SN Computer Science (2024) 5:559
https://doi.org/10.1007/s42979-024-02837-x

ORIGINAL RESEARCH

Exploring AI-based Computational Models of Novelty to Encourage Curiosity in Student Learning

Maryam Mohseni[1] · **Mary Lou Maher**[2] · **Kazjon Grace**[3] · **Safat Siddiqui**[2] · **Nadia Najjar**[2]

Received: 1 September 2022 / Accepted: 26 March 2024

Abstract

Innovative approaches to personalized and adaptive learning are being developed that leverage advances in AI and access to large datasets. In this paper, we focus on computational models of novelty in large datasets of documents with the goal to encourage curiosity in student learning. While encouraging curiosity is important in all learning experiences, we focus on learning experiences that are open ended and do not have a right or wrong answer, as is common in project based learning and graduate research courses. Through the identification of a generalized framework for AI-based recommendations and the development of the Pique model, we describe the role of computational models of novelty in personalized educational systems, and how computational novelty models can be leveraged to stimulate student curiosity and expand their learning interests. Pique is a web-based application that applies computational models of novelty to encourage curiosity and self-directed learning by presenting a sequence of learning materials that are both novel and personalized to learners' interests, inspiring learners' intrinsic motivation to explore. We describe a generalized framework for computational models of novelty as the basis for the AI component of the Pique learning system and developed two computational models of novelty using Natural Language Processing techniques and concepts from recommender systems. The contributions in this paper include: a generalized framework for integrating computational models of novelty in course recommendations and the Pique model for encouraging curiosity in learners in project-based open ended course experiences. The framework is described to provide structure for the use of computational novelty in Pique and is generalized to inspire this approach in other domains and courses. We report the student experiences with Pique in four courses over two semesters showing that the cumulative interests of the students continued to grow until the end of the semester, and the percentage of students that expanded their interests has a different temporal pattern in a graduate course when compared to an undergraduate course. Based on a qualitative analysis, the students experience with Pique encouraged their curiosity and led them to unexpected topics in their projects.

Keywords Personalized learning · Computational models of novelty · Curiosity · Recommender systems for education

This article is part of thetopical collection "Computer Supported Education" guest edited by James Uhomoibhi andBeno Csapó.

✉ Maryam Mohseni
maryam.mohseni@asu.edu

Mary Lou Maher
m.maher@uncc.edu

Kazjon Grace
kazjon.grace@sydney.edu.au

Safat Siddiqui
ssiddiq6@uncc.edu

Nadia Najjar
Nadia.Najjar@uncc.edu

1 School of Mathematical and Natural Sciences, Arizona State University, Glendale, AZ, USA

2 Department of Software and Information Systems, University of North Carolina at Charlotte, Charlotte, NC, USA

3 School of Architecture, Design and Planning, University of Sydney, Sydney, Australia

Published online: 17 May 2024

Introduction

How to personalize learning at scale has been of significant interest to researchers in education in recent decades. Typically instructors lack the time to personalize learning materials and must progress based on overall learning goals rather than individual student needs. Recommender systems and other information retrieval aids have been studied as a way to help address this challenge. Advances in natural language processing (NLP) on learning materials enable educational environments to personalize the presentation of learning materials to students with content that is responsive to their learning experience, knowledge, and interests. This paper discusses novel approaches to personalized and adaptive learning that use AI models of document novelty to recommend content to students. We contextualize our models of novelty in an interactive system for recommending course relevant publications to University students.

Most research involving recommender systems for personalized learning assumes—explicitly or implicitly—a fixed sequence, tree or other structure of knowledge to be gained by all students. The system then algorithmically assesses each student's progress through those prerequisites. For example, Imran et. al. [1] developed personalized learning in an LMS that responds to the individual profile of achievement of the learner by matching the profile with an object representation of the learning materials. This has been shown to work well in domains with highly structured, hierarchical knowledge, such as mathematics, but is less effective in domains or courses in which the knowledge to be gained is more open ended. In many real-world classroom contexts there are a multitude of possible paths through course knowledge. Algorithmically profiling students in these contexts in order to personalize their learning means not just knowing where they are in a sequence, but what areas of a vast, multi-dimensional space they are familiar with, interested in, challenged by, and bored of. This is a much more complex problem computationally than personalizing learning in well-structured domains.

To address this challenge we leveraged a principle from cognitive psychology: intrinsic motivation. We model, in our AI-based personalized learning system called Pique, what will make students curious, based on the notion that intrinsic motivation follows curiosity about novel and unexpected objects [2]. Pique supports learners by providing recommendations of academic papers based on their interests and computational models of the paper novelty. The goal of Pique is to help students build their knowledge on topics of interest and encourage them to be curious and eager to discover new interests. Curiosity can arise from exposure to appropriately novel stimuli, with insufficiently novel stimuli being boring and overly novel stimuli being overwhelming. An early version of this notion can be found in the writings of 19th-century philosopher Wundt [3], who suggests a region of moderate novelty within the space of possible stimuli, within which what he calls "hedonic value" will be maximized. The parameters of this region are dependent on experiences, context, and personal preference for novelty [4, 5]. The ultimate goal of Pique is to stimulate the curiosity of the student and recommend resources that place them in a state of maximal curiosity by considering their interests and the novelty score of course materials as the basis for recommending the next sequence of course materials.

The distance between what is known and what is desired to be known can be described as an information gap [6]. From the perspective of encouraging curiosity, small amounts of new knowledge prime further desire to learn leading to intellectual hunger [7], but larger amounts have a satiating effect. Vygotsky [8] describes the Zone of Proximal Development as the space of all knowledge that is adjacent to current knowledge and thus is comprehensible to the learner. These theories of knowledge acquisition apply metaphors to describe curiosity and suggest approaches for how curiosity might be operationalized. These approaches are the basis for computational models of curiosity resulting in AI systems that value the unexplored areas [9–12]. The Pique model contributes to this computational modeling of curiosity by setting a goal to stimulate the curiosity of individual students based on a representation of the student's background and preferences to produce personalized recommendations that are both novel and interesting.

Identifying novel and valuable content, designs, and articles can lead to surprising recommendations and consequently stimulate a curiosity to explore beyond what we already know [13, 14]. Surprise is an observer's reaction to novelty, and it has been argued that the same computational models may be applicable to modelling both [15]. Theories of intrinsic motivation consider novelty and surprise as two of the main factors that evoke interest, motivate exploratory behaviors, and consequently drive learning and creativity [16]. Novelty and surprise have been incorporated into several recommender systems in recent years, with similar goals to this research—driving user adoption of new material and thus the broadening of users' preferences [2, 17].

Pique applies AI-based computational models to identify novel documents from a data set of learning resources, then generates a sequence of learning materials personalized to an individual's knowledge and interests. This approach enables instructors to set a class-wide task with a fixed corpus of learning materials, but for each student's experiences to be personalized in open-ended student-led and/or project-based learning [18, 19]. This paper describes the role of computational models of novelty in educational recommender systems to encouraging students' curiosity. To demonstrate how computational models of novelty can be leveraged to

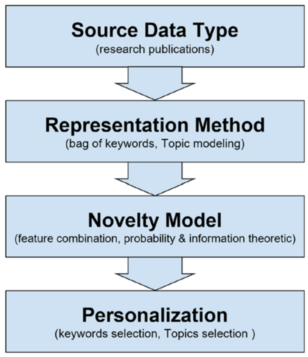

Fig. 1 Framework for novelty-based recommendation indicating the approaches in Pique

encourage curiosity, we describe the Pique system from the perspective of a generalized framework for computational models of novelty. The framework shown in Fig. 1 provides an ontology for computational novelty with respect to its four components: the source of text data, methods for representing the data, models for measuring novelty, and personalization [15]. We elaborate on this framework and discuss Pique from the viewpoint of the framework ("A Framework for Exploring AI-basedComputational Models of Novelty forRecommender Systems" and "The Pique Learning System" sections) as well as on students' experience of using Pique ("The Student Experience Using Pique inSpecific Courses" section).

We have explored a variety of algorithms in the Pique project for representing the source data (learning resources), novelty, and sequence generation of learning resources. We present two data representation methods, two computational models of novelty, and three ways to select learning resources to stimulate students' curiosity. Our novelty models are based on the concept of unexpectedness as a cause of novelty and surprise, which consequently leads to curiosity [13, 20]. For example, learning materials containing interesting and unexpected information can create a surprise response, which may drive students to explore those concepts further. Presenting students with novel learning resources related to but distinct from their knowledge can inspire their curiosity to explore more in the domain.

