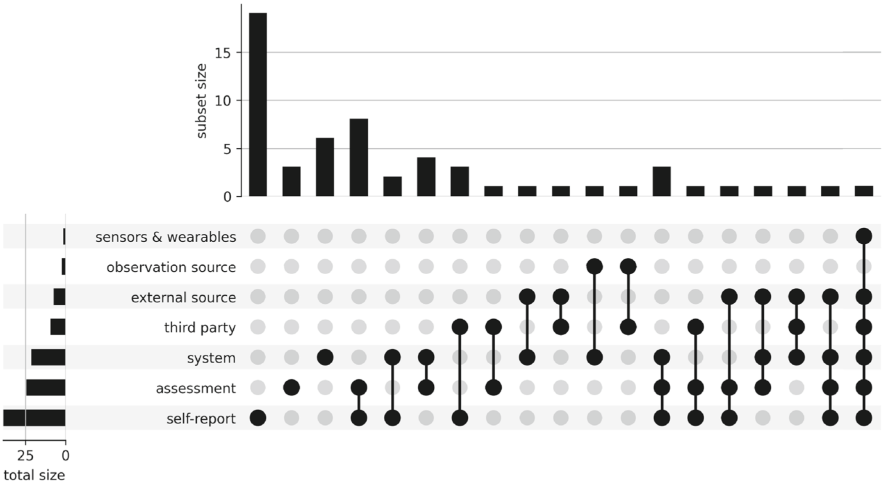

Fig. 16 Frequency of use of (combinations of) data sources for initialising the learner model. Overall source frequency—left histogram; frequency of combination of sources—top histogram.

information can help address the cold-start problem (Bobadilla et al. 2012; Park and Chu 2009) for new learners.

What becomes evident is that the research papers provide significantly fewer details about initialising than updating the learner models. Of the ($n = 127$) papers relevant to RQ2, only 44% ($n = 56$) address the initialisation of the model, and if so, only vaguely in many cases, while the step of updating is addressed in 91% ($n = 115$) of the papers.

Looking at the data sources used to initialise learner models, we see a distribution as visualised by Fig. 16. As we can see, the major sources that are relevant for initialising learner models are *self-report*, *assessment* and the *system* itself. Overall, only a few possible data sources (mostly just two) are combined in the initialisation step.

As we can see, the most common data source for initialising learner models is *self-report* as a single data source and, thus, information provided by the learners themselves explicitly. These can range from initial personal information provided in a registration form, such as details about their academic or demographic past (Ayari et al. 2023; Wang et al. 2021; Janati et al. 2017) to *cognitive characteristics*, *metacognitive characteristics* or *affectional & motivational characteristics*, which can help in assessing the individual student's state or features via questionnaires or surveys (Sun et al. 2023; Pustovalova et al. 2022; Ramírez Luelmo et al. 2020). The second most popular source is *assessment*, such as prior knowledge tests, which usually help identify the initial cognitive state of the students (cf. Sun et al. (2023)). In particular, the combination of *self-report and assessment* is a popular approach to initialise a learner model. A further data source for initialisation is the *system* itself, i.e. setting initial parameters learned by historical data, which is also found to be used in the approaches we have examined (Barria-Pineda et al. 2017), helping to make assumptions about the learner before they interact with the system. While some papers also rely on different data sources such as third parties (*third party*) like instructors (Ganesan et al. 2023;

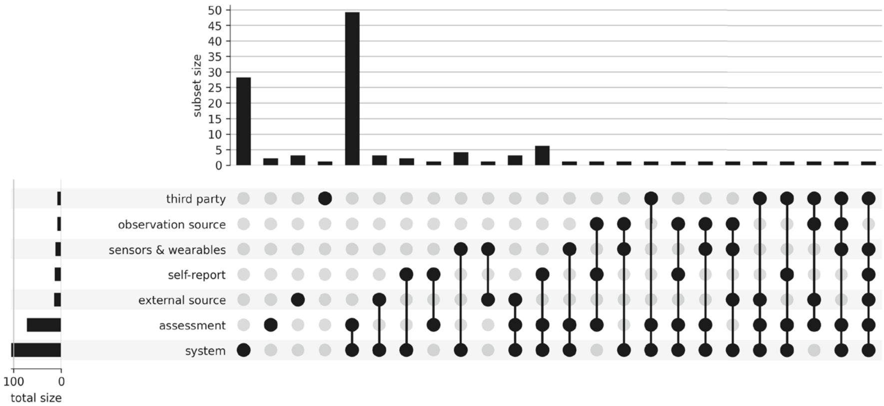

Fig. 17 Frequency of use of (combinations of) data sources for updating the learner model. Overall source frequency—left histogram; Frequency of combination of sources—top histogram.

Chrysafiadi et al. 2020; Schulz et al. 2020), peers or parents, *external sources* such as literature research data (Sun et al. 2022b), social network data (Vo et al. 2021) or observations (*observation*), e.g. teaching observation (Gooch et al. 2016)—and these all mostly in combination with one of those mentioned above three—the majority relies on either *self-report*, *assessment*, the *system*, or a combination of these.

This picture is completed by the collection methods, which correspondingly have a clear focus on *self-reporting* and *assessment-based* collection, while *automated collection*—i.e. by the system itself, sensors or external sources accessed via APIs, etc.,—comes as a second. Therefore, the large majority of these approaches use explicit methods (explicit: $n = 35$, hybrid: $n = 11$, implicit: $n = 10$), which might be related to the phenomenon that learning systems tend to lack initial user data, and the explicit method may be the most straightforward approach. Implicit methods may not be as precise or hampered by the issue that external sources like student databases may not always exist or disallow the retrieval of such information without the users noticing.

4.5.3 Updating learner models

While the initialisation step can help populate a learner model and use it to personalise a system with this initial information, such as demographic details, prior knowledge or learner preferences, it does not consider the nature of the changing learner. Most approaches acknowledge the fact that the process of learning is not static and reflect this fact in creating a dynamic model (Tetzlaff et al. 2020) and therefore address the crucial aspect of updating the learner model in their research.

A fairly unbalanced distribution can be seen if we look closely at the data sources commonly used for updating the learner model (cf. Fig. 17). The most common data source for the step of updating a learner model that could be identified in the examined research papers is the combination of the *system* itself *and assessment*, i.e. tests or quizzes that are mostly integrated within a system (mostly adaptive learning systems (Al-Chalabi et al. 2021), personalised or recommendation systems (Lin et al. 2023) or intelligent tutoring systems (Beyyoudh et al. 2018), but also other systems like ubiquitous learning environments (Ferreira et al. 2017), for instance). Secondly, a majority of the examined papers were found to update their learner models based on data from the *system* itself, such as action logs (Qiao and Hu 2018) and access or *interaction data* with educational resources or materials (Zhao 2021), deriving user patterns.

Some approaches take into account multimodal data which is reflected in the diverse combinations of collection techniques which include multiple sources such as sensors, cameras, and robotic technologies—reflected by the groups including *sensors & wearables*, *observation source* and others (Santos et al. 2014 and Neji et al. 2011 and Shen et al. 2009 cited by Bellarhmouch et al. 2022; Martinez-Maldonado et al. 2020 cited by Khosravi et al. 2022). For instance, Yang et al. (2022) include a multitude of data sources into the process of learner modelling, deriving preference data from a variety of inputs such as *(social) interaction data*, *(learning) output data*, and *contextual data* coming from the system as well as sensors or Martins et al. (2018), who draw on multiple distributed sources to derive learner characteristics such as robots as well as IoT sensor networks and further devices. Interestingly, we could not identify an increasing usage of such approaches through the selected period in the papers examined.

In contrast to the initialisation step, the manner of collection for updating learner models was predominantly hybrid or implicit (hybrid: $n = 60$, implicit: $n = 49$, explicit: $n = 6$). The prevalence of implicit techniques is not surprising as many of the more state-of-the-art manners of collecting and analysing data are unobtrusive and do not need attention from the user, such as sensor-collected data or automatic system logs. Nevertheless, explicit methods are represented in a significant number of hybrid approaches, where *automated collection* is complemented by *self-reporting*—i.e. in the context of user feedback in negotiable or editable open learner models (OLM) (Kay et al. 2022; Ramírez Luelmo et al. 2020; Bull et al. 2016; Ginon 2016 and Van Labeke 2007 and Bull and Pain 1995 cited by Bull 2020)—or *third party reporting*—i.e. via cognitive assessments of peers or instructors (Schulz et al. 2020; Al-Jadaa et al. 2017).

4.5.4 Modelling approaches

While the previous section examined the data sources and how the data is retrieved, this section takes a closer look at the second part of the research question, i.e. *What is the overall modelling approach that was utilised?* In total, 103 of the 127 relevant papers for RQ2 answer this question and utilise one of the modelling approaches presented in Sect. 4.4.5. Among those that specify an approach, the majority of 56% could be assigned to *predictive modelling*, followed by 32% of papers that include *uncertainty modelling*. *Classification* is used by 23%, *ontology* by 21% and *clustering* by 17%. Less frequent is the use of *overlay modelling* 12%, *cognitive theories* 11% and least, *perturbation modelling* and *constraint-based modelling* with 4% of the publications.

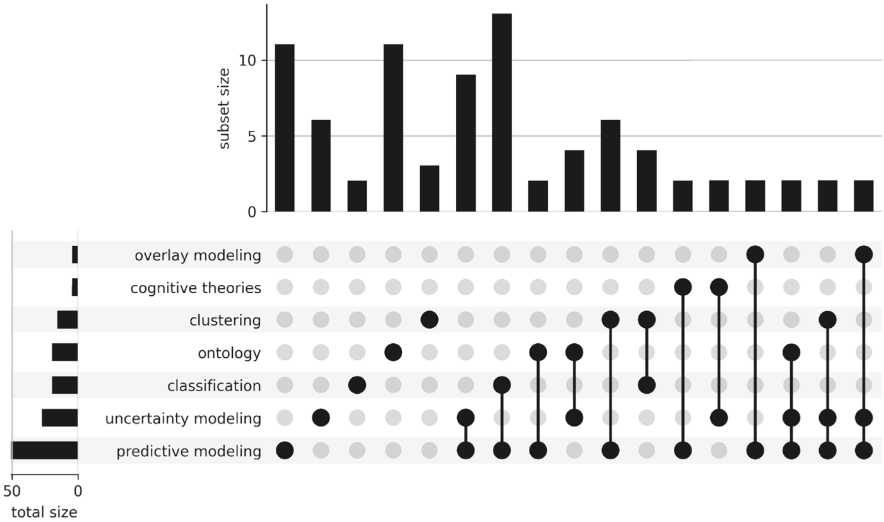

Fig. 18 Frequency of use of (combinations of) modelling approaches that appear more than once. Overall approach frequency—left histogram; Frequency of combination of approaches—top histogram.

Figure 18[14] demonstrates frequent combinations of approaches beyond these single-approach numbers[15] In particular, hybrid models seem to be popular (cf., for example, Sani and Aris (2014b)), and we could observe a high significance of combined approaches, often including *predictive modelling*, such as *classification and predictive modelling* (Bernard et al. 2022; Lwande et al. 2021; Abyaa et al. 2018; Ortigosa et al. 2014a), *uncertainty modelling and predictive modelling* (Almeida et al. 2021; Faucon et al. 2020; Martins et al. 2018; Wagner et al. 2014) or *clustering and predictive modelling* (Dijkstra et al. 2023; Sun et al. 2018; Chatti et al. 2014). Other combinations are less frequently used and it becomes evident that more traditional approaches, such as *overlay modelling* (Greer and McCalla 1994) and *cognitive theories*, tend to appear in combination with other approaches such as *uncertainty modelling* (Chrysafiadi and Virvou 2015; Liu et al. 2022 and Chrysafiadi and Virvou 2012 mentioned by Liang et al. 2022) or *predictive modelling*. Often, hybrid models can be found with such approaches. For instance, Chrysafiadi and Virvou (2014) combine *overlay modelling* with other approaches like *perturbation modelling*, *cognitive theories* (OCC model) and *uncertainty modelling* (fuzzy logic).

When we examine the usage of the modelling approaches throughout the last ten years as in Fig. 19, we see a rather consistent development for many approaches such as *classification*, *clustering*, *ontology* or *cognitive theories*. The significance of *constraint-based modelling*, *overlay modelling*, *perturbation modelling*, and—to a

[14] Combinations that appear only once have been excluded from the plot for the sake of readability. A complete version of the plot can be found in "Appendix Fig. 22".

[15] Combinations that appear only once have been excluded from the diagram for the sake of readability. Therefore, the frequencies on the left-hand side histogram differ slightly from those presented above, and perturbation and constraint-based modelling are not included due to their appearance in unique combinations only.

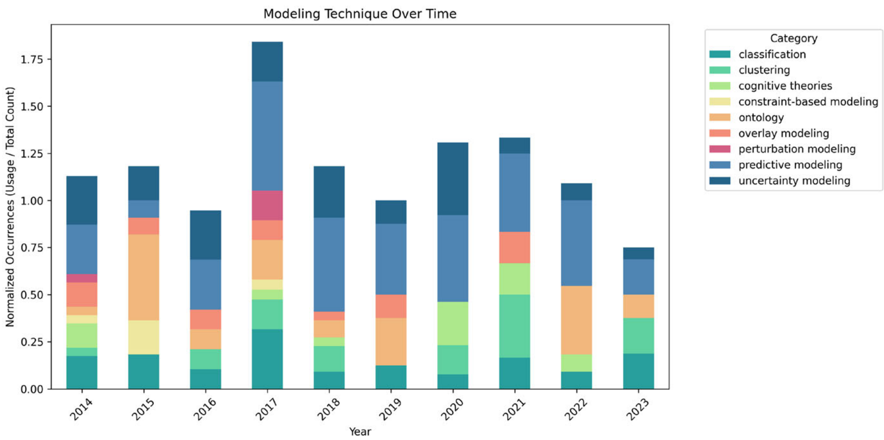

Fig. 19 Modelling approaches over time (occurrences have been normalised as the count of papers per year varies.

certain extent—*uncertainty modelling* has decreased, while *predictive modelling* has been highly common. Considering the work of Chrysafiadi and Virvou (2013), who examined the modelling approaches that were utilised between 2002 and 2012, we can confirm some trends they have observed, which seem to continue in this last decade we have examined. Thus, the significance of the overlay and perturbation model has decreased throughout the years, while cognitive theories, for instance, appear to be a mostly constant modelling approach. Techniques they described to be more recent, such as Bayesian modelling—subsumed under *uncertainty modelling* in our work—and *ontology* both have been highly significant throughout recent years.

4.5.5 Discussion of results

To conclude this subsection, we collect some of the most striking aspects that emerged from the literature review.

First, only 38 out of 127 papers (30%) fully address all aspects of RQ2, and if parts of RQ2 appear in papers, the information on the data sources or gathering is often sparse, and the results of such a major study thus have to remain on a rather abstract level as has become evident in Fig. 17. While the initialisation mostly draws on traditional data sources such as *self-report* or *assessment*, a variety of data sources was found to be used to update the learner models. What has become evident in the research is that data preprocessing as a major step between gathering the data and modelling is hardly addressed in detail, nor is the data's precise nature, which, again, defies precise classification. For the modelling approaches, there is a rise in *predictive modelling*; other more traditional methods, such as *overlay modelling* or *perturbation modelling*, are not as commonly used as they had been around the early 2000s (Chrysafiadi and Virvou 2013; Stash et al. 2006 and Alfonseca et al. 2006 and Papanikolaou et al. 2003 cited by Alshammari et al. 2014). However, they are still utilised and are predominantly combined with further approaches. Last, many of the examined approaches, such as *ontology* or *cognitive theories*, have been used constantly throughout the last decade.

As introduced in this section, updating the learner model is the major step after initialising it. We have seen how the systems themselves, as well as assessments, are common approaches to populating such a learner model. For this purpose, many research papers utilise data related to knowledge and the cognitive aspect, as we have described in Sect. 4.4. However, we have not considered the changing nature of this data by phenomena such as forgetting, which will be addressed in the following section.

4.6 RQ3: Ephemeral knowledge and forgetting

Accordingly, a subsequent question is: *How do learner models evolve during their lifetime concerning the learner's ephemeral knowledge?* (RQ3). Thus, this subsection addresses knowledge development over time and the important topic of forgetting. This process of forgetting has been explored in manifold ways, one of the most prominent ones being David Ausubel's idea of subsumption, which attributes forgetting not to the loss of knowledge but to its meaningful integration into the learner's existing cognitive structures, whereby overarching concepts and relationships are retained while specific details fade (Ausubel et al. 1968).

Of the 315 scientific publications analysed that deal more intensively with learner models, only 25 (<8%) address the topic of knowledge development over time and, thus, the important topic of ephemeral knowledge and forgetting. This allows conclusions to be drawn about the current state of research. Of the few scientific publications that do address the topic, most are very superficial and only touch on the subject. Three groups of publications can be roughly categorised:

- Those that mention the importance of the topic but do not provide any details about its specific implementation or cite any further references.
- Those that refer to already known and proven approaches.
- Those that do describe details of their implementations, including specific parameter settings or assumptions.

For example, the first group points out that time (Martynov et al. 2023) is an important factor in learning. Therefore, the temporal impact of previously learnt material must be modelled (Liang et al. 2022), as the elapsed time after the learner has learnt the knowledge increases the probability of forgetting the knowledge (Chrysafiadi and Virvou 2015; Oeda and Shimizu 2021; Wang et al. 2021).

On the other hand, some publications draw on already proven approaches and build on and reference them. For example, Oeda and Hasegawa (2023) Item Response Theory (IRT) is expanded to include the well-known Ebbinghaus Forgetting Curve (Ebbinghaus 1885) to take the topic of fleeting knowledge into account. Similarly, Abyaa et al. (2017) mention spaced repetition algorithms, which can address the issue of forgetting through repeated quizzes. However, they also observe that forgetting is hardly addressed in related approaches. Pelánek (2016) made a case study about extending the Elo rating system with learning. Another common approach is to model the knowledge and temporal modelling using fuzzy rules so that the system can infer as soon as the learner's knowledge level decreases (Chrysafiadi and Virvou 2014, 2015). Last, Bull (2020) describe that the aspect of forgetting can be addressed by interactive models in the context of open learner models that allow for user contribution; however, they do not provide further detailed information on the specifics of this in their article.

In contrast, some works describe how forgetting is addressed in more detail. They describe parameter settings and/or assumptions made in modelling the fading of knowledge. For instance, Abdi et al. (2021) consider time lags by adjusting the model's confidence parameter accordingly. Similarly, Qinghong et al. (2015) use a forgetting factor that gradually reduces the weight of older interests, ensuring the user model stays updated with the current ones. Last, Kurup et al. (2016), who use Bayesian knowledge tracing, assume that the forgotten skill is set to 0 after a student's state of knowledge evolves from learned to unlearned.

In summary, ephemeral knowledge has little representation in the analysed literature. The temporal aspect should also be considered for a digital twin of the learner that is as accurate and up to date as possible. This applies particularly to dynamic learner characteristics (cf. RQ1) that change over time. Alternatively, the learner would have to prove his current knowledge cyclically via assessments, for example, so that the learner model remains up to date. However, this is almost impossible due to the many different subject areas and would also be to the detriment of the learner, as they would not be able to acquire any new content during this time.

In the future, this means that existing research would have to be examined more closely in a further step to evaluate the already known approaches and weigh up which realisation most closely resembles the nature of ephemeral knowledge. More concrete implementations must be discussed to incorporate the important topic of forgetting into the learner model. This was hardly described in sufficient detail in our analysed publications.

4.7 Summary and overview

From the perspective of RQ1, it can be summarised that there is no detailed description of an ideal learner model. Most publications list a few possible attributes without describing how these are modelled in detail and where the data for them comes from in detail. Unfortunately, this does not allow for a direct implementation based on the publications because none of them provide a ready-made data or learner model via a research data management system. Other authors also see the problem of learner modelling as a challenge, as it is based on the combination of different disciplines such as educational science, psychology and information technology (Abyaa et al. 2019 cited by Zanellati et al. 2024). Therefore, for example, Abyaa et al. (2019) propose to develop an ideal learner model through the following three steps: (1) identifying and selecting the attributes of the learner that influence their learning; (2) considering the psychological states of the learner during their learning process; (3) selecting the most appropriate technologies that enable accurate modelling of each selected attribute. These attributes[16] just mentioned have been studied individually in this publication. Based on the detailed analysis, we have designed a taxonomy for categorising indi-

[16] Attributes are referred to as components in this publication.

vidual possible components to make it easier for designers and developers to start constructing a learner model.

Regarding RQ2, we found that few papers explicitly address all parts of the modelling process, including details such as data sources, the nature of the data, data collection, preprocessing and modelling approach. While detailing every step may not be necessary for all papers, specifying such details would enhance transparency in the field. Since pipelines are likely to become more complex with the popularity of approaches rooted in data analytics, we encourage future authors to address these aspects more thoroughly.

To adapt the learner model as faithfully as possible to the real learner, knowledge with all its facets must be considered. In addition to acquiring knowledge, this also includes the topic of forgetting: ephemeral knowledge. It has already been scientifically proven that knowledge does not remain constant over time but changes. From this point of view, the temporal aspect must also be considered. As we saw in Sect. 4.6 on RQ3, this topic is (still) underrepresented in science. This can be derived from the small number of mentions of the concept of forgetting and the even smaller number of papers that deal with the topic (even if only superficially). The topic will require more attention to make major scientific progress in the future.

Last, to provide an overview of the publications that were analysed in this work, we collected the metadata of the papers as described in Sect. 3.5, as well as their relevance for each respective research question, which is visualised by the tables in the Appendix. To highlight recent papers, we have sorted the papers according to recency. As we acknowledge that surveys usually have a larger scope and tend to have more references than other works, we labelled surveys explicitly as such in the overview.

As demonstrated in Sect. 4.1, the topic of learner models has become increasingly important in recent years, also regarding the shift of teaching to the digital space due to the pandemic. This can also be seen in the number of published articles. This makes it all the more important to present and compare an overview of the various options for creating a learner model. For this purpose, we have created a digital repository, which we have made available via a research information system (Böck and Ochs 2025). It contains all the information needed to reproduce this systematic literature research in a traceable way. In addition, further (interactive) forms of presentation of the processed data are available in the repository, such as a rule-based classifier, which can be used to classify a new attribute into the taxonomy of this work, which may help find related approaches. Therefore, this repository is intended as an introductory aid for designers and developers of learner models to help them construct them according to their needs and requirements.

5 Discussion

After examining the results extracted for each research question in the SLR, this section extends these findings by first discussing the purposes and applications of learner models and second, by examining the challenges and opportunities that emerge from analysing this large body of current literature. Both sections take a larger perspective on the topic and aim to derive research opportunities for future work.

5.1 Purpose of use and applications

As we have seen throughout the SLR, learner models appear in diverse forms and play a crucial role in various subfields of individual learning and adaptive systems in education. Learner models have been an important part of various related—and intersecting—research communities such as Intelligent Tutoring Systems (ITS) or Artificial Intelligence in Education (AIED), among others. The relationship between these two subfields illustrates this intersection. While ITS largely focus on modelling the students' knowledge or cognitive state (Greer and McCalla 1994), AIED aims at integrating various techniques of AI into educational processes or systems, which results in diverse possibilities of student support (Khosravi et al. 2022). Such support can enhance ITS with functionalities such as AI-generated feedback, personalised learning recommendations or the automatic prediction of possible learning outcomes. Innovative developments in AIED may reinforce innovations in ITS, and innovation in the learner modelling field contributes to both of these research areas, as well as others. While their appearance in different communities of adaptive learning systems shows their multi-faceted relevance in research, learner models are highly relevant for diverse applications across different levels of contexts in practice.

Levels of Application—Micro, Meso, Macro. Learner models are relevant across diverse educational contexts and levels:

- **Micro level:** At the micro level, learner models facilitate individualisation by providing tailored insights into each learner's progress and needs.
- **Meso level:** At the meso level, these models inform instructional design, allowing educators to make course-based adjustments that enhance the learning experience.
- **Macro level:** At the macro level, learner models can influence institutional strategies and curriculum design.

Due to our focus on individual student modelling rather than cohorts of students, as they are usually considered in learning analytics (cf. Sect. 2), our research is mostly related to the micro level. Future research should delve into specific modelling practices relevant to each application, particularly highlighting areas that have not been thoroughly examined, such as the following.

Open Learner Models and Learning Analytics Dashboards. An extension to the learner models we have considered and a significant advancement in the field is the open learner model. Open learner models are visual representations of machine-readable representations of learner models and their components (Bull and Kay 2010). They aim to support the learner in learning in various ways (Bull and Kay 2007) by giving them a tool for self-reflection in different ways. In addition to only visualising the personal data, some models also offer the option of interacting with it and, for example, negotiating with the model if—in the learner's opinion—there are inadequacies in the data (Bull et al. 2016; Bull and Kay 2010). Thereby, OLM give learners ownership of their individual data and progress. A common tool closely related to these OLM are Learning Analytics Dashboards (LAD) (Bodily et al. 2018). LAD provide a clear and user-friendly presentation of learning data through transparent data visualisations, which would often be difficult to access or confusing without such representations. LAD, however, are based on a static representation of behavioural metrics from interaction data and not—as with learner models—on modelling knowledge and other individual characteristics of the learner (Bodily et al. 2018; Kay and Bull 2015).

Tool for Instructors. However, OLM and LAD cannot only serve learners but also their instructors. These tools empower instructors to tailor teaching strategies or implement targeted learning interventions by providing insights into students' learning states. In this way, instructors gain access to relevant learner data, typically presented in pseudonymised or aggregated forms, to uphold data protection standards, specifically on the macro level. One typical example of how such insights may be used is dropout prevention, where a major challenge is identifying trends such as increasing dropout rates without compromising student data privacy in the intervention. Observing the decreasing performance of students in a course, instructors can develop strategies to address these issues, such as implementing targeted resources and ensuring that individual student data remains protected and anonymised. The applications of OLM and LAD demonstrate that although the disciplines of learner modelling and learning analytics are delineated in theory, they tend to overlap in practical applications.

Applications Beyond. As we have observed in the analysis of domains addressed in the publications, a significant proportion of the publications are not explicitly associated with a specific domain. Yet, in less frequent cases, we also saw that learner models show potential for use beyond academic or educational contexts such as workforce training (Hansen et al. 2017), adult learning (Corredor et al. 2023; Abyaa et al. 2017, 2016) or lifelong learning (Ramírez Luelmo et al. 2020; Nurjanah 2018). These applications may introduce further challenges that were only marginally considered in the scope of this work with a focus on higher education. Future research should not hesitate to explore these less examined application areas, acknowledging the manifold applications of learning and the importance of individualisation in education.

5.2 Challenges and opportunities

To complement and extend our observations in the SLR, this subsection discusses some crucial challenges and opportunities related to the examined topic, highlighting further research opportunities for future work.

Privacy and Data Sovereignty. First, data privacy and data sovereignty are crucial aspects in learner modelling, as learner models can hold significant amounts of often sensitive data. Privacy regulations such as the GDPR (EU) or CCPA (US) have significantly impacted the field by establishing provisions that protect individuals' data and enforce user control (e.g. through user consent) in data collection and storage, for instance. However, the notion of privacy has been a part of this strand of research for several years before these acts were established. A prominent example is provided by Anwar (2021), who offers an overview of mechanisms that support privacy and trust in online learning based on prior work with Jim Greer. Their work introduces different privacy mechanisms, such as identity management and privacy preference settings. We observed some of these mechanisms in the research we examined. Many researchers rely on identity management, facilitating students' pseudonymity within their systems. In these cases, privacy measures are often mentioned only marginally, presumably because they are already established in the system architecture that the user model is part of. Similarly, user consent is seldom explicitly mentioned, possibly because it typically occurs during onboarding and is not perceived as central to the research presented. More explicitly, we observed research where users can set privacy preferences (Tang et al. 2020), to choose between different privacy modes for activity recording within a collaborative learning environment. More specialised approaches like federated learning are also present in the examined literature, where the users' data resides locally on the client (Liu et al. 2023). Privacy is often a major concern in the context of multimodal approaches within ubiquitous systems—possibly because modern multimodal approaches, in particular, may rely on sensors and similar technologies to process *contextual data* or affective data. Here, we found that privacy concerns may lead the decision in favour of sensor-free approaches (Rahman et al. 2023) to circumvent these issues. Although implicit implementations of privacy regulations, such as using identity management or eliciting only necessary learner details, could be observed, we believe data privacy and sovereignty may need to be addressed more explicitly. Conscious, privacy-related decisions could be conveyed more clearly so that privacy mechanisms become more prominent when designing user models or adaptive systems.

Ethical Considerations and Interpretability. The second aspect—ethics—is closely tied to the topic of privacy and data sovereignty and was thus already touched upon to a certain degree through the ideas of user control and user consent, for instance. The development of adaptive systems is increasingly guided by further concepts such as the FATE principles—fairness, accountability, transparency and ethics (Woolf 2022). These principles address crucial ethical questions such as how to ensure fairness in algorithms, who is to be held accountable for model outcomes, how to reduce biases and how to make complex models more transparent to learners so they can understand the decisions that these models make.

The growing relevance of predictive models, as observed in this article, highlights the increasing importance of model transparency and interpretability. We adopt Miller (2019) definition of interpretability, who equates it with explainability and defines it as the understandability of a model's or an agent's decisions by humans. This understandability or interpretability is crucial for fostering learners' trust in the model or the adaptive system (Mannekote et al. 2024; Zapata-Rivera and Arslan 2024). Recent research demonstrates a clear interest in this topic of interpretability. For example, Khosravi et al. (2022) introduce a framework (XAI-ED) aimed at developing trustworthy and explainable AI tools for educational purposes. Oeda and Hasegawa (2023) focus on Item Response Theory (IRT) with interpretability as their primary motivation. The growing interest in the topic is especially prominent in the context of OLM research. OLM are often developed to enhance model transparency and improve the models' understandability. In the case of editable models, OLM also aim to increase user control over the model's information. Given their pioneer role in this area, OLM may serve as a conceptual guide for creating interpretable models in related fields (Conati et al. 2018). Looking ahead, we expect the topic of FATE to remain prominent in research. In particular, the T (transparency) aspect will continue to be a relevant topic regarding the increasing complexity of learner models and the prevalence of artificial intelligence in the field.

Interoperability & Reuse. Another important aspect of learner models is their ability to facilitate personalised recommendations. When launching a recommendation platform that incorporates learner models, the cold-start problem arises, as there may be insufficient authentic user data to populate the learner model. This lack of data hinders the generation of suitable initial recommendations and the effective testing of the system. Generating synthetic, authentic learner model data could enhance the shareability and reusability of these models, which in turn would promote more reproducible open science. However, our review has revealed a limited number of synthetic, authentic learner model data. Those we found were utilised for experiments to validate the approaches and were generated either using Item Response Theory (IRT) (Oeda and Hasegawa 2023; Oeda and Shimizu 2021) or by simulating interactions among diverse user profiles (Martins et al. 2018). Beyond synthetic data, common learner datasets are referenced in multiple instances—Minn, for instance, names some of the most common ones in the context of knowledge assessment (Minn 2022). Improving interoperability can also occur on the data acquisition level where standards such as xAPI aim to facilitate educational data sharing, particularly in the MOOC context, e.g. (Ramírez Luelmo et al. 2020). At the model level, shared ontologies (Rahayu et al. 2022) could be identified. However, relying on standardised approaches or the provision of interoperable solutions remains uncommon in the literature we reviewed, although the issue had already prevailed a decade ago (Desmarais and Baker 2011). Mostly, it is in the context of open learner modelling where we could identify such efforts (Conejo et al. 2012 and Kay 2008 and Dolog and Schäfer 2005 and Zapata-Rivera and Greer 2001 cited by Bull 2020). Also, none of the considerable number of papers reviewed contained any operational models that were implemented concretely in the form of software programs or programming libraries that can be executed, shared and reused. Given the potential of current learner models, we believe there is significant opportunity for growth in this area moving forward.

6 Current research

As research and development in the field of learner models has continued to evolve after our specific analysis period, numerous publications have emerged in 2024. Applying our methodology from Sect. 3 to these publications, this results in further hits—IEEE Xplore (701), ACM Digital Library (81), Elsevier ScienceDirect (409), pedocs (0) as of 01/12/2024. Since this new data would require further processing time, we have decided to select a set of works that have been cited frequently in the short period that fit and extend our research based on their titles and abstracts. We use them to indicate further development throughout the year.

At first glance, some trends persist in these new works. As indicated by examining the data gathering and modelling process for RQ2, data-intensive approaches such as artificial intelligence or data mining have become highly prevalent in the field. This applies to the modelling process in general and the application level, such as course recommendation systems (Narimani and Barberà 2024). Similarly, the increased use of large language models (LLMs) has created new possibilities, including creating and updating learner models. Mannekote et al. highlight this potential but also emphasise the significant challenges of interpretability and explainability (Mannekote et al. 2024), which we already addressed in the discussion section. Nevertheless, they see the combination of LLMs in interaction with learner models as a promising approach and suggest starting with the initialisation of a learner model without LLMs and passing its inferences to the LLM as an additional component of prompts. This strategy is already being tested with GenAI models such as LLMs (Zhang et al. 2024a mentioned by Mannekote et al. 2024). Other possible applications for integrating LLMs in the field of ITS include generating feedback for learners (Stamper et al. 2024) and improving the prediction of learning performance (Zhang et al. 2024a). In addition, numerous surveys deal with LLMs in dialogue systems (Yi et al. 2024; Wang et al. 2023).

Another emerging focus is incorporating multiple learner characteristics (and/or data) into the learner models. For instance, Narimani and Barberà (2024) emphasise the importance of avoiding oversimplification of learners by incorporating multiple learner data (e.g. learner actions) into the learner model rather than relying on single components. Wang et al. (2025) extend this idea of multiple learner characteristics to the adaptation level. They observe that existing adaptive e-learning systems tend to use single learner characteristics—mostly learning style and knowledge level—and promote the inclusion of multiple learner characteristics to improve adaptability.

While examining several learner characteristics within a model is a central topic in recent research, the detailed examination of individual components (e.g. knowledge) remains relevant. For instance, Zanellati et al. (2024) conducted a survey on hybrid models for knowledge tracing, building on the detailed classification of knowledge by von Rueden et al. (von Rueden et al. 2023 cited by Zanellati et al. 2024), which identifies three main sources of knowledge, i.e. scientific knowledge, world knowledge and expert knowledge. Purely data-driven knowledge modelling approaches, such as hidden Markov models, factor analysis models, learning factor analysis and deep learning-based models, can be extended by integrating alternative sources of knowledge into the data (e.g. learning context) (Zanellati et al. 2024). This in-depth research should be carried out for all analysed components of learner models individually and in combination to uncover causal relationships between the components. However, not only the individual components must be considered when designing and developing learner models, but also the purpose and the target audience.

A valuable path in recent research is the more detailed exploration of stakeholder involvement in learner modelling and the development of adaptive systems. Alajlani et al. (2024) reviewed stakeholder involvement across several stages of adaptive system development, including how learners, educators, developers and academic researchers are involved in the design, development and testing phase of the development of an adaptive system. They identified gaps in the involvement of learners during the development phase and of educators during the design phase, for example, and emphasised the need for further evaluation of the effectiveness of including the respective stakeholders to verify if the inclusion is advantageous.

Ongoing work also studies the application of learner models in specific contexts such as pedagogical agents. Zhang et al. (2024b) highlight that learner characteristics can influence the effectiveness of pedagogical agents and should, therefore, be considered when designing a pedagogical agent. For example, Heidig and Clarebout's model (Heidig and Clarebout 2011 mentioned by Zhang et al. 2024b) includes learner characteristics across various dimensions, including cognitive factors such as prior knowledge, academic competencies such as GPA scores, emotional factors such as boredom, pride, joy and shame, motivational factors such as interest, achievement motivation, self-efficacy and metacognitive factors such as self-regulation.

7 Conclusion and outlook

Throughout this work, we have explored the varied terminology, forms and structures associated with learner models. These variations reflect the richness of learner modelling approaches but hinder the accessibility and understanding of this particular research field. This creates the need for analysis as well as categorisation of existing approaches. For this purpose, we systematically developed an extensive taxonomy built on a large collection of relevant literature to structure the data and simplify access for learner model designers and developers so they can create learner models more easily. This includes providing information on how the individual components of learner models are modelled and where the necessary data for the initialisation and updating come from. By providing this comprehensive systematic overview spanning the last decade (2014–2023), we aim to advance the state of research and facilitate the design and development of future learner models.