In this paper, we shed light on the role of computational models of novelty in personalized educational systems such as Pique, and how computational novelty models could be leveraged to stimulate student curiosity and expand their learning interests. We describe a generalized framework for computational models of novelty as the basis for the AI component of the Pique system. Including this framework as a generalization and guidance for personalized learning extends our presentation of Pique [21]. The contributions in this paper include: (1) a generalized framework for integrating computational model of novelty in course recommendations and (2) the Pique model for encouraging curiosity in learners in a project-based open ended course experience. The framework is described to provide structure for the use of computational novelty in Pique and is generalized to inspire this approach in other domains and courses. Pique is presented as a model and an implementation, with an evaluation of this approach based on students' experiences with Pique in 2 semesters of 2 different courses. We describe the architecture of the Pique system and its implementation in personalizing learning materials. The contribution of this paper is a unique approach to personalized learning that encourages curiosity.

The rest of this paper is organized as follows. In "AI-based Approaches to Personalizing LearnerExperiences" section we describe recent research in AI-based approaches to personalized learning. In "A Framework for Exploring AI-basedComputational Models of Novelty forRecommender Systems" section, we describe the framework for computational models of novelty and identify specific computational approaches to each component used in Pique. In "The Pique Learning System" section we describe the Pique model as well as the development and implementation of the Pique system. In "The Student Experience Using Pique inSpecific Courses" section we report on the experiences of university students who used Pique in the classroom, and in "Conclusion and Future Work" section we conclude the paper and discuss directions for extending this research.

AI-based Approaches to Personalizing Learner Experiences

The development and study of Pique contribute to the broad field of AI in education, which has been the subject of significant and diverse work over several decades. Baker and Smith [22] describe three perspectives on educational AI: learner-facing (focusing on assisting students), teacher-facing (focusing on reducing teachers' workloads), and system-facing (focusing on institutions' administrative and management capabilities). Similarly, Zawacki-Richter et al. [23] identify four areas of AI applications in higher education: adaptive systems and personalization, assessment and evaluation, profiling and prediction, and intelligent tutoring systems (ITSs). Pique, while sharing some characteristics with the latter two, is a unique approach to the first category, as it adaptively personalizes learner experiences through a model of curiosity.

The Pique system incorporates elements of educational recommender systems [24], in that it recommends customized course materials, as well as elements of an ITS approach, in that it profiles students to figure out what to teach them next. These elements are underlain by our computational models of novelty, which provide the rankings by which we recommend personalized learning activities. These models create the difference between Pique and existing educational recommender systems: rather than focusing on explicit learning goals or pre-existing sequences of material, we are motivated to increase student curiosity by any means necessary. To this end, we recommend sequences of resources that are predicted to surprise the user and are thus intended to evoke and sustain learner curiosity.

Intelligent Tutoring Systems and Personalized Learning Technologies

Intelligent tutoring systems leverage AI techniques to provide instructional capability according to the needs of individual students. Some models have been proposed to distinguish students' search behaviors for supporting students in effective ways [25]. Ai et al. [26] suggested applying deep knowledge models and reinforcement learning for tracing students' knowledge status and recommending exercises. ITS have been applied to provide personalized feedback in learning programming courses as well as mathematical problem solving [26, 27]. Related to but distinct from ITS are Learning Management Systems (LMS) which track students' learning activities, and in some cases classify students by their behavioral patterns [28, 29]. We do not present Pique as an ITS or an LMS, however, it can be thought of as analogous to one. It models student preferences and represents the learning resources to be recommended with a strategy for composing resource sequences that focus on students' motivation and maximizing their curiosity in learning.

Previous studies applied models of student motivation to enhance the instructional capabilities of ITSs. Del Soldato et al. [30] describe an approach for planning communication with a student performing a series of learning tasks based on a model of the student's motivations. The motivational reasoning in their study is based on Keller's model of motivation [31] comprising curiosity, challenge, confidence, and control, previously applied in computer-supported collaborative learning [32]. The Pique system focuses on curiosity, the first of these motivational factors, which is stimulated by novelty and surprise. We explore the computational models of novelty and recommend resources calculated to increase student familiarity with concepts, and additionally aim to evoke their curiosity.

Educational Recommender Systems (ERS)

ERSs can be applied by students or instructors in either formal or informal educational contexts aiming to recommend learning resources. The goal of the Pique system is being applied in open-ended learning tasks which share many of the features of informal learning while being part of a formal educational context. In open-ended learning tasks, students are asked to choose a focus scope for their work within the proposed problem space [33]. This indicates an aspect shared with informal educational contexts that is students being self-directed to a degree.

Educational recommender systems have been suggested to recommend course material for students in Computer Science [34] and Business and Administration studies [35]. Cobos et al. [36] proposed a recommendation system to prepare course content for the instructors. Educational recommender systems include explainability to justify the recommendation. Barria-Pineda et al. [37] suggested approaches to minimize students' misunderstandings related to solving programming problems, and explained the reasoning behind the recommended learning activities. Barria Pineda and Brusilovsk [37] found that students spend more time on the exploratory interface and suggest the effectiveness of the transparent recommendation process. In Pique, we show students the computed novelty scores of the papers to facilitate students' paper selection process.

Novelty and Surprise in Recommender Systems

Traditional recommender systems have the problem of overspecialization: they suggest to users items that are popular or a close match with prior or current searches [15, 17]. They provide content that is estimated to fit with a user's preferences but do not seek to expand those preferences. There are many domains in which it would be beneficial to provide users with information that goes beyond what a user is familiar with in order to inspire exploration, including education, literature, nutrition [38], and creative practice. Unexpectedness has been a noted cause of novelty and surprise and consequently curiosity [13, 20]. An efficient novelty model can save the user a great deal of time when accessing information by exposing the user only to the most novel and surprising information [15]. From this perspective, modeling and measuring computational novelty and surprise within different types of content plays an essential role in personalized recommender systems and educational systems. Novelty and surprise have been proposed as components of a new kind of recommender system that attempts to expand its users' preferences [2, 17, 39]. Infusing novelty and surprise in recommender systems can prevent fixation by providing a broader set of recommendations to the user. In our approach, we present an innovative model for assigning a novelty score to research papers based on co-occurrence of keywords and topics, and a sequence generator to produce a set of recommendations of research papers to students.

Niu et al. [2] applied two different computational models of surprise in a health news recommendation system to leverage serendipity in recommending health news to users based on their preferences. They define serendipity as happening when something is surprising and also valuable and discuss that their approach presents information that users were not looking for initially but is valuable to their unexpressed requirements [2]. Adamopoulos et al. [17] propose an approach to enhance user satisfaction by generating recommendations that are novel and unexpected for the user and of high quality based on users' interests. They define the concept of unexpectedness as recommending items which are not exactly what users expect from the system. They present methods for identifying the users' expectations and suggest performance metrics to measure the unexpectedness of sets of recommendation items. By considering accuracy simultaneously with unexpectedness and diversity, their approach leads to recommending unexpected items that are also useful to the users [15]. In our model, we adapt this concept to present students with unexpected content in which we model their interests rather than their expectations and we develop computational models of novelty based on topic and keywords co-occurrence in the entire corpus of content. This model of novelty represents expectation in the dataset, and uses student interest to guide the generation of a sequence of novel papers.

A Framework for Exploring AI-based Computational Models of Novelty for Recommender Systems

Advances in Natural Language Processing provide opportunities in recommender systems and education for extracting useful knowledge and measuring the computational novelty in unstructured text documents which are increasingly available as digital content [15]. Approaches to computational novelty and surprise in unstructured text data can benefit scientific innovation, the design of learning materials, and expand the role of recommender systems. Effective and meaningful novelty models can play a key role in identifying content that is both relevant and interesting to users, which is a central goal in recommender and educational learning systems. In this section, we describe a framework for computational models of novelty to have a better understanding of the AI component of the Pique system.

Figure 1 shows the framework for exploring and categorizing computational models of novelty which is an extension of the framework in [15] and extends it for personalized learning. This framework provides a common ontology for exploring and categorizing different existing computational novelty models as well as a guideline for research in developing computational novelty measures for new applications. The framework establishes a sequential process for computational novelty that includes four major components: (1) type of source data, (2) representation methods, and (3) novelty models and (4) personalization. This framework facilitates exploring different approaches to modeling novelty in unstructured text data. Each of these components reflects one major aspect in the analysis of novelty. This framework is independent of the type of data in the items and can be used as a tool for researchers even in other domains to study, compare, and extend existing computational novelty models and applications. In Fig. 1, the approaches used in Pique are shown in each component and described further in this section and "The Pique Learning System" section.