Last, there are inherent limitations to this survey. Although the methodological foundation of this work was constructed carefully, systematically, and guided by senior researchers as described in Sect. 3, this research is based on a selection of papers from the current literature on learner modelling, which, by nature, is not representative of the entire body of literature on the topic. Since our goal was to provide a comprehensive contribution that allows for systematic access to the field, and we have, therefore, integrated numerous approaches into this work, it is impractical to provide an in-depth analysis of each paper within the scope of this article. Nevertheless, we have attempted to highlight the diversity of the papers as comprehensively as possible in the scope of this work and to the best of our knowledge. We will further provide the data on which this work is based (Böck and Ochs 2025) as it is impossible to convey the details of every approach in this article.

We would like to conclude by pointing out three critical aspects that came to our attention during the analysis. We propose these aspects as foundational directions for future research.

In-Depth Analysis of Learner Models. Learner models have often been explored only superficially, with limited attention to their specific characteristics and interrelationships—possibly because the focus is frequently on other aspects such as constructing an adaptive learning system, where learner models serve as a means to an end. Therefore, learner models require more in-depth analysis. In addition to the possible learner characteristics listed in the literature, the analysed publications provide little information about the influence and significance of the individual attributes. For research, as well as for the development of learner models, it is essential to demonstrate and substantiate causal relationships between individual learner characteristics (Abyaa et al. 2019) and learning outcomes, as well as their combinations and the resulting added value and synergies. This would significantly simplify the creation of new learner models. However, the analysis would need to be validated in several large-scale studies.

Requirement Specification. We have not encountered any publication during the analysis that deals with a detailed list and justification of the technical and non-technical requirements. Clearly defined requirements for what a learner model has to achieve are based on the fact that there is no generally valid definition for learner models, and the scope of what belongs to it and what an accessory is not generally defined. Following our detailed definition (cf. Sect. 2), the next logical step would be a requirements analysis. The issue of stakeholders also needs to be considered in more detail as to whether it relates exclusively to learners and teachers or whether there are other authorised stakeholders, such as parents, administrative bodies or peers.

Expansion of Research Areas. Many aspects central to the research on learner models remain underrepresented in the literature. This includes the various research areas outside of computer science and specialised didactics. Future investigations should continue to address challenges such as data privacy and sovereignty, ethics and interpretability of models, and interoperability and reuse of learner data.

By addressing these challenges, researchers can enhance the applicability and effectiveness of learner models in various educational contexts. It is imperative that these challenges and research opportunities are investigated systematically and that the findings are validated through robust research methodologies. Conducting controlled evaluations in the field of learner modelling (Pelánek 2015) and adaptive systems is a crucial step for future studies, as the results may be highly insightful and inform further development (Shute and Zapata-Rivera 2012). Thereby, the field of learner modelling and its diverse applications in adaptive learning environments will be advanced.

Appendix A

Overview of all analysed publications—systematic literature review

Overview of all analysed publications—first iteration

See Table 7.

Table 7 Overview of all analysed publications—first iteration

References	Metadata				Research Questions		
	Survey	#Ref	#Cited	Domain	RQ1	RQ2	RQ3
2023							
Sun et al. (2023)	×	32	6	Higher Education	✓	✓	×
Miranda et al. (2023)	×	30	0	Higher Education	✓	✓	×
Lin and Wang (2023)	×	13	0	Not Specified	✓	✓	×
Ayari et al. (2023)	×	40	–	Not Specified	✓	✓	×
Kakizaki and Oeda (2023)	×	8	0	Not Specified	✓	✓	×
Oeda and Hasegawa (2023)	×	14	0	Not Specified	×	×	✓
Vildjiounaite et al. (2023)	✓	35	4	Adult Learning	✓	✓	×
Pratiwi et al. (2023)	×	30	0	School	✓	✓	×
Corredor et al. (2023)	×	13	0	Adult Learning	✓	✓	×
Hu (2023)	×	20	0	Higher Education	✓	✓	×
Rahman et al. (2023)	✓	60	0	Not Specified	✓	✓	×
Martynov et al. (2023)	×	13	0	Adult Learning	✓	✓	✓
Lin et al. (2023)	×	11	0	Not Specified	✓	✓	×
Gautam et al. (2023)	×	37	0	Higher Education	✓	✓	×
Dijkstra et al. (2023)	×	40	6	School	✓	✓	×
Liu et al. (2023)	×	110	3	Not Specified	✓	✓	×
Ganesan et al. (2023)	×	10	2	Higher Education	✓	✓	×
2022							
Thohiroh and Nurjanah (2022)	×	15	0	Higher Education	✓	✓	×
Bernard et al. (2022)	×	64	10	Not Specified	✓	✓	×
Sun et al. (2022b)	×	22	2	Not Specified	✓	✓	×
Sun et al. (2022a)	×	25	1	Prof. Training	✓	×	×
Oeda and Kakizaki (2022)	×	9	0	Not Specified	✓	✓	×
Pucheta et al. (2022)	✓	13	0	Higher Education	✓	✓	×
Chrysafiadi et al. (2022)	×	46	19	Not Specified	✓	✓	×
Yang et al. (2022)	×	14	4	Not Specified	✓	✓	×
Pustovalova et al. (2022)	×	19	3	Higher Education	✓	✓	×
Khosravi et al. (2022)	✓	164	289	Not Specified	✓	✓	×
Bellarhmouch et al. (2022)	✓	39	0	Not Specified	✓	✓	×
Kay et al. (2022)	×	97	28	Not Specified	✓	✓	×
Kaliwal and Deshpande (2022)	×	14	1	Not Specified	✓	✓	×
Alatrash et al. (2022)	×	55	6	Not Specified	✓	✓	×
Ezaldeen et al. (2022)	×	32	7	Not Specified	✓	✓	×
Liang et al. (2022)	✓	98	13	Not Specified	✓	✓	✓
Minn (2022)	✓	102	72	Not Specified	×	✓	×
Rahayu et al. (2022)	✓	88	95	Not Specified	✓	×	×
2021							
Tang et al. (2021)	✓	42	1	Not Specified	✓	×	×

Table 7 continued

References	Metadata				Research Questions		
	Survey	#Ref	#Cited	Domain	RQ1	RQ2	RQ3
Tetzlaff et al. (2020)	×	84	142	Not Specified	✓	×	×
Gasmi et al. (2021)	×	18	1	Not Specified	✓	✓	×
Almeida et al. (2021)	×	11	1	Not Specified	✓	✓	×
Almeida et al. (2021)	×	21	0	Not Specified	✓	✓	×
Lei and Mendes (2021)	✓	36	6	Higher Education	✓	✓	×
Oeda and Shimizu (2021)	×	14	2	Not Specified	✓	✓	✓
Zhao (2021)	×	5	3	Not Specified	✓	✓	×
Lwande et al. (2021)	×	41	85	Higher Education	✓	✓	×
Al-Chalabi et al. (2021)	×	19	8	Higher Education	✓	✓	×
Wang et al. (2021)	×	12	2	School	✓	✓	✓
Abdi et al. (2021)	×	29	13	Not Specified	✓	✓	✓
Li and He (2021)	×	10	11	Higher Education	✓	✓	×
Miraz et al. (2021)	✓	172	98	Not Specified	✓	✓	×
2020							
Vo et al. (2021)	×	68	12	Not Specified	✓	✓	×
Chrysafiadi et al. (2020)	×	37	30	Higher Education	✓	✓	×
Jing et al. (2020a)	×	28	22	School	✓	✓	×
Jing et al. (2020b)	×	73	6	Not Specified	✓	✓	×
Luxing (2020)	×	5	1	School	✓	✓	×
Tang et al. (2020)	×	28	0	Not Specified	✓	✓	×
Cully and Demiris (2020)	×	62	22	School	✓	✓	×
Qiu et al. (2020)	×	5	2	School	✓	×	×
Schulz et al. (2020)	×	15	2	School	✓	✓	×
Rowe and Lester (2020)	×	40	50	School	✓	✓	×
Ramírez Luelmo et al. (2020)	×	43	6	Lifelong Learning	✓	✓	×
Hooshyar et al. (2020)	✓	82	141	Higher Education	✓	✓	×
Bull (2020)	✓	171	84	Not Specified	✓	✓	✓
Faucon et al. (2020)	×	37	4	School	✓	✓	×
Tacoma et al. (2020)	×	47	11	Higher Education	✓	✓	×
Taufik & Nurjanah (2019)	×	13	1	Not Specified	✓	×	×
Yuan and Yang (2019)	×	15	1	Not Specified	✓	×	×
Al-Shabandar et al. (2019)	×	52	41	Higher Education	✓	✓	×
Zhang and Cheng (2019)	×	17	6	Higher Education	×	✓	✓
Saurabh et al. (2019)	×	32	1	Not Specified	✓	✓	×
Schmidmaier et al. (2019)	×	17	4	Not Specified	✓	✓	×
2019							
Ochukut and Oboko (2019)	✓	32	15	Not Specified	✓	×	×
Ahmadaliev et al. (2019)	×	14	5	Not Specified	✓	✓	×
Ennouamani and Mahani (2019)	✓	24	4	Not Specified	✓	✓	×

Table 7 continued

References	Metadata				Research Questions		
	Survey	#Ref	#Cited	Domain	RQ1	RQ2	RQ3
David et al. (2019)	×	8	23	Not Specified	✓	✓	×
2018							
Afini Normadhi et al. (2019)	✓	120	193	Not Specified	✓	✓	×
Minn et al. (2018a)	×	23	15	Not Specified	✓	✓	×
Minn et al. (2018b)	×	12	21	Not Specified	✓	×	×
Martins et al. (2018)	×	26	7	Not Specified	✓	✓	×
Abyaa et al. (2018)	×	13	13	Not Specified	✓	✓	×
Zhao et al. (2018)	×	14	0	Not Specified	✓	✓	×
Raj and Renumol (2018)	×	27	19	Not Specified	✓	✓	×
Almeida et al. (2018)	✓	30	1	Higher Education	✓	✓	×
Alday (2018)	×	9	4	Not Specified	✓	✓	✓
Papadimitriou et al. (2018)	×	17	3	Not Specified	✓	✓	✓
Qiao and Hu (2018)	×	18	10	Higher Education	✓	✓	×
Barria-Pineda et al. (2018)	×	25	20	Higher Education	✓	×	×
Emerson et al. (2018)	×	47	21	Higher Education	✓	✓	×
Jun-min et al. (2018)	×	7	4	Not Specified	✓	×	×
Lu et al. (2018)	×	40	4	Not Specified	✓	✓	×
Yago et al. (2018)	×	60	39	Not Specified	✓	✓	×
Yang and Li (2018)	×	32	216	Higher Education	✓	×	×
Ding et al. (2018)	×	19	14	Higher Education	✓	✓	×
Beyyoudh et al. (2018)	×	23	8	Not Specified	✓	✓	×
Akharraz et al. (2018)	×	42	15	Not Specified	✓	×	×
Nurjanah (2018)	×	15	13	Not Specified	✓	×	×
Aissaoui et al. (2018)	×	21	99	Not Specified	✓	×	×
Sun et al. (2018)	×	58	95	Not Specified	✓	✓	×
Kavitha et al. (2018)	×	7	8	Not Specified	✓	✓	✓
2017							
Fei Zhou et al. (2017)	×	11	7	Not Specified	✓	✓	×
Janati et al. (2017)	×	34	51	Not Specified	×	✓	×
Al-Abri et al. (2017)	×	28	8	Not Specified	×	✓	×
Abyaa et al. (2017)	×	25	9	Adult Learning	✓	✓	✓
Ferreira et al. (2017)	×	36	9	Higher Education	✓	✓	×
Chaturvedi and Ezeife (2017)	×	20	8	Higher Education	✓	✓	×
Zhu et al. (2017)	×	6	6	Not Specified	✓	✓	×
Ferreira et al. (2017)	×	17	9	Higher Education	✓	✓	×
Mejia et al. (2017)	×	8	16	Higher Education	✓	✓	×
He et al. (2017)	×	10	15	Higher Education	✓	×	×
Huang et al. (2017)	×	18	8	Not Specified	✓	✓	×
Yi et al. (2017a)	×	11	3	Higher Education	×	✓	×

Table 7 continued

References	Metadata				Research Questions		
	Survey	#Ref	#Cited	Domain	RQ1	RQ2	RQ3
Yi et al. (2017b)	×	10	12	Not Specified	✓	✓	×
Kotova (2017)	×	13	15	Not Specified	✓	✓	×
Meenakshi et al. (2017)	×	14	18	Not Specified	✓	✓	×
Maravanyika et al. (2017)	✓	28	52	Not Specified	✓	✓	×
Ouf et al. (2017)	✓	114	154	Not Specified	✓	✓	×
Andaloussi et al. (2017)	✓	70	14	Not Specified	✓	✓	×
Zhao et al. (2017)	×	11	2	Not Specified	✓	×	×
Al-Jadaa et al. (2017)	×	14	6	School	✓	✓	×
Hansen et al. (2017)	×	7	3	Adult Learning	×	✓	✓
Barria-Pineda et al. (2017)	✓	5	16	Higher Education	✓	✓	×
2016							
Bent et al. (2017)	×	32	16	Not Specified	✓	✓	×
Zhu et al. (2016)	×	5	1	Prof. Training	✓	✓	×
Ripin et al. (2016)	✓	21	0	Higher Education	✓	✓	×
Ferreira et al. (2016)	×	42	31	Higher Education	✓	✓	×
Cheng et al. (2016b)	×	11	0	Higher Education	✓	✓	×
Elghouch et al. (2016)	×	13	10	Higher Education	✓	✓	×
Abdo and Noureldien (2016)	✓	37	4	Not Specified	✓	×	×
Abyaa et al. (2016)	×	16	12	Adult Learning	✓	✓	×
Paiva et al. (2016)	×	47	81	Not Specified	✓	✓	×
Huang (2016)	×	15	13	Higher Education	✓	✓	✓
Huang et al. (2016)	×	44	30	Not Specified	✓	✓	×
Cheng et al. (2016a)	×	12	0	Not Specified	✓	×	×
Ishak and Ahmad (2016)	×	16	9	Not Specified	✓	✓	×
Iqbal and Nurjanah (2016)	×	16	1	Not Specified	✓	✓	×
Hashemi and Herbert (2016)	×	8	3	Not Specified	✓	✓	×
Gooch et al. (2016)	×	45	9	Not Specified	✓	✓	×
Bull et al. (2016)	×	46	89	Higher Education	✓	✓	×
Chaplot et al. (2016)	×	6	32	Not Specified	✓	✓	✓
Pelánek (2016)	✓	44	202	Not Specified	✓	✓	✓
Kurup et al. (2016)	✓	13	8	Not Specified	✓	×	✓
2015							
Aitdaoud et al. (2015)	×	32	5	Higher Education	✓	✓	×
Guettat and Farhat (2015)	×	31	9	Not Specified	✓	✓	×
Khenissi and Essalmi (2015)	×	22	8	Not Specified	✓	✓	×
Chou et al. (2015b)	✓	59	56	Higher Education	✓	✓	×
Cocea and Magoulas (2015)	×	92	16	Not Specified	✓	✓	×
Wang et al. (2015)	×	9	0	Not Specified	✓	✓	×
Luna et al. (2015)	×	33	68	School	✓	✓	×

Table 7 continued

References	Metadata				Research Questions		
	Survey	#Ref	#Cited	Domain	RQ1	RQ2	RQ3
Hu and Huang (2015)	×	20	6	Higher Education	✓	✓	×
Qinghong et al. (2015)	×	5	1	Not Specified	✓	✓	✓
Zielinski (2015)	×	9	6	Not Specified	✓	✓	×
Özyurt and Özyurt (2015)	✓	44	234	Not Specified	✓	✓	×
Descalço and Carvalho (2015)	×	19	5	Higher Education	×	✓	×
Chou et al. (2015a)	×	73	51	Not Specified	✓	✓	×
Lucht et al. (2015)	×	16	1	Not Specified	✓	✓	×
Nedungadi and Remya (2015)	×	19	29	Not Specified	✓	×	✓
Jones et al. (2015)	×	5	5	School	✓	✓	×
Chrysafiadi and Virvou (2015)	×	52	98	Not Specified	✓	✓	✓
Muldner and Burleson (2015)	×	88	74	Not Specified	✓	✓	×
2014							
Nguyen (2014a)	×	2	6	Not Specified	✓	×	×
Nguyen (2014b)	×	13	17	Not Specified	✓	✓	×
Ait Adda and Amar (2014)	×	19	2	Not Specified	×	✓	×
Mohamed et al. (2014)	×	25	5	Not Specified	✓	✓	×
Chaturvedi and Ezeife (2014)	×	8	4	Higher Education	×	✓	×
Aballay et al. (2014)	×	3	5	Not Specified	✓	✓	×
Thinakaran & Ali (2014)	✓	26	4	Not Specified	✓	×	×
Alotaibi and Bull (2014)	×	6	0	Higher Education	✓	✓	×
Bull et al. (2014)	×	10	21	Not Specified	✓	✓	×
Chatti et al. (2014)	×	8	13	Higher Education	✓	✓	×
Chrysafiadi and Virvou (2014)	×	33	10	Not Specified	✓	✓	✓
Khenissi et al. (2014)	×	8	13	Not Specified	✓	✓	×
Nesbit et al. (2014)	✓	32	74	Higher Education	✓	✓	×
Alshammari et al. (2014)	✓	60	66	Not Specified	✓	✓	×
Al-Rajhi et al. (2014)	×	20	17	Higher Education	✓	✓	×
Koedinger et al. (2014)	×	23	9	Not Specified	✓	×	×
Garkavijs et al. (2014)	×	3	2	Not Specified	✓	✓	×
Swertz et al. (2014)	×	13	6	Prof. Training	×	✓	×
Seffrin et al. (2014)	×	2	9	School	✓	×	×
Wagner et al. (2014)	×	37	110	Higher Education	✓	✓	×
Clemente et al. (2014)	×	21	38	Not Specified	✓	×	✓
Ortigosa et al. (2014b)	×	57	792	Not Specified	✓	✓	×
Sani and Aris (2014a)	×	15	8	Not Specified	✓	✓	×
Sani and Aris (2014b)	✓	59	27	Not Specified	✓	✓	×
Peña-Ayala et al. (2014)	×	36	96	Not Specified	✓	✓	×

Overview of analysed publications—second iteration (reporting date 01/12/2024)

See Table 8.

Table 8 Overview of analysed publications—second iteration (reporting date 01/12/2024)

References	Metadata				Research Questions		
	Survey	#Ref	#Cited	Domain	RQ1	RQ2	RQ3
Karuru et al. (2024)	×	11	12	Higher Education	✓	×	×
Abdennour et al. (2023)	×	38	–	Not Specified	✓	✓	×
Cheung et al. (2023)	×	–	–	Not Specified	✓	✓	×
Abyaa et al. (2019)	✓	187	133	Not Specified	✓	✓	×
Mawas et al. (2018)	×	18	30	Not Specified	✓	✓	×
Crow et al. (2018)	✓	29	242	Not Specified	✓	✓	×
Bodily et al. (2018)	×	56	231	Not Specified	✓	✓	×
Tarus et al. (2017)	×	68	330	Not Specified	✓	✓	×
Hlioui et al. (2016)	✓	45	13	Not Specified	✓	✓	×
Bull and Kay (2016)	×	110	232	Not Specified	✓	✓	×
Rani et al. (2015)	×	52	212	Not Specified	×	✓	×
Nakic et al. (2015)	✓	111	152	Not Specified	✓	✓	×
Cook et al. (2015)	×	27	37	Lifelong Learning	×	✓	×
Yudelson et al. (2013)	×	8	675	Not Specified	×	✓	×
Grubišić et al. (2013)	×	25	38	Not Specified	×	✓	×
Chrysafiadi and Virvou (2013)	✓	159	471	Not Specified	✓	✓	×
Jeremić et al. (2012)	×	52	136	Not Specified	✓	✓	×
Desmarais and Baker (2011)	✓	136	594	Not Specified	✓	✓	×
Guangbing and Graf (2010)	×	16	47	Not Specified	✓	✓	×
Bull and Kay (2010)	×	53	335	Not Specified	×	✓	×
Reformat and Golmohammadi (2009)	×	23	12	Not Specified	×	✓	✓
Brusilovsky and Millán (2007)	×	210	1552	Not Specified	✓	✓	×
Gil-Olarte Márquez et al. (2006)	×	55	711	School	×	✓	×
Wenger (1987)	×	–	3408	Not Specified	✓	✓	✓
Woolf and McDonald (1984)	×	4	217	Not Specified	✓	×	×
Rich (1979)	×	14	1844	Not Specified	✓	✓	×
Ausubel et al. (1968)	×	1443	22676	Not Specified	×	×	✓

Overview of Complete UpSet Plots

UpSet plots without any thresholds that filter out minor occurrences.

See Figs. 20, 21 and 22.

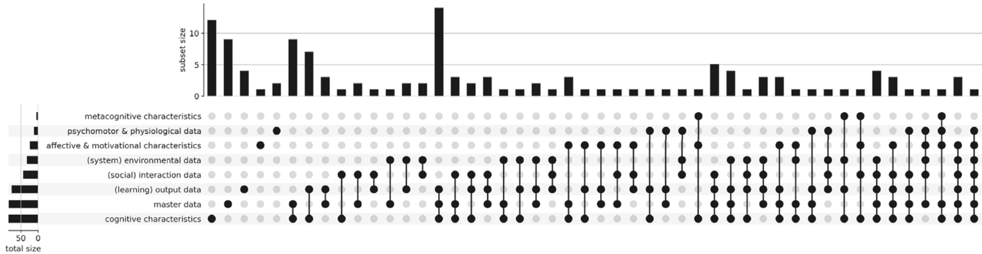

Fig. 20 Frequency of use of (combinations of) main data categories for components of the learner model. Overall source frequency—left histogram; frequency of combination of sources—top histogram.

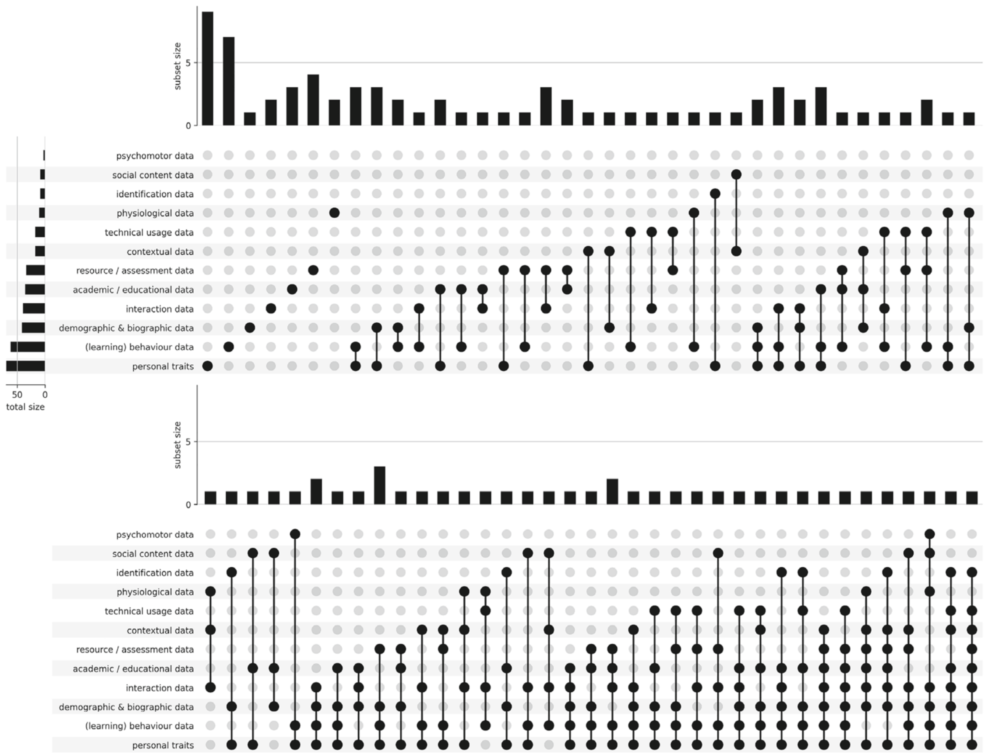

Fig. 21 Frequency of use of (combinations of) data categories for components of the learner model. Overall source frequency—left histogram; frequency of combination of sources—top histogram.

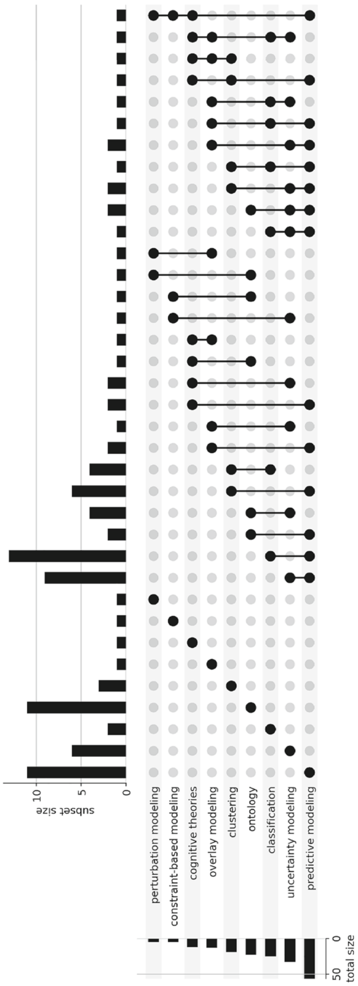

Fig. 22 Frequency of use of (combinations of) modelling approaches that appear more than once. Overall approach frequency—left histogram; Frequency of combination of approaches—top histogram.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11257-025-09434-4.

Acknowledgements This research is part of the VoLL-KI project and funded by the German Ministry of Education and Research (*Bundesministerium für Bildung und Forschung*) under grant 16DHBKI090.

Author Contributions F. Böck and M. Ochs wrote the main manuscript and contributed equally, while all remaining authors reviewed the manuscript.

Data availability All research data on which this systematic literature review is based are available via the university information system (https://fis.uni-bamberg.de/handle/uniba/109163) and the university research data repository (https://fd-repo.uni-bamberg.de/records/ctdrk-0rr96) at: https://doi.org/10.48564/unibafd-ctdrk-0rr96.

Declarations

Competing interests The authors declare no competing interests.

References

1EdTech. 1EdTech learner information package specification (2001)

Aballay, L.N., Collazos, C.A., Aciar, S.V., González, C.S.: Adaptation model content based in cultural profile into learning environment. In: International Conference on Human Computer Interaction, Interacción '14, New York, NY, USA. Association for Computing Machinery (2014). https://doi.org/10.1145/2662253.2662306

Abdennour, O., Kemouss, H., Erradi, M., Khaldi, M.: From learner profile management to learner model modeling. In: Khaldi, M. (ed.) Advances in Educational Technologies and Instructional Design, pp. 71–93. IGI Global (2023). https://doi.org/10.4018/978-1-6684-7634-5.ch004

Abdi, S., Khosravi,H., Sadiq, S.: Modelling learners in adaptive educational systems: A multivariate Glicko-based approach. In: LAK21: 11th International Learning Analytics and Knowledge Conference, LAK21, New York, NY, USA, pp 497–503. Association for Computing Machinery (2021). https://doi.org/10.1145/3448139.3448189

Abdo, N., Noureldien, N.A.: Evolution of dynamic user models for adaptive educational hypermedia system (AEHS). In: 2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT), pp. 1–5 (2016). https://doi.org/10.1109/ICAICT.2016.7991810

Abrams, M.: The origins of personal data and its implications for governance (2014)

Abyaa, A., Idrissi, M.K., Bennani. S.: Towards an adult learner model in an online learning environment. In: International Conference on Information Technology Based Higher Education and Training (ITHET), pp. 1–5 (2016). https://doi.org/10.1109/ITHET.2016.7760735

Abyaa, A., Idrissi, M.K., Bennani, S.: An adult learner's knowledge model based on ontologies and rule reasoning. In: Mediterranean Symposium on Smart City Application, SCAMS '17, New York, NY, USA. Association for Computing Machinery (2017). https://doi.org/10.1145/3175628.3175656

Abyaa, A., Idrissi,M.K., Bennani, S.: Predicting the learner's personality from educational data using supervised learning. In: Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications, SITA'18, New York, NY, USA, pp. 1–7. Association for Computing Machinery (2018). https://doi.org/10.1145/3289402.3289519

Abyaa, A., Khalidi Idrissi, M., Bennani, S.: Learner modelling: systematic review of the literature from the last 5 years. Educ. Technol. Res. Dev. **67**(5), 1105–1143 (2019). https://doi.org/10.1007/s11423-018-09644-1

Afini Normadhi, N.B., Shuib, L., Md Nasir, H.N., Bimba, A., Idris, N., Balakrishnan, V.: 03. Identification of personal traits in adaptive learning environment: Systematic literature review. Comput. Educ. **130**, 168–190 (2019). https://doi.org/10.1016/j.compedu.2018.11.005

Ahmadaliev, D., Xiaohui, C., Abduvohidov,M., Medatov, A., Temirova,G.: An adaptive activity sequencing instrument to enhance e-learning: an integrated application of overlay user model and mathematical

programming on the web. In: 2019 International Conference on Computer and Information Sciences (ICCIS), pp. 1–4 (2019). https://doi.org/10.1109/ICCISci.2019.8716473
Aissaoui, O.E., Madani, Y.E.A.E., Oughdir, L., Allioui, Y.E.: 03. Combining supervised and unsupervised machine learning algorithms to predict the learners' learning styles. In: International Conference on Intelligent Computing in Data Sciences (ICDS) **148**, 87–96 (2018). https://doi.org/10.1016/j.procs.2019.01.012
Ait Adda, S., Amar, B.: Enrichment of learner profile through the semantic analysis of the learner's query on the web. In: International Conference on the Innovative Computing Technology (INTECH 2014), pp. 7–12 (2014). https://doi.org/10.1109/INTECH.2014.6927773
Aitdaoud, M., Guerss, F.Z., Lmati,I., Namir, A., Talbi,M., Benlahmar, H., Achtaich, N., Douzi, K.: Standardized modeling learners ontology ≪SLMonto≫ to support interactive learning environment. In: International Conference on Information & Communication Technology and Accessibility (ICTA), pp. 1–6 (2015). https://doi.org/10.1109/ICTA.2015.7426899
Akharraz, L., El Mezouary,A., Mahani, Z.: To context-aware learner modeling based on ontology. In: Global Engineering Education Conference (EDUCON), pp. 1326–1334 (2018).https://doi.org/10.1109/EDUCON.2018.8363383
Al-Abri, A., Al-Khanjari,Z., Kraiem, N., Jamoussi, Y.: A scheme for extracting information from collaborative social interaction tools for personalized educational environments. In: 2017 International Conference on Computing Networking and Informatics (ICCNI), pp. 1–6 (2017). https://doi.org/10.1109/ICCNI.2017.8123795
Al-Chalabi, H.K.M., Hussein,A.M., Apoki, U.C.: An adaptive learning system based on learner's knowledge level. In: 2021 13th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), pp. 1–4 (2021).https://doi.org/10.1109/ECAI52376.2021.9515158
Al-Jadaa, A.A., Abu-Issa, A.S., Ghanem, W.T., Hussein, M.S.: Enhancing the intelligence of web tutoring systems using a multi-entry based open learner model. In: International Conference on Internet of Things, Data and Cloud Computing, ICC '17, New York, NY, USA. Association for Computing Machinery (2017). https://doi.org/10.1145/3018896.3036389
Al-Rajhi, L., Salama, R., Gamalel-Din, S.: Personalized intelligent assessment model for measuring initial students abilities. In: Workshop on Interaction Design in Educational Environments, IDEE '14, New York, NY, USA, pp. 41–48. Association for Computing Machinery (2014). https://doi.org/10.1145/2643604.2643606
Al-Shabandar, R., Hussain, A.J., Liatsis, P., Keight, R.: Detecting at-risk students with early interventions using machine learning techniques. IEEE Access **7**, 149464–149478 (2019). https://doi.org/10.1109/ACCESS.2019.2943351
Alajlani, N., Crabb, M., Murray, I.: A systematic review in understanding stakeholders' role in developing adaptive learning systems. J. Comput. Educ. **11**(3), 901–920 (2024). https://doi.org/10.1007/s40692-023-00283-x
Alatrash, R., Priyadarshini, R., Ezaldeen, H., Alhinnawi, A.: A hybrid recommendation integrating semantic learner modelling and sentiment multi-classification. J. Web Eng. **21**(4), 941–988 (2022). https://doi.org/10.13052/jwe1540-9589.2141
Alday, R.B.: Bayesian networks in intelligent tutoring systems as an assessment of student performance using student modeling. In: International Conference on Algorithms, Computing and Systems, ICACS '18, New York, NY, USA, pp. 119–122. Association for Computing Machinery (2018). https://doi.org/10.1145/3242840.3242867
Alfonseca, E., Carro, R.M., Martín, E., Ortigosa, A., Paredes, P.: The impact of learning styles on student grouping for collaborative learning: a case study. User Model. User-Adap. Inter. **16**, 377–401 (2006). https://doi.org/10.1007/s11257-006-9012-7
Almeida, T.O., Netto, J.F. de M.: Adaptation content in robotic systems: A systematic mapping study. In: Frontiers in Education Conference (FIE), pp. 1–9 (2018). https://doi.org/10.1109/FIE.2018.8658539
Almeida, T.O., De Magalhães Netto, J.F., Lopes, A.M.M.: Multi-agent system for recommending learning objects in e-learning environments. In: 2021 IEEE Frontiers in Education Conference (FIE), pp. 1–5 (2021). https://doi.org/10.1109/FIE49875.2021.9637258
Almeida, T.O., de Magalhães Netto, J.F., Lopes, A.M.M.: Recommendation model of personalized teaching materials in e-learning environments. In: Frontiers in Education Conference (FIE), pp. 1–7 (2021). https://doi.org/10.1109/FIE49875.2021.9637286

Alotaibi, M., Bull, S.: Discussion around individual open learner models: understanding or copying? In: International Conference on Advanced Learning Technologies (ICALT), pp. 82–83 (2014). https://doi.org/10.1109/ICALT.2014.34

Alshammari, M., Anane, R., Hendley, R.J.: Adaptivity in e-learning systems. In: International Conference on Complex, Intelligent and Software Intensive Systems, pp. 79–86 (2014). https://doi.org/10.1109/CISIS.2014.12

Andaloussi, K.S., Capus, L., Berrada, I.: Adaptive educational hypermedia systems: current developments and challenges. In: Proceedings of the 2nd International Conference on Big Data, Cloud and Applications, BDCA'17, New York, NY, USA, pp. 1–8. Association for Computing Machinery (2017). https://doi.org/10.1145/3090354.3090448

Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y.: An integrated theory of the mind. Psychol. Rev. **111**(4), 1036–1060 (2004). https://doi.org/10.1037/0033-295X.111.4.1036

Anwar, M.: Supporting privacy, trust, and personalization in online learning. Int. J. Artif. Intell. Educ. **31**(4), 769–783 (2021). https://doi.org/10.1007/s40593-020-00216-0

Ausubel, D.P., Novak, J.D., Hanesian, H., Ausubel, D.F.: Educational Psychology: A Cognitive View, 2nd edn. Holt, Rinehart & Winston, New York (1968)

Ayari, A., Chaabouni, M., Ghezala, H.B.: New learner model for intelligent and adaptive e-learning system. In: International Conference on Innovations in Intelligent Systems and Applications (INISTA), pp. 1–7 (2023). https://doi.org/10.1109/INISTA59065.2023.10310323

Barria-Pineda, J., Guerra, J., Huang, Y., Brusilovsky, P.: Concept-level knowledge visualization for supporting self-regulated learning. In: International Conference on Intelligent User Interfaces Companion, IUI '17 Companion, New York, NY, USA, pp. 141–144. Association for Computing Machinery (2017). https://doi.org/10.1145/3030024.3038262

Barria-Pineda, J., Guerra-Hollstein, J., Brusilovsky, P.: A fine-grained open learner model for an introductory programming course. In: Conference on User Modeling, Adaptation and Personalization, UMAP '18, New York, NY, USA, pp. 53–61. Association for Computing Machinery (2018). https://doi.org/10.1145/3209219.3209242