Source Data Type

The first component of the framework is the type of source data: the raw input data which we want to analyze in order to use in our novelty modeling and measurement. The type of input data can affect the ways we can measure novelty in that corpus or domain [15]. For example, when modeling novelty, news articles are different from scientific publications even though they are both represented as unstructured text. The two corpora differ in breadth, depth, intended audience, word use, and purpose. Novelty in news is also time sensitive while a scientific publication written 30 years ago may still be novel. A corpus of news documents contains repetitive content but each scientific publication is likely to be at least minimally different from other publications [15]. Different perspectives are required depending on the type of source data and the kind of novelty being modeled.

In Pique, the source data is scientific publications from relevant conferences, journals, and digital libraries represented as unstructured text. These are provided by the instructor for the specific course in which Pique is being implemented, and constitute the relevant body of learning materials for that course. In the following section, we discuss how data can be represented for use in computational modeling of novelty.

Representation Method

As with any AI task, the choice of representation is critical. This is especially true when measuring novelty because the representation implies what inter-object differences are considered meaningfully new. In particular, unstructured data must first be converted into some kind of structured representation so that novelty models can be built on top of it. In other words, the representation method provides a bridge between the source (raw) data and the novelty model. It provides a processed and structured version of the raw data that can be efficiently used to measure novelty [15]. Building meaningful text representations to be used for novelty calculations can be challenging. Methods for representing text data to be applied for computational novelty models include but are not limited to: bag of words/keywords [40, 41], TF-IDF models [42], word embeddings [43, 44], and topic modeling [45–47].

In Pique, we applied bag-of-keywords and topic modeling to produce representations, as two of the most prominent features of a scientific article are its keywords and main themes. The advantage of using topic modeling compared to author-defined keywords is that there is consistency in the identification of features across the entire data set in topic modeling, where author-defined keywords provide features relevant to the author of a single item in the corpus. On the other hand, author defined keywords are more distinguishable and human interpretable compared to the topics automatically extracted from a large corpus of text documents. Therefore, we decided to apply both methods for representing the source data in Pique in order to benefit the advantages of the two. We will describe applying these representation methods for Pique in more detail in "Data Representation Methods for Pique" section.

Novelty Model

Modeling novelty is the third and most important piece of the framework which is applied in the Pique system for inspiring curiosity and expanding user preferences. There have been different definitions for novelty based on different domains and perspectives. Novelty can be defined as a measure of the difference between an item and a collection of the other items [48]. In some cases, novelty arises from a comparison in a descriptive space such as finding the distance of two points in the space [49]. In the context of recommender systems, novelty means an item that is unknown to the user [15, 50]. A good measure of novelty can help us in educational learning systems like Pique to recommend more diverse and interesting contents to the users to stimulate their curiosity, encourage them to explore more in the domain, and inspire creativity.

One approach for measuring novelty is considering atypical combination of features (i.e. items co-occurrence) [15]. By considering the frequency of co-occurrence of any two (or more) items, novelty can be defined as any rare (new) combination of items that is not similar to the past observed frequent combinations. This also can be explained as observing any combination of items with low probability of co-occurrence which we call atypical combination. Carayol et al. [40] propose a measurement for novelty of scientific articles based on keyword pairwise combination frequencies which is computed on the set of all research articles in the WoS that have at least two keywords during fifteen years from 1999 to 2013. In our study, we use atypical co-occurrence of keywords and topics extracted automatically by topic modeling in modeling novelty of papers, and then we personalize the presentation of novel papers to the interests of our student users.

Probability and information theory have been also applied in modeling the novelty of items. Information theory originally proposed by Shannon [51], examines the properties of information such as quantification, storage, and communication of the information. Some studies use entropy as a metric for measuring novelty by computing the information content of a dataset [52]. Entropy in information theory is a measure of the uncertainty that is associated with a random variable. An entropy function can be applied by researchers to gauge the level of disorder of the remaining dataset after removal of points with high entropy which are considered as novel [15, 53]. It is assumed that novel data contain more information to convey and consequently make the observer surprised [52]. Baldi et al. [52] use relative entropy [54] or Kullback Liebler divergence [55] as one way of measuring the surprise. Niu et al. [2] used two different computational models of surprise for health news articles. One is a variation of Mutual Information (MI) [56] that gauges how much information various random variables share [57]. In their study, the random variables are the topics/labels assigned to an article by health experts. In this approach, each news article is represented as a bag of topics defined by health experts. An infrequent topic combination in an article is considered as a novel combination and gives a higher surprise score for that article. The other surprise calculation method in their study uses KL divergence and LDA Probabilistic Topic Modeling algorithm [45] to discover the themes in the health news articles. Our first novelty model which is based on the keywords co-occurrence of the papers, also benefits from a variation of the Mutual Information in its novelty calculation (see "Computational Models of Novelty for Pique" section).

In Pique, we explore two different approaches for modeling novelty of text data items in a corpus, considering paper keywords and (main) topics as the most prominent features for a scientific article. One approach is based on co-occurring of paper keywords and probabilistic and information theoretic techniques, the other approach is based on co-occurring of paper topics and correlations between topics in the corpus. In the next section, we discuss more about the representation methods and novelty models of the two approaches for computational models of novelty in Pique.

Personalization and Sequence Generation for Recommendation

The last piece of the framework is personalization and sequence generation to prepare a set of recommendations for the user/learner in recommender and learning systems. Using the novelty scores assigned to each document (learning resource) by the novelty model from the third component of the framework, an algorithm will be defined to generate a sequence of item recommendations based on the user topic/keyword selection.

In Pique, we explored three personalization and sequence generator approaches during our project. All of these approaches are based on the novelty score obtained from the third component of the framework as well as student's keyword or topic selection as their interests. Each of these three models named as "Origin–Destination model" , "Destination model" , and "User-Directed model" has its own strategy to produce a sequence of learning materials to recommend to the student. These models and related algorithms are described in more detail in "Personalization and Sequence Generation Algorithmsin Pique" section.

This section showed how the computational novelty framework is the basis for the AI in Pique. In "AI in Pique" section about the Pique AI element, we refer again to the computational novelty framework and describe Pique from the perspective of this framework.

The Pique Learning System

Pique as an educational learning system consists of four main elements of learning materials, artificial intelligence methods (AI), learner model, and user experience (UX). As illustrated in Fig. 2, all of these elements have close interrelation with the components of the computational novelty framework in the AI element. In this section, we describe different parts of the Pique system as depicted in Fig. 2. Additionally, we explore the four components of the computational novelty framework pertaining Pique.

Learning Materials

The instructor provides the source of documents as the learning material for a specific course. The learning material for our deployment of Pique is selected based on its relevance to the courses in which we used the Pique system. We included Pique in two courses in a Computer Science program. The first course titled, "Human Centered Design" , has a focus on human computer interaction. The learning materials for this course are articles published in the ACM Digital Library under the classification of Human Centered Computing. The second course titled "Graduate Teaching Seminar" , has a focus on educational research in computer science, and the relevant learning materials are articles published in the ACM SIGCSE (Special Interest Group on Computer Science Education) proceedings. For the Human Centered Design Course, we collected a total of 9452 conference, journal and magazine papers with publication dates between 2008 and 2018. For each publication we extracted the title, ISSN, location, abstract, publisher, address, ACM ID, journal, URL, volume, issue date, DOI, number, month, year, pages, and tags/keywords as metadata. For the Graduate Teaching Seminar we collected a total of 1172 papers with publication dates between 2008 and 2018, with the following metadata: title, author, conference, year, DOI, keywords, and abstract.

AI in Pique

The AI element in Pique is the most important part, i.e. heart of the Pique system which includes the four components of the computational novelty framework described in "A Framework for Exploring AI-basedComputational Models of Novelty forRecommender Systems" section. In this section we elaborate on the framework components and the approaches/methods we used in each component for Pique.