Bellarhmouch, Y., Jeghal, A., Benjelloun, N., Tairi, H.: Overview of personalization approaches in MOOCs. In: International Conference on Intelligent Systems and Computer Vision (ISCV), pp. 1–7 (2022). https://doi.org/10.1109/ISCV54655.2022.9806079

Bent, O., Dey, P., Weldemariam, K., Mohania, M.K.: Modeling user behavior data in systems of engagement. Future Gener. Comput. Syst. **68**, 456–464 (2017). https://doi.org/10.1016/j.future.2016.05.038

Bernard, J., Popescu, E., Graf, S.: Improving online education through automatic learning style identification using a multi-step architecture with ant colony system and artificial neural networks. Appl. Soft Comput. **131**, 109779 (2022). https://doi.org/10.1016/j.asoc.2022.109779

Beyyoudh, M., Idrissi, M.K., Bennani, S.: A new approach of designing an intelligent tutoring system based on adaptive workflows and pedagogical games. In: International Conference on Information Technology Based Higher Education and Training (ITHET), pp. 1–7 (2018). https://doi.org/10.1109/ITHET.2018.8424619

Bobadilla, J., Ortega, F., Hernando, A., Bernal, J.: A collaborative filtering approach to mitigate the new user cold start problem. Knowl.-Based Syst. **26**, 225–238 (2012). https://doi.org/10.1016/j.knosys.2011.07.021

Böck, F.: A research agenda for learner models for adaptive educational digital learning environments. In: Auer, M.E., Rüütmann, T. (eds.) Futureproofing Engineering Education for Global Responsibility. Springer Nature Switzerland, Cham (2025)

Böck, F., Ochs, M.: Research Data Repository. Research information system of the University of Bamberg: https://fis.uni-bamberg.de/handle/uniba/109163. Research Data Repository of the University of Bamberg: https://fd-repo.uni-bamberg.de/records/ctdrk-0rr96. (2025). https://doi.org/10.48564/unibafd-68754-p3573

Böck, F., Landes, D., Sedelmaier, Y.: Learner models: A systematic literature research in norms and standards:. In: Proceedings of the 16th International Conference on Computer Supported Education, Angers, France, pp. 187–196. SCITEPRESS—Science and Technology Publications (2024). https://doi.org/10.5220/0012556100003693

Böck, F., Deuerling, A., Landes, D.: Adaptive learning environment reference architecture for an optimised learning process. In: 2025 IEEE Global Engineering Education Conference (EDUCON). IEEE (2025). https://doi.org/10.1109/EDUCON62633.2025.11016496

Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., Verbert, K. Open learner models and learning analytics dashboards: a systematic review. In: Proceedings of the 8th International Conference on Learning Analytics and Knowledge, Sydney New South Wales Australia, pp. 41–50. ACM (2018). https://doi.org/10.1145/3170358.3170409

Brereton, P., Kitchenham, B.A., Budgen, D., Turner, M., Khalil, M.: Lessons from applying the systematic literature review process within the software engineering domain. J. Syst. Softw. **80**(4), 571–583 (2007). https://doi.org/10.1016/j.jss.2006.07.009

Brickley, D., Miller, L.: FOAF vocabulary specification 0.99—namespace document 14 January 2014—paddington edition (2014)

Brusilovsky, P.: Methods and techniques of adaptive hypermedia. User Model. User-Adapt. Interact. **6**(2), 87–129 (1996). https://doi.org/10.1007/BF00143964

Brusilovsky, P.: Adaptive hypermedia. User Model. User-Adap. Inter. **11**(1), 87–110 (2001). https://doi.org/10.1023/A:1011143116306

Brusilovsky, P., Millán, E.: User models for adaptive hypermedia and adaptive educational systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web, vol. 4321, pp. 3–53. Springer, Berlin (2007). https://doi.org/10.1007/978-3-540-72079-9_1

Brusilovsky, P., Sosnovsky, S., Shcherbinina, O.: User modeling in a distributed e-learning architecture. In: Ardissono, L., Brna, P., and Mitrovic, A. (eds.) User Modeling 2005. pp. 387–391. Springer, Berlin, Heidelberg (2005).https://doi.org/10.1007/11527886_50

Bull, S.: Supporting learning with open learner models. In: Proceedings of fourth Hellenic Conference on Information and Communication Technologies in Education, Athen, Griechenland, pp. 47–61 (2004)

Bull, S.: There are open learner models about! IEEE Trans. Learn. Technol. **13**(2), 425–44 (2020). https://doi.org/10.1109/TLT.2020.2978473

Bull, S., Kay, J.: Student models that invite the learner in: The SMILI:() open learner modelling framework. Int. J. Artif. Intell. Educ. **17**(2), 89–120 (2007)

Bull, S., Kay, J.: Open learner models. In: Nkambou, R., Bourdeau, J., Mizoguchi, R. (eds.) Advances in Intelligent Tutoring Systems, vol. 308, pp. 301–322. Springer, Berlin (2010). https://doi.org/10.1007/978-3-642-14363-2_15

Bull, S., Kay, J.: Smili: a framework for interfaces to learning data in open learner models, learning analytics and related fields. Int. J. Artif. Intell. Educ. **26**(1), 293–33 (2016). https://doi.org/10.1007/s40593-015-0090-8

Bull, S., Pain, H.: 'Did I say what I think I said, and do you agree with me?': Inspecting and questioning the student Model. In: Proceedings of World Conference on Artificial Intelligence in Education, Charlottesville, VA, pp. 501–508. AACE (1995)

Bull, S., Pain, H.: "Did I say what I think I said, and do you agree with me?": Inspecting and questioning the student model. University of Edinburgh, Department of Artificial Intelligence (1995)

Bull, S., Greer, J., McCalla,G., Kettel, L., Bowes, J.L: User modelling in i-help: What, why, when and how. In: Bauer, M., Gmytrasiewicz, P.J., Vassileva, J. (eds.) User Modeling 2001, vol. 2109, pp. 117–126. Springer, Berlin (2001). https://doi.org/10.1007/3-540-44566-8_12

Bull, S., Johnson, M.D., Epp, C.D., Masci, D., Alotaibi, M., Girard., S.: Formative assessment and meaningful learning analytics. In: 2014 IEEE 14th International Conference on Advanced Learning Technologies, pp. 327–32 (2014).https://doi.org/10.1109/ICALT.2014.100

Bull, S., Ginon,B., Boscolo, C., Johnson, M.: Introduction of learning visualisations and metacognitive support in a persuadable open learner model. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, LAK '16, New York, NY, USA, pp. 30–39. Association for Computing Machinery (2016). https://doi.org/10.1145/2883851.2883853

Cassidy, S.: Learning styles: an overview of theories, models, and measures. Educ. Psychol. **24**, 419–444 (2004). https://doi.org/10.1080/0144341042000228834

CELTS-11. CELTS-11 - learner model specification (2012)

Chaplot, D.S., Rhim, E., Kim, J.: Personalized adaptive learning using neural networks. In: Proceedings of the Third (2016) ACM Conference on Learning @ Scale, L@S '16, New York, NY, USA, pp. 165–168. Association for Computing Machinery (2016). https://doi.org/10.1145/2876034.2893397

Chatti, M.A., Dugoija,D., Thüs, H., Schroeder, U..: Learner modeling in academic networks. In: 2014 IEEE 14th International Conference on Advanced Learning Technologies, pp. 117–121 (2014). https://doi.org/10.1109/ICALT.2014.42

Chaturvedi, R., Ezeife, C.: Mining relevant examples for learning in ITS student models. In: International Conference on Computer and Information Technology (CIT), pp. 743–750 (2014). https://doi.org/10.1109/CIT.2014.31
Chaturvedi, R., Ezeife, C.I.: Predicting student performance in an ITS using task-driven features. In: International Conference on Computer and Information Technology (CIT), pp. 168–175 (2017). https://doi.org/10.1109/CIT.2017.34
Cheng, Y., Zeng, G., Xu, W.S., Weng, Q.Y.: Personalized learning research on interest-driven in virtual learning community. In: 2016 International Conference on Information System and Artificial Intelligence (ISAI), pp. 561–565 (2016a).https://doi.org/10.1109/ISAI.2016.0125
Cheng, Y., Zeng, G., Xu, W.S., YuanWeng, Q.: Research on the interest mining and cluster grouping based on the knowledge point. In: 2016 6th International Conference on Information Communication and Management (ICICM), pp. 7–12 (2016b). https://doi.org/10.1109/INFOCOMAN.2016.7784206
Cheung, S.K., Wang, F.L., Kwok, L.F., Poulová, P.: Personalized Learning: Approaches, Methods and Practices, 1st edn. Routledge, London (2023)
Chignell, M., Hancock, P.: Intelligent Interface Design, Handbook of Human–Computer Interaction, pp. 969–995. Elsevier, New York (1988). https://doi.org/10.1016/B978-0-444-70536-5.50051-8
Chou, C.Y., Lai, K.R., Chao, P.Y., Lan, C.H., Chen, T.H.: Negotiation based adaptive learning sequences: combining adaptivity and adaptability. Comput. Educ. **88**, 215–226 (2015). https://doi.org/10.1016/j.compedu.2015.05.007
Chou, C.Y., Tseng, S.F., Chih, W.C., Chen, Z.H., Chao, P.Y., Lai, K.R., Chan, C.L., Yu, L.C., Lin, Y.L.: Open student models of core competencies at the curriculum level: using learning analytics for student reflection. IEEE Trans. Emerg. Top. Comput. **5**(1), 32–44 (2015). https://doi.org/10.1109/TETC.2015.2501805
Chrysafiadi, K., Virvou, M.: Evaluating the integration of fuzzy logic into the student model of a web-based learning environment. Expert Syst. Appl. **39**, 13127–13134 (2012). https://doi.org/10.1016/j.eswa.2012.05.089
Chrysafiadi, K., Virvou, M.: Student modeling approaches: a literature review for the last decade. Expert Syst. Appl. **40**(11), 4715–4729 (2013). https://doi.org/10.1016/j.eswa.2013.02.007
Chrysafiadi, K., Virvou, M.: K.E.M.Cs: A set of student's characteristics for modeling in adaptive programming tutoring systems. In: International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 106–110 (2014). https://doi.org/10.1109/IISA.2014.6878786
Chrysafiadi, K., Virvou, M.: Fuzzy logic for adaptive instruction in an e-learning environment for computer programming. IEEE Trans. Fuzzy Syst. **23**(1), 164–177 (2015). https://doi.org/10.1109/TFUZZ.2014.2310242
Chrysafiadi, K., Troussas, C., Virvou, M.: Combination of fuzzy and cognitive theories for adaptive e-assessment. Expert Syst. Appl. **161**, 113614 (2020). https://doi.org/10.1016/j.eswa.2020.113614
Chrysafiadi, K., Papadimitriou, S., Virvou, M.: Cognitive-based adaptive scenarios in educational games using fuzzy reasoning. Knowl.-Based Syst. **250**, 109111 (2022). https://doi.org/10.1016/j.knosys.2022.109111
Clabaugh, C., Ragusa, G., Sha, F., Mataric, M.: Designing a socially assistive robot for personalized number concepts learning in preschool children. In: 2015 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDLEpiRob). pp. 314–319. IEEE, Providence, RI, USA (2015). https://doi.org/10.1109/DEVLRN.2015.7346164
Clemente, J., Ramírez, J., Antonio, A.d.: Applying a student modeling with non-monotonic diagnosis to intelligent virtual environment for training/instruction. Expert Syst. Appl. **41**(2), 508–520 (2014). https://doi.org/10.1016/j.eswa.2013.07.077
Cocea, M., Magoulas, G.D.: Participatory learner modelling design: a methodology for iterative learner models development. Inf. Sci. **321**, 48–70 (2015). https://doi.org/10.1016/j.ins.2015.05.032
Conati, C., Gertner, A., Vanlehn, K.: Using bayesian networks to manage uncertainty in student modeling. User Model. User-Adapt. Interact. **12**, 371–417 (2002). https://doi.org/10.1023/A:1021258506583
Conati, C., Porayska-Pomsta, K., Mavrikis, M.: AI in education needs interpretable machine learning: lessons from Open Learner Modelling. In: ICML Workshop on Human Interpretability in Machine Learning (WHI 2018), Stockholm, Sweden (2018)
Conejo, R., Trella, M., Cruces, I., Garcia, R.: INGRID: A web service tool for hierarchical open learner model visualization. In: Ardissono, L. and Kuflik, T. (eds.) Advances in User Modeling. pp. 406–409. Springer, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28509-7_38

Cook, R., Kay, J., Kummerfeld, B.: MOOClm: user modelling for MOOCs. In: Ricci, F., Bontcheva, K., Conlan, O., Lawless, S. (eds.) User Modeling, Adaptation and Personalization, pp. 80–91 (2015). Springer, Cham.https://doi.org/10.1007/978-3-319-20267-9_7

Corbett, A.T., Anderson, J.R.: Knowledge tracing: modeling the acquisition of procedural knowledge. User Model. User-Adapt. Interact. **4**(4), 253–278 (1994). https://doi.org/10.1007/BF01099821

Corredor, C.M., Franco, M.A., del Pilar García Gutiérrez, Z.: Adaptive system for non-literate older adult learning. In: International Conference on Advanced Learning Technologies (ICALT), pp. 33–35 (2023). https://doi.org/10.1109/ICALT58122.2023.00015

Crow, T., Luxton-Reilly, A., Wuensche, B.: Intelligent tutoring systems for programming education: a systematic review. In: Proceedings of the 20th Australasian Computing Education Conference, Brisbane, Queensland, Australia, pp. 53–62. ACM (2018). https://doi.org/10.1145/3160489.3160492

Cully, A., Demiris, Y.: Online knowledge level tracking with data-driven student models and collaborative filtering. IEEE Trans. Knowl. Data Eng. **32**(10), 2000–2013 (2020). https://doi.org/10.1109/TKDE.2019.2912367

David, J., Lobov, A., Lanz, M.: Attaining learning objectives by ontological reasoning using digital twins. In: Conference on Learning Factories, vol. 31, pp. 349–355. Procedia Manufacturing (2019). https://doi.org/10.1016/j.promfg.2019.03.055

de Koch, N.P.: Software engineering for adaptive hypermedia systems—reference model, modeling techniques and development process. Ph.D. thesis, Ludwig-Maximilians-Universität München, München (2000)

Descalço, L., Carvalho, P.: Using parameterized calculus questions for learning and assessment. In: Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–5 (2015). https://doi.org/10.1109/CISTI.2015.7170428

Desmarais, M.C., Baker, R.S.J.D.: A review of recent advances in learner and skill modeling in intelligent learning environments. User Model. User-Adapt. Interact. **22**(1), 9–38 (2011). https://doi.org/10.1007/s11257-011-9106-8

Dijkstra, S.H.E., Hinne, M., Segers, E., Molenaar, I.: Clustering children's learning behaviour to identify self-regulated learning support needs. Comput. Hum. Behav. **145**, 1–22 (2023). https://doi.org/10.1016/j.chb.2023.107754

Dillenbourg, P., Self, J.: A framework for learner modelling. Interact. Learn. Environ. **2**(2), 111–137 (1992). https://doi.org/10.1080/1049482920020202

Ding, W., Zhu, Z., Guo, Q.: A new learner model in adaptive learning system. In: 2018 3rd International Conference on Computer and Communication Systems (ICCCS), pp. 440–443 (2018). https://doi.org/10.1109/CCOMS.2018.8463316

Dolog, P., Schäfer, M.: A framework for browsing, manipulating and maintaining interoperable learner profiles. In: Ardissono, L., Brna, P., and Mitrovic, A. (eds.) User Modeling 2005. pp. 397–401. Springer, Berlin, Heidelberg (2005). https://doi.org/10.1007/11527886_52

Ebbinghaus, H.: Ueber das gedächtnis. Mind **10**(39), 454–459 (1885)

Ed-Fi Alliance.: Ed-fi unifying data model (UDM) (2023)

Effenberger, T., Pelánek, R.: Validity and reliability of student models for problem-solving activities. In: LAK21: 11th International Learning Analytics and Knowledge Conference, pp. 1–11. ACM (2021).https://doi.org/10.1145/3448139.3448140

Elghouch, N., En-Naimi, E.M., Seghroucheni, Y.Z., El Mohajir, B.E., Achhab, M.A.: ALS_corr[LP]: An adaptive learning system based on the learning styles of felder-silverman and a bayesian network. In: International Colloquium on Information Science and Technology (CiSt), pp. 494–499 (2016). https://doi.org/10.1109/CIST.2016.7805098

Emerson, A., Sawyer,R., Azevedo, R., Lester, J.: Gaze-enhanced student modeling for game-based learning. In: Conference on User Modeling, Adaptation and Personalization, UMAP '18, New York, NY, USA, pp. 63–72. Association for Computing Machinery (2018). https://doi.org/10.1145/3209219.3209238

Ennouamani, S.,Mahani, Z.: A comparative study of the learner model in adaptive mobile learning systems. In: International Conference on Networking, Information Systems & Security, NISS '19, New York, NY, USA. Association for Computing Machinery (2019). https://doi.org/10.1145/3320326.3320351

Essalmi, F., Ayed, L.J.B., Jemni, M., Kinshuk, Graf, S.: A fully personalization strategy of E-learning scenarios. Comput Human Behav. **26**, 581–591 (2010). https://doi.org/10.1016/j.chb.2009.12.010

European Commission.: Justice and consumers article 29—Guidelines on the right to "data portability" (2017)

Ezaldeen, H., Bisoy, S.K., Misra, R., Alatrash, R.: Semantics-aware context-based learner modelling using normalized PSO for personalized e-learning. J. Web Eng. **21**(4), 1187–1224 (2022). https://doi.org/10.13052/jwe1540-9589.2148

Faucon, L., Olsen, J.K., Dillenbourg, P.: A bayesian model of individual differences and flexibility in inductive reasoning for categorization of examples. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, LAK '20, New York, NY, USA, pp. 285–294. Association for Computing Machinery (2020). https://doi.org/10.1145/3375462.3375512

Fei Zhou, T., Qing Pan,Y, Huang, L.R..: Research on personalized e-learning based on decision tree and RETE algorithm. In: 2017 International Conference on Computer Systems, Electronics and Control (ICCSEC), pp. 1392–1396 (2017). https://doi.org/10.1109/ICCSEC.2017.8446741

Felder, R.: Learning and teaching styles in engineering education. J. Eng. Educ. **78**, 674–681 (1988)

Ferreira, H.N.M., Araújo, R.D., Dorça, F.A., Cattelan, R.G.: Open student modeling for academic performance visualization in ubiquitous learning environments. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 641–646 (2017). https://doi.org/10.1109/SMC.2017.8122679

Ferreira, H.N.M., Brant-Ribeiro, T., Araújo, R.D., Dorça, F.A, Cattelan, R.G..: An automatic and dynamic student modeling approach for adaptive and intelligent educational systems using ontologies and bayesian networks. In: International Conference on Tools with Artificial Intelligence (ICTAI), pp. 738–745 (2016). https://doi.org/10.1109/ICTAI.2016.0116

Ferreira, H.N.M., Brant-Ribeiro, T., Araújo, R.D., Dorça, F.A., Cattelan, R.G.: An automatic and dynamic knowledge assessment module for adaptive educational systems. In: International Conference on Advanced Learning Technologies (ICALT), pp. 517–521 (2017). https://doi.org/10.1109/ICALT.2017.86

Ganesan, S., Somasiri, N., Pokhrel, S.: The role of artificial intelligence in education. In: International Conference on Computer Communication and Informatics (ICCCI), pp. 1–6 (2023). https://doi.org/10.1109/ICCCI56745.2023.10128180

Garkavijs, V., Okamoto,R., Ishikawa, T., Toshima, M., Kando, N.: GLASE-IRUKA: Gaze feedback improves satisfaction in exploratory image search. In: International Conference on World Wide Web, WWW '14 Companion, New York, NY, USA, pp. 273–274. Association for Computing Machinery (2014). https://doi.org/10.1145/2567948.2577313

Gasmi, S., Bouhadada, T., Kamilya, L.: Recommendation of learning resources in social learning environment using data mining. In: International Conference on Networking and Advanced Systems (ICNAS), pp. 1–7 (2021). https://doi.org/10.1109/ICNAS53565.2021.9628979

Gautam, S., Rosson, M.B., Akgun, M.: Exploring potential contributions of social learning to adaptive learning systems. In: Conference on Human Factors in Computing Systems, CHI EA '23, New York, NY, USA. Association for Computing Machinery (2023). https://doi.org/10.1145/3544549.3585758

GB/T 29805-2013.: GB/t 29805-2013 information technology—learning, education and training—learner model (2014)

Germanakos, P., Belk, M.: User Modeling, pp. 79–102. Springer, Cham (2016)

Ghallabi, S., Essalmi, F., Jemni, M., Kinshuk: Optimal composition of e-leaming personalization parameters. In: 2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA). pp. 1–4. IEEE, Marrakech (2015). https://doi.org/10.1109/ICTA.2015.7426921

Gil-Olarte Márquez, P., Palomera Martín, R., Brackett, M.A.: Relating emotional intelligence to social competence and academic achievement in high school students. Psicothema **18**(Suppl), 118–123 (2006)

Ginon, B., Boscolo, C., Johnson, M.D., Bull, S.: Persuading an open learner model in the context of a university course: An exploratory study. In: Micarelli, A., Stamper, J., and Panourgia, K. (eds.) Intelligent Tutoring Systems. pp. 307–313. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-39583-8_34

Gooch, D., Benton, L., Khaled, R., Lukeš, D., Vasalou, A.: Creating bridges: the role of exploratory design research in an intelligent tutoring system project. Interact. Comput. **28**(3), 372–386 (2016). https://doi.org/10.1093/iwc/iwv009

Greer, J.E., McCalla, G.I. (eds.): Student Modelling: The Key to Individualized Knowledge-Based Instruction. Springer, Berlin (1994)

Grubišić, A., Stankov, S., Peraić, I.: Ontology based approach to bayesian student model design. Expert Syst. Appl. **40**(13), 5363–5371 (2013). https://doi.org/10.1016/j.eswa.2013.03.041

Guangbing, Y., Kinshuk, Graf, S.: A practical student model for a address-aware and context-sensitive personalized adaptive learning system. In: 2010 International Conference on Technology for Education, Mumbai, India, pp. 130–133. IEEE (2010). https://doi.org/10.1109/T4E.2010.5550048

Guettat, B., Farhat, R.: An approach to assist learners to identify their learning objectives in personal learning environment (PLE). In: International Conference on Information & Communication Technology and Accessibility (ICTA), pp. 1–6 (2015). https://doi.org/10.1109/ICTA.2015.7426934

Hadj-M'tir, R.: Ingénierie ontologique pour la réutilisation à base de cas dans les systèmes de e-learning, (2010)

Hambleton, R.K., Swaminathan, H.: Item Response Theory: Principles and Applications. Springer, Dordrecht (1985)

Hansen, C.J.S., Wasson, B., Skretting, H., Netteland, G., Hirnstein, M.: When learning is high stake. In: Proceedings of the Seventh International Learning Analytics & Knowledge Conference, LAK '17, New York, NY, USA, pp. 564–565. Association for Computing Machinery (2017). https://doi.org/10.1145/3027385.3029461

Hashemi, M., Herbert, J.: A pro-active and dynamic prediction assistance using BaranC framework. In: International Conference on Mobile Software Engineering and Systems (MOBILESoft), pp. 269–270 (2016). https://doi.org/10.1145/2897073.2897759

He, X., Liu, P., Zhang, W.: Design and implementation of a unified mooc recommendation system for social work major: experiences and lessons. In: International Conference on Computational Science and Engineering (CSE) and International Conference on Embedded and Ubiquitous Computing (EUC), vol. 1, pp. 219–223 (2017). https://doi.org/10.1109/CSE-EUC.2017.46

Heidig, S., Clarebout, G.: Do pedagogical agents make a difference to student motivation and learning? Educ. Res. Rev. **6**, 27–54 (2011). https://doi.org/10.1016/j.edurev.2010.07.004

Hlioui, F., Alioui, N., Gargouri, F.: A survey on learner models in adaptive e-learning systems. In: 13th International Conference of Computer Systems and Applications, AICCSA 2016, Proceedings of International Conference on Computer Systems and Applications. IEEE Computer Society (2016). https://doi.org/10.1109/AICCSA.2016.7945770

Honey, P., Mumford, A.: The Manual of Learning Styles, 3rd edn. P. Honey, Maidenhead (1992)

Hooshyar, D., Pedaste, M., Saks, K., Leijen, Ä., Bardone, E., Wang, M.: Open learner models in supporting self-regulated learning in higher education: a systematic literature review. Comput. Educ. **154**, 103878 (2020). https://doi.org/10.1016/j.compedu.2020.103878

Hu, P.: Investigation on smart campus management platform based on digital twin. In: International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy, vol. 228, pp. 937–945. Procedia Computer Science (2023). https://doi.org/10.1016/j.procs.2023.11.123

Hu, Q., Huang, Y.: The design of open learner model to improve interaction of peer assessment in online learning. In: International Conference on Computer Science & Education (ICCSE), pp. 310–314 (2015). https://doi.org/10.1109/ICCSE.2015.7250262

Huang, Y.: Deeper knowledge tracing by modeling skill application context for better personalized learning. In: Conference on User Modeling Adaptation and Personalization, UMAP '16, New York, NY, USA, pp. 325–328. Association for Computing Machinery (2016). https://doi.org/10.1145/2930238.2930373

Huang, Y., Yudelson, M., Han, S., He, D., Brusilovsky, P.: A framework for dynamic knowledge modeling in textbook-based learning. In: Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, UMAP '16, New York, NY, USA, pp. 141–150. Association for Computing Machinery (2016). https://doi.org/10.1145/2930238.2930258

Huang, S., Yin, B., Liu, M.: Research on individualized learner model based on context-awareness. In: 2017 International Symposium on Educational Technology (ISET), pp. 163–167 (2017). https://doi.org/10.1109/ISET.2017.45

IEEE 9274.1.1-2023. [IEEE 9274.1.1-2023] IEEE standard for learning technology–JavaScript object notation (JSON) data model format and representational state transfer (RESTful) web service for learner experience data tracking and access (2023)

IEEE P1484.2.: [IEEE p1484.2]—standard for information technology—learning systems—learner model (1997)

IEEE P9274.4.1.: [IEEE p9274.4.1] recommended practice for the implementation of the experience application programming interface (API) 1.0.3 (2022)

Internet2.: EduPerson specification (200312) (2003)

Iqbal, Y., Nurjanah, D.: Adaptive navigation with knowledge tracing method. In: 2016 4th International Conference on Information and Communication Technology (ICoICT), pp. 1–6 (2016). https://doi.org/10.1109/ICoICT.2016.7571879

Ishak, M.K., Ahmad, N.B.: Enhancement of learning management system with adaptive features. In: International Student Project Conference (ICT-ISPC), pp. 37–40 (2016). https://doi.org/10.1109/ICT-ISPC.2016.7519230

ISO/IEC 19479:2019.: [ISO/IEC 19479:2019] information technology for learning, education, and training—learner mobility achievement information (LMAI) (2019)

ISO/IEC 19788-9:2015.: [ISO/IEC 19788-9:2015] information technology—learning, education and training—metadata for learning resources—part 9: data elements for persons (2015)

ISO/IEC 29140:2021.: [ISO/IEC 29140:2021] information technology for learning, education and training—nomadicity and mobile technologies (2021)

ISO/IEC TR 20748-1:2016.: [ISO/IEC TR 20748-1:2016] information technology for learning, education and training—learning analytics interoperability—part 1: reference model (2016)

JADN.: Specification for JSON abstract data notation (JADN) version 1.0 (2021)

Janati, S.E., Maach, A., Ghanami, D.E.: SMART education framework for adaptation content presentation. In: International Conference on Intelligent Computing in Data Sciences (ICDS), vol. 127, pp. 436–443. Procedia Computer Science. https://doi.org/10.1016/j.procs.2018.01.141

Jeremić, Z., Jovanović, J., Gašević, D.: Student modeling and assessment in intelligent tutoring of software patterns. Expert Syst. Appl. **39**(1), 210–222 (2012). https://doi.org/10.1016/j.eswa.2011.07.010

Jing, S., Tang, Y., Liu, X., Gong, X.: A learner model integrating cognitive and metacognitive and its application on scratch programming projects. IFAC-PapersOnLine **53**(5), 644–649 (2020). https://doi.org/10.1016/j.ifacol.2021.04.154

Jing, S., Tang,Y., Liu, X., Gong, X., Cui, W., Liang, J.: A parallel education based intelligent tutoring systems framework. In: International Conference on Networking, Sensing and Control (ICNSC), pp. 1–6 (2020b). https://doi.org/10.1109/ICNSC48988.2020.9238052

Johnson, A., Taatgen, N.: User modeling. In: Proctor, R.W., Vu, K.P.L. (eds.) Handbook of Human Factors in Web Design, pp. 424–439. Lawrence Erlbaum Associates, Mahwah (2005)

Jonassen, D.H., Grabowski, B.L.: Handbook of Individual Differences, Learning, and Instruction, 1st edn. Routledge, London (1993)

Jones, A., Bull, S., Castellano, G.: Open learner modelling with a robotic tutor. In: International Conference on Human-Robot Interaction Extended Abstracts, HRI'15 Extended Abstracts, New York, NY, USA, pp. 237–238. Association for Computing Machinery (2015). https://doi.org/10.1145/2701973.2702713

Jun-min, Y., Song, X., Da-Xiong, L., Zhi-Feng, W., Peng-Wei, H., Chen, X.: Research on the construction and application of individual learner model. Procedia Comput. Sci. **131**, 88–92 (2018). https://doi.org/10.1016/j.procs.2018.04.189

Kakizaki, T., Oeda, S.: Student modeling considering learning behavior history with deep factorization machines. In: International Conference on Knowledge Based and Intelligent Information and Engineering Systems, vol. 225, pp. 2808–2815. Procedia Computer Science (2023). https://doi.org/10.1016/j.procs.2023.10.273

Kaliwal, R.B., Deshpande, S.L.: Chapter 16—Evaluate learner level assessment in intelligent e-learning systems using probabilistic network model. In: Pandey, R., Khatri, S.K., Singh, N.k., Verma, P. (eds.) Artificial Intelligence and Machine Learning for EDGE Computing, pp. 255–265. Academic Press, New York (2022). https://doi.org/10.1016/B978-0-12-824054-0.00029-0

Karuru, P., Sipahelut, J., Riyanti, R., Saleh, M., Makulua, K.: Development of technology-based learning models to enhance critical thinking skills in education students. Global Int. J. Innov. Res. **2**(1), 330–335 (2024). https://doi.org/10.59613/global.v2i1.53

Kavitha, P., Krishna Moorthy, B., Sudharshan, P., Aarthi, T.: Mapping artificial intelligence and education. In: 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT), pp. 165–168 (2018).https://doi.org/10.1109/IC3IoT.2018.8668123

Kay, J.: Lifelong learner modeling for lifelong personalized pervasive learning. IEEE Trans. Learn. Technol. **1**, 215–228 (2008). https://doi.org/10.1109/TLT.2009.9

Kay, J., Bull, S.: New opportunities with open learner models and visual learning analytics. In: Conati, C., Heffernan, N., Mitrovic, A., Verdejo, M.F. (eds.) Artificial Intelligence in Education, vol. 9112, pp. 666–669 (2015). Springer, Cham. https://doi.org/10.1007/978-3-319-19773-9_87

Kay, J., Bartimote, K., Kitto, K., Kummerfeld, B., Liu, D., Reimann, P.: Enhancing learning by open learner model (olm) driven data design. Comput. Educ. Artif. Intell. **3**, 100069 (2022). https://doi.org/10.1016/j.caeai.2022.100069

Khenissi, M.A., Essalmi, F.: Automatic generation of fuzzy logic components for enhancing the mechanism of learner's modeling while using educational games. In: International Conference on Information & Communication Technology and Accessibility (ICTA), pp. 1–6 (2015). https://doi.org/10.1109/ICTA.2015.7426879

Khenissi, M.A., Essalmi, F., Jemni, M., Kinshuk: Learners' working memory capacity modeling based on fuzzy logic. In: International Conference on Advanced Learning Technologies (ICALT), pp. 520–521 (2014). https://doi.org/10.1109/ICALT.2014.153

Khenissi, M.A., Essalmi, F., Jemni, M.: Kinshuk: learner modeling using educational games: a review of the literature. Smart Learn. Environ. **2**(1), 6 (2015). https://doi.org/10.1186/s40561-015-0014-y

Khosravi, H., Shum, S.B., Chen, G., Conati, C., Tsai, Y.S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., Gašević, D.: Explainable artificial intelligence in education. Comput. Educ. Artif. Intell. **3**, 100074 (2022). https://doi.org/10.1016/j.caeai.2022.100074

Kitchenham, B.: Procedures for undertaking systematic reviews. Joint technical report, Keele University (TR/SE0401) and National ICT Australia Ltd. (0400011T.1) (2004)

Kitchenham, B.A., Charters, S.: Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE 2007-001, Keele University (2007)

Koedinger, K.R., McLaughlin, E.A., Stamper, J.C.: Data-driven learner modeling to understand and improve online learning: MOOCs and technology to advance learning and learning research (ubiquity symposium). Ubiquity **2014**, 3:1-3:13 (2014). https://doi.org/10.1145/2591682

Kotova, E.E.: Intellectual data analysis in the educational process. In: International Conference on Soft Computing and Measurements (SCM), pp. 757–759 (2017). https://doi.org/10.1109/SCM.2017.7970714

Kurup, L.D., Joshi, A., Shekhokar, N.: A review on student modeling approaches in ITS. In: 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 2513–2517 (2016)

Lazarinis, F., Green, S., Koutromanos, G.: Web-based learner profile management using IMS LIP. In: INTED2009 Proceedings. pp. 1515–1521. IATED, Valencia, Spain (2009)

Lei, P.I., Mendes, A.J.: A systematic literature review on knowledge tracing in learning programming. In: 2021 IEEE Frontiers in Education Conference (FIE), pp. 1–7 (2021). https://doi.org/10.1109/FIE49875.2021.9637323

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., Pfister, H.: Upset: visualization of intersecting sets. IEEE Trans. Visual Comput. Graph. **20**(12), 1983–1992 (2014). https://doi.org/10.1109/TVCG.2014.2346248

Li, X., He, S.: Research and analysis of student portrait based on campus big data. In: International Conference on Big Data Analytics (ICBDA), pp. 23–27 (2021). https://doi.org/10.1109/ICBDA51983.2021.9403070

Li, Z., Jianqing, X., Tian, C., Xiaomeng. W.: CELTS—CELTS-11-learner model (2002)

Li, Q., Zhong, S., Wang, P., Guo, X., Quan, X.: Learner model in adaptive learning system. J. Inf. Comput. Sci. **7**, 5–1137 (2010)

Liang, J., Hare, R., Chang, T., Xu, F., Tang, Y., Wang, F.Y., Peng, S., Lei, M.: Student modeling and analysis in adaptive instructional systems. IEEE Access **10**, 59359–59372 (2022). https://doi.org/10.1109/ACCESS.2022.3178744

Lin, L., Wang, F.: Adaptive learning system based on knowledge graph. In: Proceedings of the 9th International Conference on Education and Training Technologies, ICETT '23, New York, NY, USA, pp. 1–7. Association for Computing Machinery (2023). https://doi.org/10.1145/3599640.3599647

Lin, L., Wang, F., Wang, F.: Research on learning resource recommendation based on learner model. In: Proceedings of the 2022 5th International Conference on Education Technology Management, ICETM '22, New York, NY, USA, pp. 45–50. Association for Computing Machinery (2023). https://doi.org/10.1145/3582580.3582589

Liu, F., Hu, X., Bu, C., Yu, K.: Fuzzy bayesian knowledge tracing. IEEE Trans. Fuzzy Syst. **30**, 2412–2425 (2022). https://doi.org/10.1109/TFUZZ.2021.3083177

Liu, Q., Wu, J., Huang, Z., Wang, H., Ning, Y., Chen, M., Chen, E., Yi, J., Zhou, B.: Federated user modeling from hierarchical information. ACM Trans. Inf. Syst. **41**(2), 1–33 (2023). https://doi.org/10.1145/3560485

Lu, H., Niu, W., Caverlee, J.: Learning geo-social user topical profiles with bayesian hierarchical user factorization. In: International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR '18, New York, NY, USA, pp. 205–214. Association for Computing Machinery (2018). https://doi.org/10.1145/3209978.3210044