Source Data in Pique

The source data or the first component of the computational novelty framework in Pique is research publications. As described in "Learning Materials" section, the documents in the datasets for the two courses in which we included Pique are unstructured text extracted from conferences, journals,

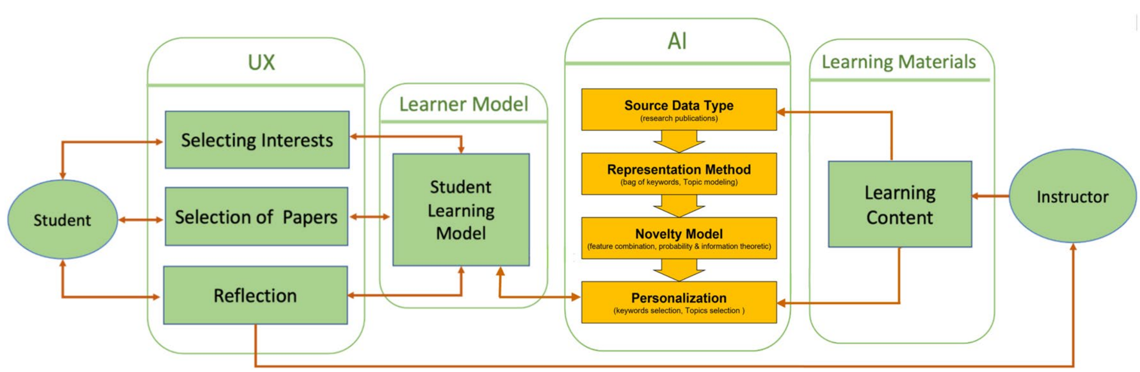

Fig. 2 Architecture of the Pique Learning System

and digital libraries relevant to each course. Following we describe the methods we used in Pique for preparing and representing data to be applied in novelty model.

Data Representation Methods for Pique

The approaches to computational novelty are dependent on the representation of the items for which we are measuring novelty. The representation of unstructured text documents plays an important role in achieving an effective novelty measurement. Two representation methods we applied in Pique to represent the source data include topic modeling and bag of keywords. Considering the paper keywords and (main) topics as the most prominent types of features for a scientific article, each learning item was represented by extracting a list of features based on keywords or topic models associated with each item/document. These features provide the basis for computing a novelty score for each item discussed in the next part.

Applying Bag of Keywords for Representing Data in Pique. For the first representation approach each item of the learning materials is represented as a bag of keywords. With the keywords for each paper, we created a bag of keywords representation for measuring novelty in the next step. Identifying the keywords for the learning materials for each item was challenging in this approach. In the dataset each paper includes two fields in the metadata that can be considered as the keywords for this model. One is the keywords selected from the ACM's Computing Classification System (CCS), and the other is author defined keywords. The ACM Computing Classification System is developed as a polyhierarchical ontology resulting in common topics relevant to all papers, but they do not specifically represent the content in each paper. On the other hand, author defined keywords are defined for each specific paper without following any standard representation. To make the data representation prepared for novelty model, we synthesized the list of keywords from each paper into a master list of keywords for the dataset. We then created a mapping from a user's interests to the concepts in the learning materials by manually curating a reduced set that can be used for mapping. Considering too many keywords would be overwhelming, and inadequate keywords would not represent the dataset with enough fidelity, we tried to choose the number of keywords that are reasonable to present to students for selection. We manually replaced keywords that were not in the reduced list to be the most relevant keyword in the reduced set. Across the semesters, feedback from students indicated that our reduced set of 35–55 keywords was sufficient for students to express their interests.

Applying Topic Modeling for Representing Data. For the second representation approach we adopted a topic modeling approach for deriving concepts from the corpus. A Topic Model [45, 47], also introduced as probabilistic topic models, is a type of statistical model for learning the hidden semantic structures (topics or themes) that occur in a corpus of text documents. A topic model algorithm scans the corpus of text documents, inspects how words and phrases co-occur in them, and learns clusters of words that best characterize those documents. These sets of words often represent a coherent theme or topic. Each extracted topic consists of a probability distribution over all the words in the corpus and each document consists of a probability distribution over the topics [13, 45–47]. Words which present considerably in a topic are assigned a relatively high probability and then documents are assigned different proportions of each topic. By applying topic modeling algorithm to the corpus of learning materials in Pique, each item of the learning materials is then represented as a vector of topic distributions.

For this approach, we used the R package "STM" [58] for representing research publications and then building the second novelty model. STM or Structural Topic Model is an extension of the basic topic modeling algorithm called CTM or Correlated Topic Model algorithm [46]. Correlated topic models relax the assumption by earlier topic modeling algorithms considering all the topics in a corpus as independent and therefore no one pair of topics being more likely to occur together in a document than any other. In order to prepare the data to feed to the STM algorithm, we removed the stop words, numbers, and punctuations, all the words were converted to lowercase and then stemming was performed. After the preprocessing steps, we run the STM topic model algorithm on our dataset to extract the main themes (topics) of the corpus of published articles and obtain a vector of topic proportions for each paper. STM runs on a corpus of text documents. The three main outputs we obtained from the topic model by which we later build our novelty model are: (1) topics derived by STM algorithm, (2) document vectors including proportion of topics for each document, and (3) topics correlation matrix. Each document in the corpus is represented with a 20 dimensional vector containing the distribution of topics in that document. The correlation matrix is a 20×20 dimensional matrix including the correlation coefficient for all topic pairs.

Computational Models of Novelty for Pique

In the process of developing Pique, we implemented two computational models of novelty, based on probability and information theory, and features combination. One novelty model is referred to as 'Keyword co-occurrence model' and the other as 'Topic co-occurrence model'. Each item in the learning materials is represented as a bag of keywords in the keyword co-occurrence model, while the topic co-occurrence model applies topic modeling approach/method to represent each item in the learning materials as a vector of topic distributions. Following we describe how these computational approaches assign novelty scores to the documents in the corpus of learning materials specified by the instructor of the course.

Novelty Model Based on Keyword Co-occurrence. The first novelty model is based on the probability of co-occurring for each pair of keywords in the corpus. This model benefits from a variation of the *Mutual Information* for calculating novelty as in [2, 56]. Having a bag-of-keywords representation for each paper, we calculated the cooccurrence of keywords for measuring novelty. We removed papers with fewer than two keywords, and then measured the probability of each pair of keywords appearing together in the corpus. We applied the resulting probabilities shown in Eqs. (1) and (2) below to get the probability of co-occurring of keywords $x1$ and $x2$ in the corpus as shown in Eq. (3). By taking its logarithm we got the novelty score for that pair of keywords. A novelty matrix *NM* (Eq. 4) was then created to all pairs of keywords in the corpus, considered as the look-up table for identifying the novelty scores among the keyword pairs in the papers. The highest value of all keyword pairs present in a paper was then used to get the score for the paper as surprising combinations stand out [13] which is shown in Eq. (5).

$$prob(x_1) = \frac{\#\ of\ papers\ have\ x_1}{\#\ of\ total\ papers} \tag{1}$$

$$prob(x_2) = \frac{\#\ of\ papers\ have\ x_2}{\#\ of\ total\ papers} \tag{2}$$

$$prob(x_1, x_2) = \frac{\#\ of\ papers\ have\ both\ x_1\ and\ x_2}{\#\ of\ total\ papers} \tag{3}$$

$$NM(x_1, x_2) = \frac{log_2(prob(x_1, x_2))}{prob(x_1) * prob(x_2)} \tag{4}$$

$$NoveltyScore\_P_n = max(NM(x_1, x_2), NM(x_1, x_3), ...) \tag{5}$$

Novelty Model Based on Topic Co-occurrence. The second novelty model assesses the atypicality (unexpectedness) of the topic combinations that appear in abstracts of the published papers to build the novelty model of papers. By applying the topic modeling approach, each paper is represented as a vector of topic proportions. Novelty score for each document is computed based on proportion of topics present in the corpus and frequency of topic co-occurrence. We consider the overall novelty of a document to be equal to the most novel concept or combination of concepts within that material [13]. This model bases the novelty of a text document as the lowest (i.e. highest negative) correlation coefficient among all pairs of topics significantly present in that document, and the proportion of the document which contains that pair. To determine whether a topic is "significantly present" in a document, a topic proportion threshold of 0.1 is used, that is the document should be at least 10% comprised of that topic. This is the basis of our novelty model stating topics are concepts derived from the dataset and the co-occurrence between topics gives us a basis for what combinations of concepts are novel (unexpected). This novelty model considers topics as concepts derived from the dataset and the co-occurrence between topics giving a basis for what combinations of concepts are novel (unexpected). This model is based on previous work in topic-model approaches to novelty [13]. Equation (6) shows the novelty formula for a paper p given $p = [t_i, t_j, \ldots, t_n]$ consisting of the set of topics significantly present in p. The pair of topics in p with the lowest correlation coefficient are denoted by t_i and t_j which are considered as the most novel topic combination in p. This coefficient is normalized against the most novel pair of topics in the whole corpus that are t_a and t_f and then weighted by the proportions of t_i and t_j in p for computing the novelty score.