Lucht, M., Pharow, P., Belonovskaya, I.D.: CoALa—a collaborative analysis lab for an identification of user parameters for adaptive e-leaming systems. In: Global Engineering Education Conference (EDUCON), pp. 1008–1012 (2015). https://doi.org/10.1109/EDUCON.2015.7096096

Luna, V., Quintero, R., Torres, M., Moreno-Ibarra, M., Guzmán, G., Escamilla, I.: An ontology-based approach for representing the interaction process between user profile and its context for collaborative learning environments. Comput. Hum. Behav. **51**, 1387–1394 (2015). https://doi.org/10.1016/j.chb.2014.10.004

Luxing, L.: Student aesthetic education recommendation app based on interest. In: 2020 13th International Conference on Intelligent Computation Technology and Automation (ICICTA), pp. 631–634 (2020). https://doi.org/10.1109/ICICTA51737.2020.00140

Lwande, C., Muchemi, L., Oboko, R.: Identifying learning styles and cognitive traits in a learning management system. Heliyon **7**(8), e07701 (2021). https://doi.org/10.1016/j.heliyon.2021.e07701

MACE.: eduPerson (202208) v4.4.0 (2022)

Mannekote, A., Davies, A., Pinto, J.D., Zhang, S., Olds, D., Schroeder, N.L., Lehman, B., Zapata-Rivera, D., Zhai, C.: Large language models for whole-learner support: opportunities and challenges. Front. Artif. Intell. **7**, 1460364 (2024). https://doi.org/10.3389/frai.2024.1460364

Maravanyika, M., Dlodlo,N., Jere, N.: An adaptive recommender-system based framework for personalised teaching and learning on e-learning platforms. In: IST-Africa Week Conference, pp. 1–9 (2017). https://doi.org/10.23919/ISTAFRICA.2017.8102297

Martinez-Maldonado, R., Echeverria, V., Fernandez Nieto, G., Buckingham Shum, S.: From data to insights: A layered storytelling approach for multimodal learning analytics. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–15. ACM, Honolulu HI USA (2020). https://doi.org/10.1145/3313831.3376148

Martins, G.S., Santos, L., Dias, J.: Bum: bayesian user model for distributed learning of user characteristics from heterogeneous information. IEEE Trans. Cognit. Dev. Syst. **11**(3), 425–434 (2018). https://doi.org/10.1109/TCDS.2018.2878451

Martynov, V.V., Filosova, E.I., Zvereva, N.N.: Computer technologies for modeling the competencies of employees of industrial enterprises as part of a digital twin of a person. In: International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), pp. 911–915 (2023). https://doi.org/10.1109/ICIEAM57311.2023.10139249

Mawas, N.E., Ghergulescu, I., Moldovan, A.N., Muntean, C.H. : Pedagogical based learner model characteristics. In: Ireland International Conference on Education, Dublin, Irelandal-02251323 (2018)

Mayo, M.J.: Bayesian student modelling and decision-theoretic selection of tutorial actions in intelligent tutoring systems. Ph.D. thesis, University of Canterbury (2001)

Meenakshi, K., Sunder, R., Kumar, A., Sharma, N.: An intelligent smart tutor system based on emotion analysis and recommendation engine. In: International Conference on IoT and Application (ICIOT), pp. 1–4 (2017). https://doi.org/10.1109/ICIOTA.2017.8073608

Mejia, C., Gomez, S., Mancera, L., Taveneau, S.: Inclusive learner model for adaptive recommendations in virtual education. In: International Conference on Advanced Learning Technologies (ICALT), pp. 79–80 (2017). https://doi.org/10.1109/ICALT.2017.101

Miller, T.: Explanation in artificial intelligence: insights from the social sciences. Artif. Intell. **267**, 1–38 (2019). https://doi.org/10.1016/j.artint.2018.07.007

Minn, S.: Ai-assisted knowledge assessment techniques for adaptive learning environments. Comput. Educ. Artif. Intell. **3**, 100050 (2022). https://doi.org/10.1016/j.caeai.2022.100050

Minn, S., Yu, Y., Desmarais,M.C., Zhu, F., Vie, J.J.: Deep knowledge tracing and dynamic student classification for knowledge tracing. In: International Conference on Data Mining (ICDM), pp. 1182–1187 (2018a). https://doi.org/10.1109/ICDM.2018.00156

Minn, S., Zhu, F., Desmarais, M.C.: Improving knowledge tracing model by integrating problem difficulty. In: International Conference on Data Mining Workshops (ICDMW), pp. 1505–1506 (2018b). https://doi.org/10.1109/ICDMW.2018.00220

Miranda, S.K.O., Marcelino,M.J., Silva, P.A.: An ontology to trace the computer science student profile. In: Frontiers in Education Conference (FIE), College Station, TX, USA, pp. 1–5. IEEE (2023). https://doi.org/10.1109/FIE58773.2023.10342999

Miraz, M.H., Ali, M., Excell, P.S.: Adaptive user interfaces and universal usability through plasticity of user interface design. Comput. Sci. Rev. **40**, 100363 (2021). https://doi.org/10.1016/j.cosrev.2021.100363

Mohamed, H., Ahmad, N.B.H., Shamsuddin, S.M.H.: Bijective soft set classification of student's learning styles. In: Malaysian Software Engineering Conference (MySEC), pp. 289–294 (2014). https://doi.org/10.1109/MySec.2014.6986031

Mokhtar, R., Abdullah, S.N.H.S., Zin, N.A.M.: Classifying modality learning styles based on Production-Fuzzy Rules. In: 2011 International Conference on Pattern Analysis and Intelligence Robotics. pp. 154–159. IEEE, Kuala Lumpur, Malaysia (2011). https://doi.org/10.1109/ICPAIR.2011.5976887

Muldner, K., Burleson, W.: Utilizing sensor data to model students' creativity in a digital environment. Comput. Hum. Behav. **42**, 127–137 (2015). https://doi.org/10.1016/j.chb.2013.10.060

Nakic, J., Granic, A., Glavinic, V.: Anatomy of student models in adaptive learning systems: a systematic literature review of individual differences from 2001 to 2013. J. Educ. Comput. Res. **51**(1), 459–489 (2015). https://doi.org/10.2190/EC.51.4.e

Narimani, A., Barberà, E.: Extracting course features and learner profiling for course recommendation systems: a comprehensive literature review. Int. Rev. Res. Open Distrib. Learn. **25**(1), 197–225 (2024). https://doi.org/10.19173/irrodl.v25i1.7419

NASSP. Student Learning Styles. Diagnosing and Prescribing Programs. NASSP, 1904 Association Drive, Reston, VA 22091 (1979)

Nedungadi, P., Remya, M.S.: ncorporating forgetting in the personalized, clustered, bayesian knowledge tracing (PC-BKT) model. In: 2015 International Conference on Cognitive Computing and Information Processing(CCIP), Noida, India, pp. 1–5. IEEE (2015). https://doi.org/10.1109/CCIP.2015.7100688

Neji, M., Ben Ammar, M., Alimi, A.M.: Real-time affective learner profile analysis using an EMASPEL framework. In: 2011 IEEE Global Engineering Education Conference (EDUCON). pp. 664–670. IEEE, Amman, Jordan (2011). https://doi.org/10.1109/EDUCON.2011.5773210

Nesbit, J.C., Adesope, O.O., Liu, Q., Ma. W.: How effective are intelligent tutoring systems in computer science education? In: 2014 IEEE 14th International Conference on Advanced Learning Technologies, pp. 99–103 (2014). https://doi.org/10.1109/ICALT.2014.38

Nguyen, L.: Evaluating adaptive learning model. In: International Conference on Interactive Collaborative Learning (ICL), pp. 818–822 (2014a). https://doi.org/10.1109/ICL.2014.7017878

Nguyen, L.: A user modeling system for adaptive learning. In: 2014 International Conference on Interactive Collaborative Learning (ICL), pp. 864–866 (2014b). https://doi.org/10.1109/ICL.2014.7017887

Nurjanah, D.: LifeOn, a ubiquitous lifelong learner model ontology supporting adaptive learning. In: Global Engineering Education Conference (EDUCON), pp. 866–871 (2018). https://doi.org/10.1109/EDUCON.2018.8363321

Ochukut, S.A., Oboko, R.: A learner model for adaptive e-learning based on learning theories. In: IST-Africa Week Conference, pp. 1–8 (2019). https://doi.org/10.23919/ISTAFRICA.2019.8764826

OECD.: DSTI/ICCP/REG(2014)3 - Working party on security and privacy in the digital economy (2014)

Oeda, S., Hasegawa, S.: A student modeling method combining deep learning and forgetting models with knowledge tracing. In: International Conference on Knowledge Based and Intelligent Information and Engineering Systems, vol. 225, pp. 1191–1200. Procedia Computer Science (2023). https://doi.org/10.1016/j.procs.2023.10.107

Oeda, S., Kakizaki, T.: Latent state estimation method for students using knowledge tracing combining with irt. Procedia Comput. Sci. **207**, 1992–1999 (2022). https://doi.org/10.1016/j.procs.2022.09.258

Oeda, S., Shimizu, D.: Verification of usefulness of student modeling with real educational data using convex factorization machines. Procedia Comput. Sci. **192**, 804–811 (2021). https://doi.org/10.1016/j.procs.2021.08.083

Ohlsson, S.: Constraint-Based Student Modeling, Student Modelling: The Key to Individualized Knowledge-Based Instruction, Number 125 in NATO ASI Series. Springer, Berlin (1994)

Ortigosa, A., Carro, R.M., Quiroga, J.I.: Predicting user personality by mining social interactions in facebook. J. Comput. Syst. Sci. **80**(1), 57–71 (2014). https://doi.org/10.1016/j.jcss.2013.03.008

Ortigosa, A., Martín, J.M., Carro, R.M.: Sentiment analysis in facebook and its application to e-learning. Comput. Hum. Behav. **31**, 527–541 (2014). https://doi.org/10.1016/j.chb.2013.05.024

Ortony, A., Clore, G.L., Collins, A.: The Cognitive Structure of Emotions. Cambridge University Press, Cambridge (1988)

Oubahssi, L., Grandbastien, M.: From learner information packages to student models: Which continuum? In: Ikeda, M., Ashley, K.D., and Chan, T.-W. (eds.) Intelligent Tutoring Systems. pp. 288–Springer Berlin Heidelberg, Berlin, Heidelberg (2006). https://doi.org/10.1007/11774303_29

Ouf, S., Abd Ellatif, M., Salama, S.E., Helmy, Y.: A proposed paradigm for smart learning environment based on semantic web. Comput. Hum. Behav. **72**, 796–818 (2017). https://doi.org/10.1016/j.chb.2016.08.030
Özyurt, Ö., Özyurt, H.: Learning style based individualized adaptive e-learning environments: content analysis of the articles published from 2005 to 2014. Comput. Hum. Behav. **52**, 349–358 (2015). https://doi.org/10.1016/j.chb.2015.06.020
Paiva, R., Bittencourt, I.I., Tenório, T., Jaques, P., Isotani, S.: What do students do on-line? modeling students' interactions to improve their learning experience. Comput. Hum. Behav. **64**, 769–781 (2016). https://doi.org/10.1016/j.chb.2016.07.048
Panagiotopoulos, I., Kalou, A., Pierrakeas, C., Kameas, A.: An ontology-based model for student representation in intelligent tutoring systems for distance learning. In: Iliadis, L., Maglogiannis, I., and Papadopoulos, H. (eds.) Artificial Intelligence Applications and Innovations. pp. 296–305. Springer Berlin Heidelberg, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33409-2_31
Papanikolaou, K.A., Grigoriadou, M., Kornilakis, H., Magoulas, G.D.: Personalizing the interaction in a web-based educational hypermedia system: the case of inspire. User Model. User Adapt. Interact. **13**, 213–267 (2003). https://doi.org/10.1023/A:1024746731130
Papadimitriou, S., Chrysafiadi, K., Kamitsios, M., Virvou, M.: FSP creator: a novel web service API creator of fuzzy students progress profile. In: International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 1–7 (2018). https://doi.org/10.1109/IISA.2018.8633662
Park, S.T., Chu, W.: Pairwise preference regression for cold-start recommendation. In: Proceedings of the third ACM Conference on Recommender Systems, New York, New York, USA, pp. 21–28. ACM (2009). https://doi.org/10.1145/1639714.1639720
Pavlik, P.I., Jr., Brawner, K., Olney, A., Mitrovic, A.: Design recommendations for intelligent tutoring systems: Learner modeling, vol. 1. U.S, Army Research Laboratory (2013)
Pelánek, R.: Metrics for evaluation of student models. J. Educ. Data Min. **7**(2), 1–19 (2015). https://doi.org/10.5281/zenodo.3554665
Pelánek, R.: Applications of the elo rating system in adaptive educational systems. Comput. Educ. **98**, 169–179 (2016). https://doi.org/10.1016/j.compedu.2016.03.017
Pelánek, R.: Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques. User Model. User-Adapt. Interact. **27**(3–5), 313–350 (2017). https://doi.org/10.1007/s11257-017-9193-2
Pelánek, R.: Adaptive learning is hard: challenges, nuances, and trade-offs in modeling. Int. J. Artif. Intell. Educ. **35**(1), 304–329 (2025). https://doi.org/10.1007/s40593-024-00400-6
Peña-Ayala, A., Sossa, H., Méndez, I.: Activity theory as a framework for building adaptive e-learning systems: a case to provide empirical evidence. Comput. Hum. Behav. **30**, 131–145 (2014). https://doi.org/10.1016/j.chb.2013.07.057
Piao, G., Breslin, J.G.: Inferring user interests in microblogging social networks: a survey. User Model. User-Adapt. Interact. **28**(3), 277–329 (2018). https://doi.org/10.1007/s11257-018-9207-8
Pratiwi, O.N., Fa'rifah, R.Y., Pamungkas, I.R., Megasari, R., Callista, A.S.: Adaptive e-assessment framework. In: International Conference on Advancement in Data Science, E-learning and Information System (ICADEIS), pp. 1–5 (2023). https://doi.org/10.1109/ICADEIS58666.2023.10270915
Pucheta, H.M.F., Morales, V.Y.R., Velázquez. Y.H.: User modeling for collaborative adaptative mobile educational applications: a systematic review of the literature. In: Mexican International Conference on Computer Science (ENC), pp. 1–7 (2022). https://doi.org/10.1109/ENC56672.2022.9882919
Pustovalova, N.V., Avdeenko, T.V., Pustovalova, A.V.: University's educational environment personalization based on the ontological models. In: International Conference of Young Professionals in Electron Devices and Materials (EDM), pp. 289–294 (2022). https://doi.org/10.1109/EDM55285.2022.9855106
Qiao, C., Hu, X.: Discovering student behavior patterns from event logs: preliminary results on a novel probabilistic latent variable model. In: International Conference on Advanced Learning Technologies (ICALT), pp. 207–211 (2018). https://doi.org/10.1109/ICALT.2018.00056
Qinghong, Y., Jing,X., Dule, Y.: The research of personalized learning system based on ontology. In: 2015 10th International Conference on Computer Science & Education (ICCSE), pp. 471–475 (2015). https://doi.org/10.1109/ICCSE.2015.7250292
Qiu, W., Du, J., Li.F.: Knowledge description frame of learning resources for recommendation system: From the perspectives of learning psychology. In: International Conference on Computer Science & Education (ICCSE), pp. 438–440 (2020). https://doi.org/10.1109/ICCSE49874.2020.9201686

Rahayu, N.W., Ferdiana, R., Kusumawardani, S.S.: A systematic review of ontology use in e-learning recommender system. Comput. Educ. Artif. Intell. **3**, 100047 (2022). https://doi.org/10.1016/j.caeai.2022.100047

Rahman, M.M., Salem, H.A.A., Yeom, S., Ollington,N., Ollington, R., Rahman, M.M.: A preliminary study on learners' personal traits for modelling learner profiles in ITS: a sensor-free approach. In: Symposium on Computer Applications & Industrial Electronics (ISCAIE), pp. 287–292 (2023). https://doi.org/10.1109/ISCAIE57739.2023.10165219

Raj, N.S., Renumol, V.G.: Architecture of an adaptive personalized learning environment (APLE) for content recommendation. In: International Conference on Digital Technology in Education, ICDTE '18, New York, NY, USA, pp. 17–22. Association for Computing Machinery (2018). https://doi.org/10.1145/3284497.3284503

Ramírez Luelmo, S.I., El Mawas,N., Heutte, J.: Towards open learner models including the flow state. In: Conference on User Modeling, Adaptation and Personalization, UMAP '20 Adjunct, New York, NY, USA, pp. 305–310. Association for Computing Machinery (2020). https://doi.org/10.1145/3386392.3399295

Rani, M., Nayak, R., Vyas, O.P.: An ontology-based adaptive personalized e-learning system, assisted by software agents on cloud storage. Knowl.-Based Syst. **90**, 33–48 (2015). https://doi.org/10.1016/j.knosys.2015.10.002

Reckase, M.: Development and application of a multivariate logistic latent trait model. Ph.D. thesis, Syracuse University (1972)

Reformat, M., Golmohammadi, S.K.: Updating user profile using ontology-based semantic similarity. In: 2009 IEEE International Conference on Fuzzy Systems, Jeju Island, South Korea, pp. 1062–1067. IEEE (2009). https://doi.org/10.1109/FUZZY.2009.5277205

Rich, E.: User modeling via stereotypes. Cogn. Sci. **3**(4), 329–354 (1979). https://doi.org/10.1016/S0364-0213(79)80012-9

Ripin, F.H.M., Mokhtar, R., Zain, W.M.: Adaptive tutoring system user preferences: does gender matter? In: International Conference on Information and Communication Technology for The Muslim World (ICT4M), pp. 24–29 (2016). https://doi.org/10.1109/ICT4M.2016.018

Rowe, J.P., Lester, J.C.: Artificial intelligence for personalized preventive adolescent healthcare. J. Adolesc. Health **67**(2), S52–S58 (2020). https://doi.org/10.1016/j.jadohealth.2020.02.021

Rueden, L., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Pfrommer, J., Pick, A., Ramamurthy, R., Walczak, M., Garcke, J., Bauckhage, C., Schuecker, J.: Informed machine learning - a taxonomy and survey of integrating prior knowledge into learning systems. IEEE Trans. Knowl. Data Eng. **35**, 614–633 (2023). https://doi.org/10.1109/TKDE.2021.3079836

Rupp, A.A., Templin, J., Henson, R.A.: Diagnostic Measurement: Theory, Methods, and Applications. Theory, Methods, and Applications. Guilford Press, New York, Diagnostic Measurement (2010)

Sani, S., Aris, T.N.: Proposal for ontology based approach to fuzzy student model design. In: International Conference on Intelligent Systems, Modelling and Simulation, pp. 35–37 (2014a). https://doi.org/10.1109/ISMS.2014.14

Sani, S., Aris, T.N.B.M.: Computational intelligence approaches for student/tutor modelling: a review. In: International Conference on Intelligent Systems, Modelling and Simulation, pp. 72–76 (2014b). https://doi.org/10.1109/ISMS.2014.21

Santos, O.C., Saneiro, M., Salmeron-Majadas, S., Boticario, J.G.: A methodological approach to eliciting affective educational recommendations. In: 2014 IEEE 14th International Conference on Advanced Learning Technologies. pp. 529–533. IEEE, Athens, Greece (2014). https://doi.org/10.1109/ICALT.2014.234

Saurabh, R., Singh, N.C., Duraiappah, A.K.: The alternate education for the 21st century. In: 2019 IEEE 18th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), pp. 265–271 (2019). https://doi.org/10.1109/ICCICC46617.2019.9146061

Schiaffino, S., Garcia, P., Amandi, A.: Eteacher: providing personalized assistance to e-learning students. Computers & Education. **51**, 1744–1754 (2008). https://doi.org/10.1016/j.compedu.2008.05.008

Schmidmaier, M., Han, Z., Weber, T., Liu, Y., Hussmann, H.: Real-time personalization in adaptive IDEs. In: Conference on User Modeling, Adaptation and Personalization, UMAP'19 Adjunct, New York, NY, USA, pp. 81–86. Association for Computing Machinery (2019). https://doi.org/10.1145/3314183.3324975

Schulz, S., Lingnau. A.: An evidence-based learner model for supporting activities in robotics. In: Proceedings of the Seventh ACM Conference on Learning @ Scale, L@S '20, New York, NY, USA, pp. 397–400. Association for Computing Machinery (2020). https://doi.org/10.1145/3386527.3406760

Seffrin, H.M., Rubi, G.L., Jaques, P.A.: A dynamic bayesian network for inference of learners' algebraic knowledge. In: Symposium on Applied Computing, SAC '14, New York, NY, USA, pp. 235–240. Association for Computing Machinery (2014). https://doi.org/10.1145/2554850.2555062

Self, J.A.: Student models in computer-aided instruction. Int. J. Man Mach. Stud. **6**(2), 261–276 (1974). https://doi.org/10.1016/s0020-7373(74)80005-2

Self, J.A.: Student models and artificial intelligence. Comput. Educ. **3**(4), 309–312 (1979). https://doi.org/10.1016/0360-1315(79)90008-3

Shen, L., Wang, M., Shen, R.: Affective e-learning: Using "emotional" data to improve learning in pervasive learning environment. J. Educ. Technol. Soc. **12**, 176–189 (2009)

Shute, V.J., Zapata-Rivera, D.: Adaptive educational systems. In: Durlach, P.J., Lesgold, A.M. (eds.) Adaptive Technologies for Training and Education, 1 edn., pp. 7–27. Cambridge University Press, Cambridge (2012). https://doi.org/10.1017/CBO9781139049580.004

Smythe, C., Tansey, F., Robson, R.: IMS learner information packaging information model (2001)

Sottilare, R., Holden, H., Goldberg, B., Brawner, K.: Chapter 20: the generalized intelligent framework for tutoring (GIFT). In: Fundamental Issues in Defence Simulation & Training. Ashgate Publishing **1**(1), 223–234 (2013)

Sottilare, R.A., Schwarz, J. (eds.): Adaptive instructional systems: Second international conference, AIS 2020, held as part of the 22nd HCI international conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Volume 12214 (2020) of Lecture Notes in Computer Science. Springer, Berlin

Stamper, J., Xiao, R., Hou, X.: Enhancing LLM-based feedback: Insights from intelligent tutoring systems and the learning sciences. In: Olney, A.M., Chounta, I.A., Liu, Z., Santos, O.C., Bittencourt, I.I. (eds.) Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky, Volume 2150 of Series Title: Communications in Computer and Information Science, pp. 32–43 (2024). Springer Nature Switzerland, Cham. https://doi.org/10.1007/978-3-031-64315-6_3

Stansfield, J.L., Carr, B.P., Goldstein, I.P.: Wumpus advisor 1: A first implementation program that tutors logical and probabilistic reasoning skills. Technical report, MIT (1976)

Stash, N., Cristea, A.I., Bra, P.M.E. de: Adaptation to learning styles in E-learning : approach evaluation. In: Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. pp. 284–291. AACE, Honolulu, Hawaii (2006)

Sun, G., Cui, T., Yong, J., Shen, J., Chen, S.: Mlaas: a cloud-based system for delivering adaptive micro learning in mobile mooc learning. IEEE Trans. Serv. Comput. **11**(2), 292–30 (2018). https://doi.org/10.1109/TSC.2015.2473854

Sun, H., Huang, L., Chen, Z., Zheng, Q.: Research on collaborative filtering recommendation method of online learning resources based on learner model. In: 2022 IEEE 5th International Conference on Information Systems and Computer Aided Education (ICISCAE), pp. 202–206 (2022a). https://doi.org/10.1109/ICISCAE55891.2022.9927652

Sun, H., Zheng,Q., Zhang, H., Pan, H.: Learner modeling framework based on learning analytics. In: 2022 IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA), pp. 64–68 (2022b). https://doi.org/10.1109/ICDSCA56264.2022.9988517

Sun, J.C.Y., Tsai, H.E., Cheng, W.K.R.: Effects of integrating an open learner model with ai-enabled visualization on students' self-regulation strategies usage and behavioral patterns in an online research ethics course. Comput. Educ. Artif. Intell. **4**, 100120 (2023). https://doi.org/10.1016/j.caeai.2022.100120

Swertz, C., Barberi, A., Forstner,A., Schmölz,A., Henning, P., Heberle, F.: A SMS with a KISS. towards a pedagogical design of a metadata system for adaptive learning pathways. In: Jarmo, V., Marianna, L. (eds.) EdMedia 2014. World conference on educational media & technology, June 23–26, Tampere, Finland, EdMedia Innovate Learning—Proceedings; 2014, pp. 991–1000. Waynesville: Association for the Advancement of Computing in Education 2014 (2014). https://doi.org/10.25656/01:15980

Tacoma, S., Geurts, C., Slof, B., Jeuring, J., Drijvers, P.: Enhancing learning with inspectable student models: worth the effort? Comput. Hum. Behav. **107**, 106276 (2020). https://doi.org/10.1016/j.chb.2020.106276

Tadlaoui, M.A., Khaldi, M., Carvalho, R.N.: Bayesian networks for managing learner models in adaptive hypermedia systems: Emerging research and opportunities. Advances in Educational Technologies and Instructional Design, IGI Global (2019)

Tang, Q., Abel, M.H., Negre, E.: Towards the privacy-preserving of online recommender system in collaborative learning environment. In: International Conference on Systems, Man, and Cybernetics (SMC), pp. 1298–1303 (2020). https://doi.org/10.1109/SMC42975.2020.9283463

Tang, Q., Abel, M.H. , Negre, E.: Improve learner-based recommender system with learner's mood in online learning platform. In: International Conference on Machine Learning and Applications (ICMLA), pp. 1704–1709 (2021). https://doi.org/10.1109/ICMLA52953.2021.00271

Tarus, J.K., Niu, Z., Yousif, A.: A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. Futur. Gener. Comput. Syst. **72**, 37–48 (2017). https://doi.org/10.1016/j.future.2017.02.049

Taufik, R., Nurjanah, D.: An intelligent tutoring system with adaptive exercises based on a student's knowledge and misconception. In: International Conference on Engineering, Technology and Education (TALE), pp. 1–5 (2019). https://doi.org/10.1109/TALE48000.2019.9226001

Tetzlaff, L., Schmiedek, F., Brod, G.: Developing personalized education. A dynamic framework. Educ. Psychol. Rev. **33**(3), 863–882 (2020). https://doi.org/10.1007/s10648-020-09570-w

Thinakaran, R., Ali, R.: Review on constraint based modelling in intelligent learning environment. In: International Conference on Engineering Technology and Technopreneuship (ICE2T), pp. 32–36 (2014). https://doi.org/10.1109/ICE2T.2014.7006214

Thohiroh, E.L., Nurjanah, D.: SmartMe: An open social learner model and case example recommender. In: 2022 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE), pp. 331–336 (2022). https://doi.org/10.1109/TALE54877.2022.00061

Van Labeke, N., Brna, P., Morales, R.: Opening up the interpretation process in an open learner model. Int. J. Artificial Intell. Educ. **17**, 305–338 (2007)

Vildjiounaite, E., Kallio, J., Kantorovitch, J., Kinnula,A., Ferreira, S., Rodrigues,M.A. , Rocha, N.: Challenges of learning human digital twin: case study of mental wellbeing: using sensor data and machine learning to create HDT. In: International Conference on Pervasive Technologies Related to Assistive Environments, PETRA '23, New York, NY, USA, pp. 574–583. Association for Computing Machinery (2023). https://doi.org/10.1145/3594806.3596538

Vo, D.V., Karnjana, J., Huynh, V.N.: An integrated framework of learning and evidential reasoning for user profiling using short texts. Inf. Fus. **70**, 27–42 (2021). https://doi.org/10.1016/j.inffus.2020.12.004

W3C.: Making content usable for people with cognitive and learning disabilities (2021)

Wagner, A., Barbosa, J.L.V., Barbosa, D.N.F.: A model for profile management applied to ubiquitous learning environments. Expert Syst. Appl. **41**(4), 2023–2034 (2014). https://doi.org/10.1016/j.eswa.2013.08.098

Wang, Q., Yu, X., Li, G., Lv, G.: Ontology-based ecological system model of e-learning. Int. J. Inf. Educ. Technol. **2**(6), 595–599 (2012). https://doi.org/10.7763/IJIET.2012.V2.212

Wang, L., Zhao, W., Wei, J., Jiang, Q., Liu, H.: A framework of dynamic visual navigation for subject knowledge. In: International Conference of Educational Innovation through Technology (EITT), pp. 42–45 (2015). https://doi.org/10.1109/EITT.2015.15

Wang, S., Yuan, L., Yang, H.: User model construction of Chinese learners based on learning style. In: International Conference on Computer Science and Educational Informatization (CSEI), pp. 70–75 (2021). https://doi.org/10.1109/CSEI51395.2021.9477714

Wang, H., Wang,L., Du,Y., Chen, L., Zhou, J., Wang,Y., Wong, K.F.: A survey of the evolution of language model-based dialogue systems (2023)

Wang, X., Maeda, Y., Chang, H.H.: Development and techniques in learner model in adaptive e-learning system: a systematic review. Comput. Educ. **225**, 105184 (2025). https://doi.org/10.1016/j.compedu.2024.105184

Wenger, E.: Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge. Morgan Kaufmann Publishers Inc, San Francisco (1987)

Woolf, B.: Introduction to IJAIED Special Issue, FATE in AIED. Int. J. Artif. Intell. Educ. **32**(3), 501–503 (2022). https://doi.org/10.1007/s40593-022-00299-x

Woolf, B., McDonald, D.D.: Building a computer tutor: design issues. Computer **17**(9), 61–73 (1984). https://doi.org/10.1109/MC.1984.1659246

Yaghmaie, M., Bahreininejad, A.: A context-aware adaptive learning system using agents. Expert Syst. Appl. **38**(4), 3280–3286 (2011). https://doi.org/10.1016/j.eswa.2010.08.113

Yago, H., Clemente, J., Rodriguez, D., Fernandez-de Cordoba, P.: On-smmile: ontology network-based student model for multiple learning environments. Data Knowl. Eng. **115**, 48–67 (2018). https://doi.org/10.1016/j.datak.2018.02.002

Yang, F., Li, F.W.B.: Study on student performance estimation, student progress analysis, and student potential prediction based on data mining. Comput. Educ. **123**, 97–108 (2018). https://doi.org/10.1016/j.compedu.2018.04.006

Yang, L., Yu, Y., Wei, Y.: Data-driven artificial intelligence recommendation mechanism in online learning resources. Int. J. Crowd Sci. **6**(3), 150–157 (2022). https://doi.org/10.26599/IJCS.2022.9100020

Yi, B., Lv, Y., Liu,H., Zhang, Z., Zhang, D., Wang, Y.: LSCB-ALS: a new ALS and its application in java language learning. In: International Symposium on Educational Technology (ISET), pp. 191–195 (2017a). https://doi.org/10.1109/ISET.2017.51

Yi, B., Zhang, D., Wang, Y., Liu, H., Zhang,Z., Shu, J., Lv, Y.: Research on personalized learning model under informatization environment. In: 2017 International Symposium on Educational Technology (ISET), pp. 48–52 (2017b). https://doi.org/10.1109/ISET.2017.19

Yi, Z., Ouyang, J., Liu, Y., Liao, T., Xu, Z., Shen, Y.: A survey on recent advances in LLM-based multi-turn dialogue systems (2024)

Yuan, X., Yang, C.: Attention-based user temporal model for recommendation. In: International Conference on Computer and Communications (ICCC), pp. 1877–1880 (2019). https://doi.org/10.1109/ICCC47050.2019.9064186

Yudelson, M.V., Koedinger, K.R., Gordon, G.J.: Individualized bayesian knowledge tracing models. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) Artificial Intelligence in Education, pp. 171–180. Springer, Berlin (2013). https://doi.org/10.1007/978-3-642-39112-5_18

Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–35 (1965). https://doi.org/10.1016/S0019-9958(65)90241-X

Zanellati, A., Mitri, D.D., Gabbrielli, M., Levrini, O.: Hybrid models for knowledge tracing: a systematic literature review. IEEE Trans. Learn. Technol. **17**, 1021–1036 (2024). https://doi.org/10.1109/TLT.2023.3348690

Zapata-Rivera, D. Arslan, B.: Learner modeling interpretability and explainability in intelligent adaptive systems. In: Santoianni, F., Giannini, G., Ciasullo, A. (eds.) Mind, Body, and Digital Brains, Volume 20 of Series Title: Integrated Science, pp. 95–109. Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-58363-6_7

Zapata-Rivera, J.-D., Greer, J.E.: SMODEL Server: Student modelling in distributed multi-agent tutoring systems. Int. Conf. Artif. Intell. Edu. 446–455 (2001)

Zghibi, R., Zghidi, S., Chater, O.: Les normes e-learning comme garant de qualité de l'enseignement à distance dans le contexte éducatif tunisien: le cas de l'uvt. Frantice. net. (2012)

Zhang, X., Cheng, H.N.H.: Mining the patterns of graduate students' self-regulated learning behaviors in a negotiated online academic reading assessment. In: International Conference on E-Society, E-Education and E-Technology, ICSET 2019, New York, NY, USA, pp. 109–114. Association for Computing Machinery (2019).https://doi.org/10.1145/3355966.3355977

Zhang, L., Lin, J., Borchers, C., Sabatini, J., Hollander, J., Cao, M., Hu, X.: Predicting learning performance with large language models: A study in adult literacy. In: Sottilare, R.A., Schwarz, J. (eds.) Adaptive Instructional Systems, Volume 14727 of Series Title: Lecture Notes in Computer Science, pp. 333–353. Springer Nature Switzerland, Cham (2024a). https://doi.org/10.1007/978-3-031-60609-0_24

Zhang, S., Jaldi, C.D., Schroeder, N.L., López, A.A., Gladstone, J.R., Heidig, S.: Pedagogical agent design for k-12 education: a systematic review. Comput. Educ. **223**, 10516 (2024b). https://doi.org/10.1016/j.compedu.2024.105165

Zhao, X.: Research on learner model of adaptive learning system. In: International Conference on Computer Science & Education (ICCSE), pp. 906–909 (2021)

Zhao, J., Chen, L., Guo, J.: . Research and construction of personalized user interest model for distance learning. In: 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), pp. 2324–2327 (2017).https://doi.org/10.1109/IAEAC.2017.8054436

Zhao, Y., Fan, Y., Wang, P.: Research on student model construction of educational big data system based on CD-CAT. In: International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 303–308 (2018). https://doi.org/10.1109/WiMOB.2018.8589101

Zhu, Y., Liu, S., Guo, H., Jin, S.: Design of adaptive electrician learning system based on user model. In: International Conference on Information Technology in Medicine and Education (ITME), pp. 513–516 (2016). https://doi.org/10.1109/ITME.2016.0122

Zhu, Y., Wang, P., Fan, Y., Chen, Y.: Research of learning path recommendation algorithm based on knowledge graph. In: International Conference on Information Engineering, ICIE '17. Association for Computing Machinery, New York (2017). https://doi.org/10.1145/3078564.3078567

Zielinski, A.: A utility-based semantic recommender for technology-enhanced learning. In: International Conference on Advanced Learning Technologies (ICALT), pp. 394–396 (2015). https://doi.org/10.1109/ICALT.2015.120

Authors and Affiliations

Felix Böck[1,2] · Michaela Ochs[1] · Andreas Henrich[1] · Dieter Landes[2] · Jochen L. Leidner[2,3] · Yvonne Sedelmaier[2]

✉ Felix Böck
felix.boeck@uni-bamberg.de ; felix.boeck@hs-coburg.de

✉ Michaela Ochs
michaela.ochs@uni-bamberg.de

Andreas Henrich
andreas.henrich@uni-bamberg.de

Dieter Landes
dieter.landes@hs-coburg.de

Jochen L. Leidner
leidner@acm.org

Yvonne Sedelmaier
yvonne.sedelmaier@hs-coburg.de

1 Media Informatics, University of Bamberg, An der Weberei 5, 96047 Bamberg, Germany

2 Center for Responsible Artificial Intelligence (CRAI), Coburg University of Applied Sciences, Friedrich-Streib-Str. 2, 96450 Coburg, Germany

3 Department of Computer Science, University of Sheffield, Regent Court, 211 Portobello, Sheffield S1 4DP, England, UK