$$NoveltyScore\_P_n = \frac{CovMat(t_i, t_j)}{CovMat(t_a, t_f)} \times 2(min(prop(d, t_i), prop(d, t_j))) \tag{6}$$

CovMat is the covariance matrix obtained from the topic model, $CovMat(t_i, t_j)$ is the correlation of the document's most atypical topic combination (t_i, t_j), and CovMat(t_a, t_f) is the correlation of the most atypical topic combination in the whole corpus (i.e. t_a and t_f, which are the pair with least correlation in *CovMat*). $prop(d, t)$ is the proportion of document d that consists of topic t. The expression in the parentheses is the novelty of the document's most novel topic combination, represented as a proportion of the most novel topic combination in the model. The second expression of the formula is indeed twice the smaller of the two proportions of the document d that are from the novel topics t_i and t_j. The product of the two expressions provides the normalized novelty (unexpectedness) of the most novel (unexpected) topic combination weighted by how much of the document consists of that combination. The reason for using the minimum of the two topic proportions of t_i and t_j rather than the sum of them is to prevent favoring documents that just passed the significance threshold with one topic, and were thus not particularly novel [13]. Therefore, the minimum of the two topic proportions is used in the formula to weight the novelty measure towards documents containing significant amounts of both unexpected topics. The formula assigns a novelty score of 1 to a document that is made up of 50% of each t_i and t_j. It assigns a novelty score of 0 to the document that contains at most independent topics, and assigns a negative novelty score to documents consisting of only topics with positive correlation. In this novelty model, novelty rating for documents containing a lot of relatively novel topic combination will be higher than for documents containing only a little of the most novel pair of topics [13]. In the next section we describe the process and algorithms we developed in Pique for finding appropriate papers to recommend to students which are both novel and match their interest.

Personalization and Sequence Generation Algorithms in Pique

Pique personalizes its recommendations by including student's selection of keywords/topics of their interest in the process of generating the sequence of papers for recommendation. In generating the recommendation sequence, Pique takes the novelty ratings of each document in the corpus and constructs a sequence of learning resources that maximize the chance of a student experiencing optimal novelty. The goal is generating a sequence of learning resources to support student-directed learning and to stimulate students' curiosity about learning. We explored three sequence generator approaches during the course of our project. We named these three models as "Origin–Destination model" , "Destination model" , and "User-Directed model" .

Pique generates a personalized sequence of nine documents in sets of three papers from the corpus of learning resources based on student information and preferences. Students select one paper from each set of three, read it, and reflect on it. Then students are presented with the next set. The different sequence generator models are based on different representations of student interests. The Destination Model uses one set of student-specified interests as the input to the algorithm. In the Origin–Destination Model two student-specified sets of keywords are used: one that they self-report as already knowing about or the "origin" set, and one that they want to learn more about or the "destination" set. The User-Directed Model extends the Origin–Destination Model to include other keywords from the papers most recently selected by the student. The sequence generator uses these keywords to represent student preference, and combines that with the novelty score for each paper to select and sequence learning resources with the goal of inspiring curiosity.

Destination Model. The Destination Model asks for what students desire to learn and recommends a set of nine novel documents relevant to their stated desires. When applying with our keyword co-occurrence novelty model, the student interests are directly mapped to the corpus keywords, but in the case of applying the topic co-occurrence model, a mapping was manually built between the topics automatically generated by topic model and the keyword set we had created. Here we refer to 'novel documents' generally, without specifying which novelty model labeled them as such. Initially students select their learning interests, which is considered as the destination set, D. Then the destination model identifies documents in the learning materials corpus for which the top N topics within that document include at least one of the user's selections. We decided on $N = 3$, as we found most documents in the corpus included at least this many topics at reasonable proportions. From the set of identified documents the nine most novel papers are selected and sorted ascending based on their novelty score, from the moderate novel to the most novel one. The goal in the Destination Model is recommending nine documents with information that students want to learn, starting with a moderate novel document and then scaling up to highly novel documents as the student reads more and learns about their interested topics.

Origin–Destination Model. The goal of Origin–Destination model is inspiring students to explore learning materials that contain some information that they already know, combined with some new information that they don't. The idea of this model arises from educational psychology that discuss new material is only learnable if it is at least somewhat connected to topics already known (Vygotsky's Zone of Proximal Development [8]). This model generates a recommendation sequence that moves from what the student already knows to what they want to know. The algorithm is inspired by the surprise walks algorithm [20] that moves from an unsurprising source to a surprising destination in the recommendation sequence.

In the Origin–Destination Model the learning materials are presented in three steps of "close" , "far" , and "farther" to stimulate learners' curiosity. Recommending the learning materials step by step, helps students to gradually learn new materials similar to what they already know and inspires them to explore without recommending materials that are so novel as to be unfamiliar and overwhelming for them [59]. In the first step, papers similar to student's familiarity, which are labeled as the "close" category of learning materials for that student, are recommended by the model. In the second step, papers that are similar to both what the students already know (their familiarity) and what they want to learn, labeled as the "far" category, are recommended by the model. In the third step, papers containing materials related only to what students want to learn, labeled as the "farther" category, are recommended to student by the model.

For the "close" category, the model identifies candidate papers containing at least one common keyword (or topic) from the students' initial interest set (source set). By applying the k-means algorithm the model clusters the candidate papers based on their novelty scores to distinguish the papers with three novelty levels of high, medium, and low. The model computes the paper's familiarity score as well, denoting the number of keywords in common between the paper and the "origin" set of keywords/topics the student already knows. Then in each novelty level the papers with highest familiarity scores are selected, and finally the algorithm recommends one low, one medium, and one high novelty paper. Regarding the "far" category, the model recommends another three papers for expanding students' learning from what they are familiar with to the new topics they would like to learn. Candidate papers in this category include at least one common keyword from the origin keywords set (O) and at least one common keyword from the destination keywords set (D). The same clustering approach is applied to identify low, medium, and high novelty candidate papers, and the candidate papers are identified in each level with the highest number of common keywords. For the "farther" category, papers containing information that students desire to learn are presented by the model. Candidate papers in this category include at least one keyword from the destination keywords set (D), and similar to the other two sets the candidate papers are categorized into three levels of novelty.

User-Directed Model. User-Directed model extends the Origin–Destination Model by considering students' decisions during the recommendation process in order to recommend materials aligned with their evolving interests. As in the Origin–Destination model the papers are recommended step by step by the three categories of close, far, and farther, but this model additionally keeps track of students' selections of papers from the previous step. Keywords of papers in the previous step are applied in order to prioritize similar resources in the recommendations of the next step. The User-Directed model filters the candidate papers for the far step to those that share at least one keyword with papers selected in the close step. The model first identifies candidate papers for the farther step that contain at least one keyword in common with the keywords of the paper selected in the far step. This model is identical to the Origin–Destination model except the aforementioned filtering step. That is, it recommends one low, one medium, and one high novelty paper in each of the close, far, and farther steps.

Learner Model

The Learner Model represents information about the learner to support the selection and presentation of learning materials. The learner model also includes information from the learner during their interaction with Pique that is used to analyze the user experience. The Learner model is not a comprehensive model of the learner. The model includes two kinds of information: information provided by the user about their interests and reflections and information about their use of Pique. Identifying information about the user includes their name, participant ID, email address, and course. The IDs are automatically generated by the Pique system and are linked to the identifying information, then the linking data is stored separately from the data collected about the student's interests and use of Pique, as required by our IRB approval. The student's interests are self-defined and change during the semester. Students select their interests at the beginning of the semester and are prompted to update their interests in each recommendation cycle. Each time the student uses Pique is a new cycle of recommendations. For each cycle Pique records timestamps, the interests of the student, the papers selected by the student, the options the student chose, and the student's reflections. The reflections include the student's responses to three questions: (1) why did you select the paper? (2) is the selected paper on a topic matching your interests? (3) what topics do you expect to learn from the paper?

UX for Pique

The User Experience of Pique supports students' interaction with the following three steps: Selecting interests, Selection of papers, and Reflection. We designed the student experience of using Pique as a cycle of recommendation followed by reflection. The students log in so that Pique can track their selection of topics/keywords, their selection of papers,

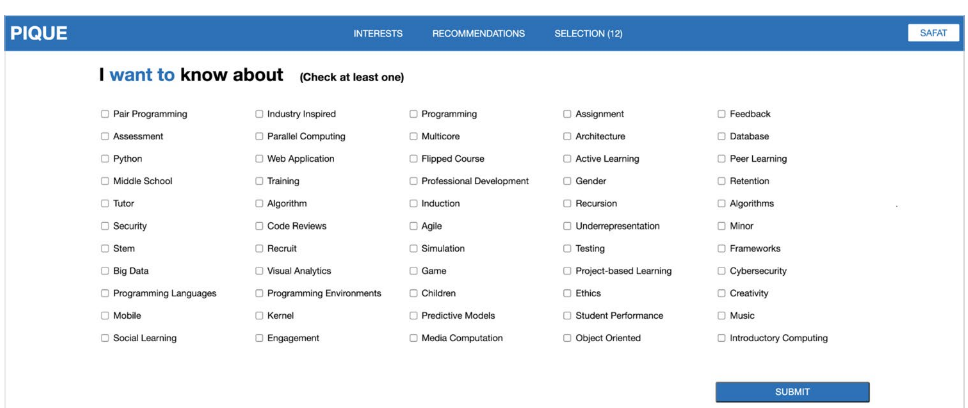

Fig. 3 UX for selecting interests in Pique [21]

and their reflection on the papers they read over the course of the semester.

Selecting Interests. Pique captures students' interest by prompting them to identify what they want to know. This prompt assists students to formulate their learning goals and provides them more control over their learning choices and enables self-directed learning. After students log in to the system, they are prompted to select the topics/keywords they would like to know about from a checkbox interface. Figure 3 shows the user interface with the learning options for the students as they were in the Graduate Teaching Seminar course.

Recommendation and Paper Selection. After students submit their selected topics/keywords, they are navigated to the recommendation page (Fig. 4) showing a sequence of nine papers in the area of their selected topics/keywords. For each paper, the students are presented with the title and novelty score of the paper and papers are sorted by their novelty score. Students can view and download the pdf file of the paper by selecting/clicking it. This step of paper selection in Pique enables students' self-regulated learning, with the intention of stimulating their intrinsic motivation to learn and explore. This stage presents the papers that are recommended by the student selection and sequence generation of Pique (the fourth component of the framework in "A Framework for Exploring AI-basedComputational Models of Novelty forRecommender Systems" section). Pique presents the nine papers in sets of three, based on the sequence generation algorithm (see "UX for Pique" section). Figure 4 shows an example of papers being recommended in the Graduate Teaching Seminar course based on the Orign-Destination sequence composition model. The top three papers are closely related to what the student already knows, the middle three are related to both what they know and what they are interested in, and the bottom three are related to what they are interested in only. The step of paper selection informs students about how novel a particular paper is, and allows them to manually choose more or less novel papers by selecting the drop-down menu labeled 'show me papers' in the top right corner of Fig. 4. In this way students have the option to explore a wider range of papers in their selected topic/interest category.

Reflection. The third step of the Pique UX is Reflection. It has been shown in cognitive studies of students that reflection is key to effective learning [60–62]. There are two types of reflection in the Pique system. One is requested when students select a paper to read as shown in Fig. 5, and one is requested at the end of the semester. The first reflection asks the student to answer 3 questions about the paper they selected (Fig. 5). The first question asks about why they selected this paper. The second question asks whether the selected paper is on a topic the student

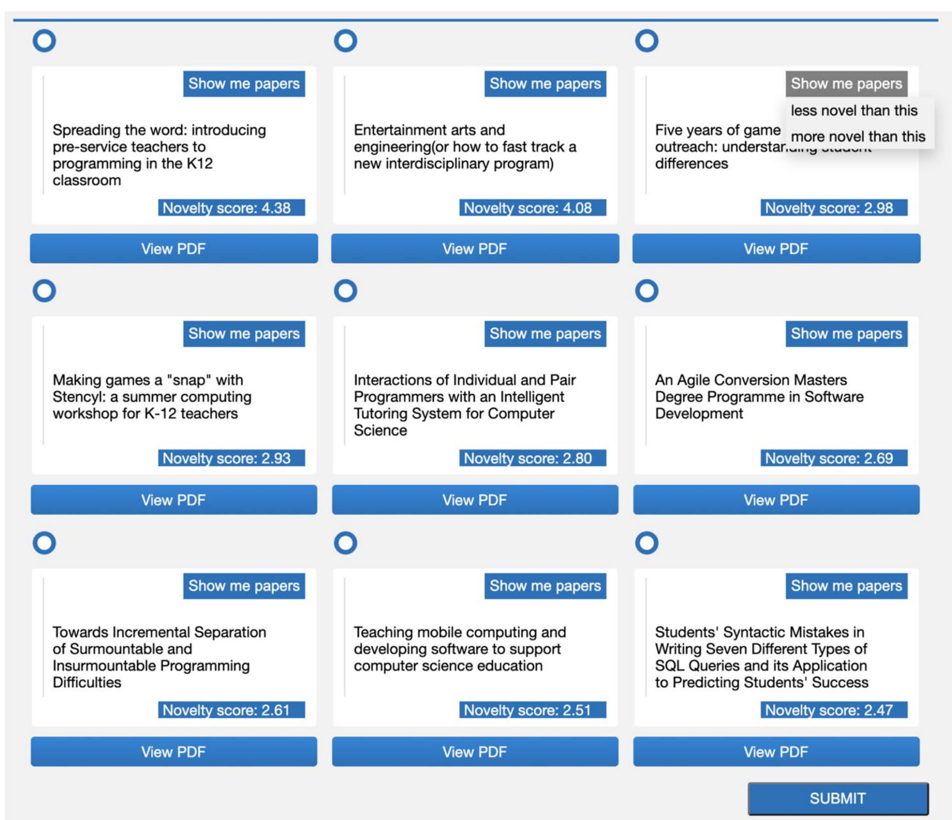

Fig. 4 The UX for recommendation sequence, and selecting learning content based on interests and novelty scores [21]

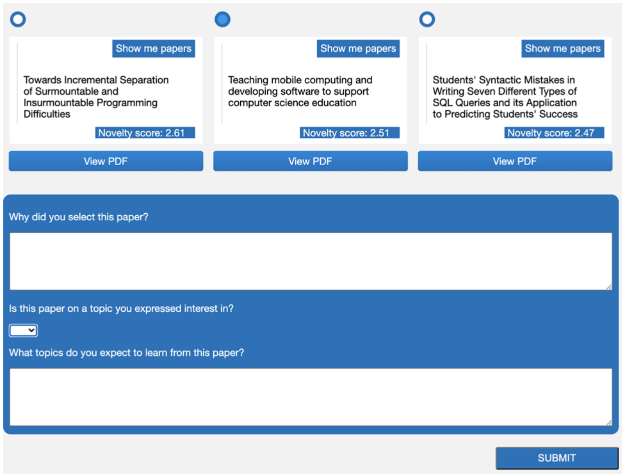

Fig. 5 Pique asks students to reflect on their paper selection and learning expectations [21]

expressed interest in, and the third question asks about what the student expects to learn from this paper. After completing the survey, the student can log out or continue to the next round of the recommendation cycle. The second type of reflection asks students to reflect on their overall learning experience. Students were asked to summarize the papers they read and categorize those papers into groups. Students are asked to identify the paper they found most interesting and justify why. This reflection allows students to organize their newly acquired knowledge where the learning paths are constructed by the students rather than the instructors. It was also critical for evaluating the impact of this educational innovation on the student experience.

The Student Experience Using Pique in Specific Courses

We applied Pique in an undergraduate human-centered design (HCD) capstone course and a graduate teaching seminar course for PhD students. These 2 courses are project-based, where the HCD course requires research relevant to a design project and the graduate teaching seminar has a focus on reviewing research for a project report on graduate teaching. We used Pique over several semesters and continually developed the models of novelty and sequence generation based on student and instructor feedback. Our goal was to address the following research questions based on students' experience with Pique and their reflections on the recommended learning content:

- RQ1: How does the experience of using Pique enable self-directed exploration and personalized learning?
- RQ2: How does the experience of using Pique assist students in expanding their learning interests?

In this section, we describe the deployment of Pique and the experiences of students who have used Pique in the classroom through a quantitative and qualitative analysis of student data collected during the course experience. We have IRB approval for the data collected by Pique.

Quantitative Analysis of Students Experience

We used Pique over four semesters in both undergraduate and graduate courses in Human Centered Design as well as the Graduate Teaching Seminar PhD course. In the Human Centered Design course students were asked to use Pique for six weeks, and had to submit weekly and end of semester writing assignments about the papers they had read. Each week they were asked to submit a summary of the three papers they downloaded and read, and identify the most interesting paper among the three. For the end of semester report, the students were asked to explain their experience of using Pique, what they learned, the most interesting paper they found among all and their reason for why they found the paper the most interesting. For the Graduate Teaching Seminar course, students were asked to use the Pique system for the entire semester, but submitted only a final report without any weekly submissions. This was due to the PhD students' greater familiarity with reading published articles, as well as their overall greater autonomy as learners.

Regarding our first research question concerning how the use of Pique assisted enabling self-directed exploration, we investigated how the student cohort differed in the resources they explored, as a measure of how self-directed their experiences were. Table 1 shows the summarization of our results. Though students' options for selecting interests were unchanged, that is 39 interests in Human Centered Design and 55 in Teaching Seminar course, we found that students were presented with very diverse sequences of learning resources. A total of 621 unique papers were recommended by Pique in the Graduate Teaching Seminar course for one semester, even though this course included just five students. The results showed 55% of those papers were recommended to at least two students, due to overlaps in topics of interest. Those five students selected a total of 66 papers to read, showing 86% of the selected papers were selected by just one student. Across all four courses we observed 72% of recommended papers were recommended to at least another student, but the students' selections were highly diverse, showing 70% of the selected papers were unique to that individual student.

Table 1 Distribution of learning materials to personalize learning

Course name	Graduate teaching seminar	Graduate teaching seminar	Human centered design	Human centered design
Semester	Spring 2020	Fall 2020	Spring 2020	Fall 2020
Number of students	24	5	12	12
Number of unique learning sequences generated by students	24	5	12	12
Total papers selected by students over the Pique cycles	221	66	76	77
% of selected papers uniquely picked by individuals	50% (111 papers)	86% (57 papers)	71% (54 papers)	74% (57 papers)
Total papers recommended by Pique	1987	612	729	774
% of papers recommended to at least one other	84% (1669 papers)	54% (333 papers)	75% (548 papers)	72% (558 papers)
% of papers recommended to only one student	16% (318 papers)	46% (279 papers)	25% (181 papers)	28% (216 papers)

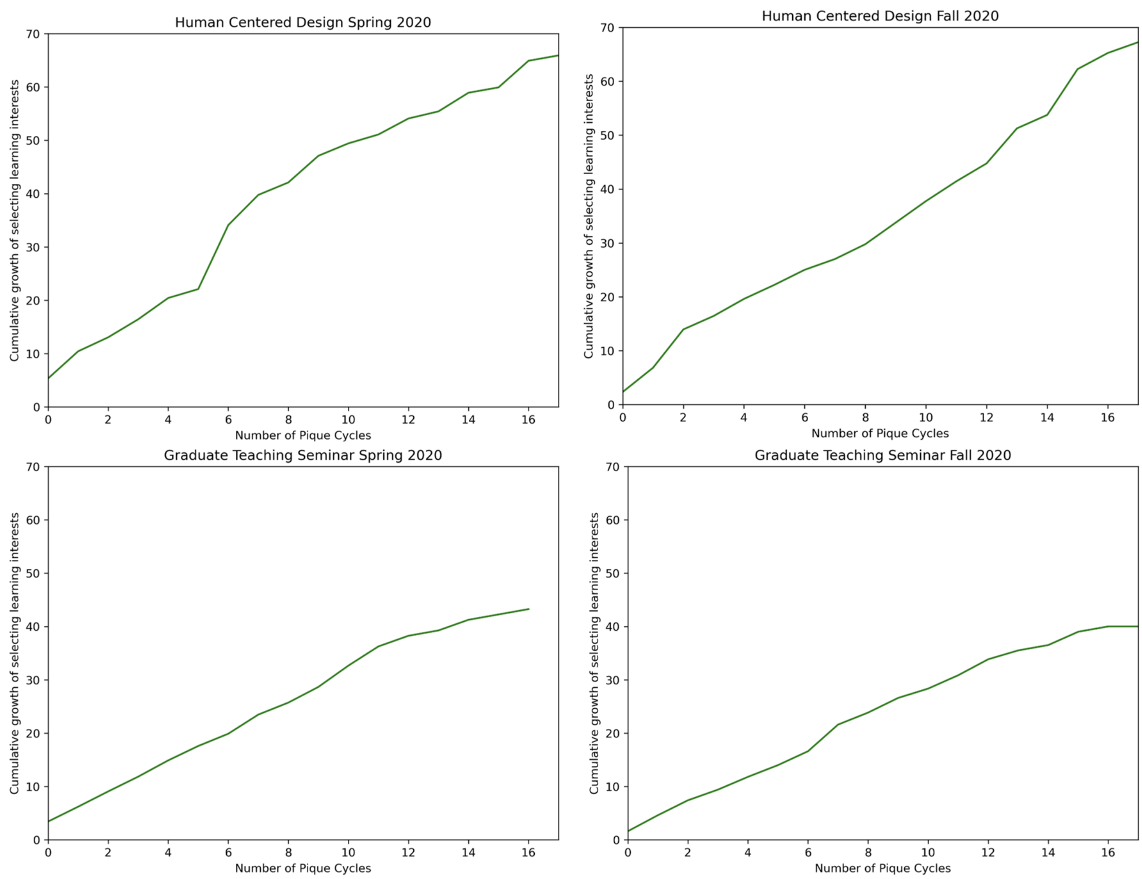

Fig. 6 The increase in the selection of learning interests while using Pique in 2 semesters of the HCD course and 2 semesters of the Graduate Teaching Seminar [21]

Our second research question asked how using Pique helped students in expanding their learning interests. We investigated the change in students' interests over time for responding to this question as illustrated in Fig. 6. The top two charts in Fig. 6 are related to the HCD courses in Spring and Fall semesters. The bottom two charts are related to the Graduate Teaching Seminar courses in the Spring and Fall semesters. In all semesters in which Pique was used, we observed an increase in the interests selected by the students. The X-axis shows the number of Pique cycles and

Y-axis shows the average cumulative growth of the interest selections. We computed number of interests selected by each student for each cycle. We aggregated this for all students within a cohort to give the average number of interests selected by the students in that cycle. The cumulative number of interests in Fig. 6 demonstrates the expansion of stated interests over the semester. The total students in the HCD courses selected an average of only four interests at the beginning of applying Pique. As students used the Pique system over the semester, we observed that searching of learning interests increased as well. At the end of the semester all students had explored an average of 67 interests. Regarding the Graduate Teaching Seminar, total students started with just two interests on average, and over the semester the average number of their searching learning interests raised to 42.

As demonstrated in Fig. 7, we found a difference between the students in the HCD courses and those in the Graduate Teaching Seminar. The top two charts are related to HCD courses in Spring and Fall semesters. The bottom two charts are related to Graduate Teaching Seminar courses in the Spring and Fall semesters. The X-axis shows the number of Pique cycles and Y-axis shows the percentage of students searched for new interests that they had not selected in earlier cycles of using Pique. The students in the HCD courses were undergraduate and graduate students who initially expanded their learning interests and over the time they reduced the number of new interests. The students in the Graduate Teaching Seminar courses were conversely PhD students who kept exploring new interests. For example, we observed that all the PhD students in the Fall semester of Graduate Teaching Seminar continued to add new interests until the end of the semester. We observed that 71% of PhD students in the Graduate Teaching Seminar for the Spring semester had new interest in their 8th cycle of using Pique, but just 18% of the undergraduate students in the HCI course continued exploring in the 8th cycle. This result suggests that students apply the Pique system differently for expanding their learning selections.

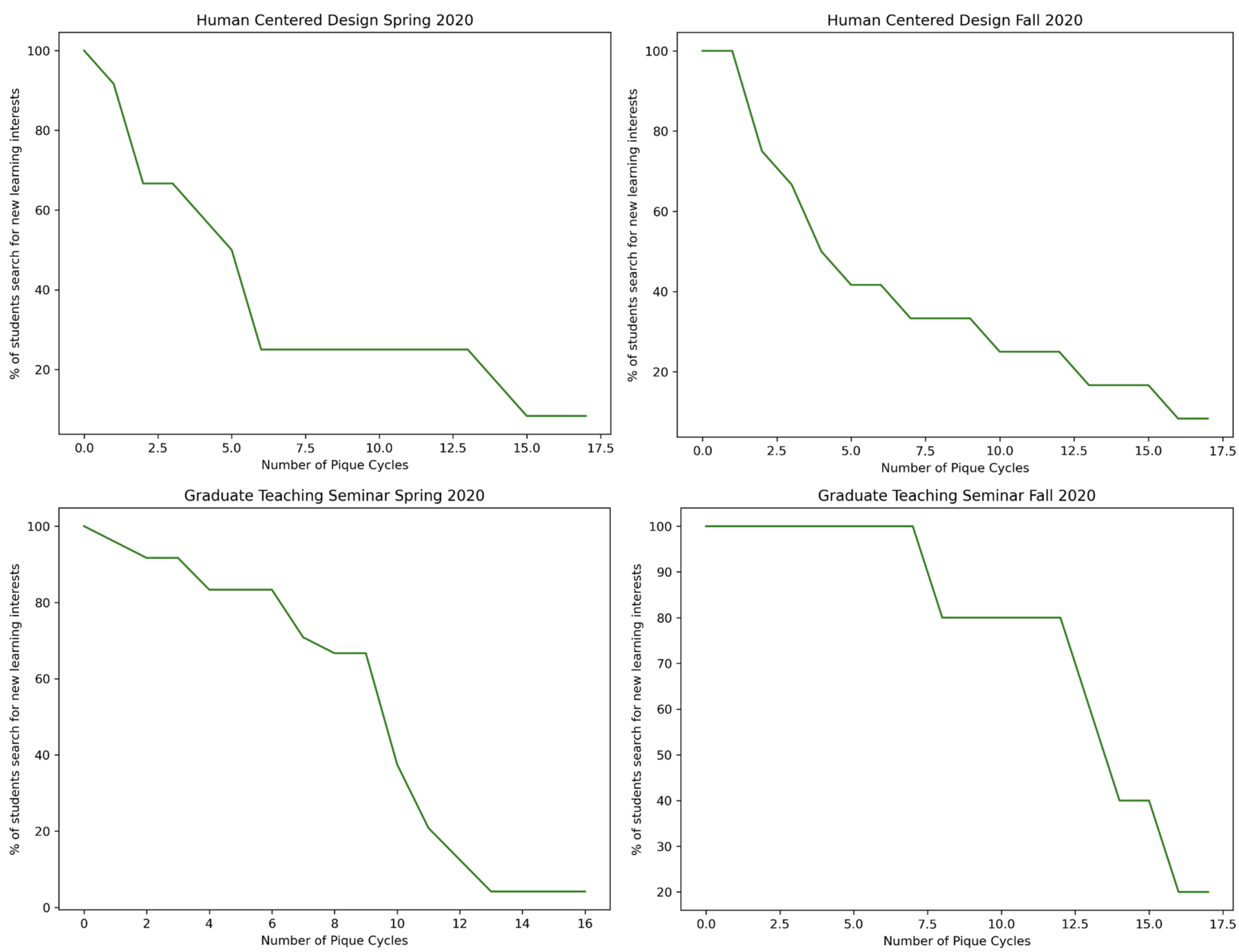

Fig. 7 Percentage of students searching for new learning interests for each cycle while using Pique in 2 semesters of the HCD course and 2 semesters of the Graduate Teaching Seminar [21]

Qualitative Analysis of Students Experience

At the end of applying Pique by students, they were asked to reflect on which paper from the system they found most interesting and their reason for that. In order to discover meaningful patterns in the data, two researchers from our team performed a thematic analysis on students' written responses [32]. Applying multiple coders provided investigator triangulation to our analysis [33]. For having a broad consensus, initially the two researchers conducted a parallel coding workshop on the first 10% of the written responses. After discovering the initial set of themes, each of them coded the rest of the data separately and then converged on a set of collaboratively authored themes through follow-up workshops.

We discovered three major themes of novelty, personal relevance, and curiosity, underlying why most students found papers interesting. The first theme captured how students found papers interesting because of the innovation and novelty of the idea proposed in the paper. Finding novelty as one of the main themes in students' reports as their reason for why they found a paper most interesting and surprising shows that recommended sequence of personalized and novel papers catches the students' interest and makes them surprised. The second theme captured how students found papers interesting specifically when they found its contributions and implications related to their personal life and experience. For instance, one student found a paper about a VR gaming application called "Spider Hero" interesting because he was a fan of Spiderman. Another student found a research idea of another paper so interesting because it presented new approaches for assisting disability and he had a disabled sister. Students also liked the recommended papers because they found those were aligned with their personal belief. For example, another student liked a paper that discussed young parents shared information about their child online because he believed it is exactly what is happening in our society. The second theme indicates that Pique recommended personalized papers that students found interesting to read. This is consistent with the system's goal of personalization. The third theme captured how the recommended papers make students curious about the research field of HCI and computing education, and assisted in growing their interest in the field. Students get the opportunity to know about the broad research area of the field. For example, one student stated he learned something new from each of the recommended papers. He became so curious that he did extensive personal research to learn more about specific topics. We found that the curiosity theme was related to the idea of students connecting their class lessons with the recommended papers. For instance, a student learned the concept of "Wizard-of-Oz" in the HCI class sessions, and later finding the same concept in a research paper excited him a lot. The students' written responses and discovered themes indicate that the recommended papers motivated students to explore and learn more in the domain. The three main themes we identified through thematic analysis (novelty, personal relevance, and curiosity), are all consistent with the goal of Pique to recommend novel and personalized papers and consequently encourage students' curiosity to explore more and expand their knowledge.

Limitations

The exploratory study of the use of Pique in a classroom setting has methodological limitations compared to a controlled study. First, we studied the use of Pique in different courses which may have caused different levels of interest in the course materials. Second, the number of participants was too small for us to perform a quantitative analysis with significance testing. We relied mostly on qualitative analysis and trends in the quantitative results. Third, we explored different computational models of novelty in different semesters as part of our efforts to develop multiple approaches to measuring novelty in test-based documents. Despite these limitations, the use of novelty as a basis for recommending learning materials to encourage curiosity in open-ended learning tasks was successful in expanding the interests of the students. These limitations identify the potential for future studies to refine the approach for recommendations in Pique.

Conclusion and Future Work

Pique is a cognitively-inspired system architecture that presents materials to learners that are personalized to encourage curiosity. We present Pique as an educational recommendation system that uses AI techniques, including NLP, to present students with personalized sequences of novel learning resources. We show these sequences encourage curiosity and support self-directed learning. Pique applies computational models of novelty for identifying documents from a corpus of learning materials that are both relevant to the student's interest and novel with respect to the corpus. Rather than steering students through a specified curriculum, Pique aims to inspire individuals' curiosity to learn by selecting their own interests. Pique encourages students to expand their knowledge and trigger new ideas for their course projects and/or research projects by reading newly recommended learning materials.

Computational models of novelty can play a key role as a major component of the AI element in educational recommender systems for engaging learners and evoking their curiosity to explore more in the learning process. Applying an efficient novelty model in educational and recommender systems can benefit the user when accessing information by presenting the user with the most novel and surprising information among the increasingly large repositories of documents and learning materials [15]. We present a framework for computational models of novelty to describe the inner processes of the AI module in our Pique system. The framework, which is an extension of the framework presented in [15], consists of four components including source data, representation method, novelty model, and personalization. This framework provides a structure for exploring and categorizing different approaches to novelty detection from the perspective of each of the 4 components in the framework, and is a basis for leveraging this technology in educational recommender systems. The proposed approach in this paper has the potential to be used in some other research fields about AI, such as image segmentation [63], image analysis with ANNs [64], medical image analysis [65], feature extraction [66], and video analysis [67].

We developed and deployed Pique during a four semester exploration of how to inspire students' curiosity. We chose novelty as a measure for content that encourages students' curiosity to explore more in the domain of study, and developed two AI-based computational models of novelty: one based on keyword co-occurrence and another based on the co-occurrence of topics from a topic modeling algorithm. These computational novelty models are based on the same underlying information theoretic approach to novelty or surprise as features that negatively correlate but differ in the way we generated a keyword or topic representation of the documents. Both models have their own pros and cons, each capturing one aspect of novelty and capable of identifying some surprising-seeming papers that the other missed. In our computational novelty models, we used the keywords and topics of the papers as two of the most prominent features for a scientific paper. For the first model, we used bag of keywords as the data representation method to be applied for modeling novelty. In the second model, we used topic modeling as the data representation method by extracting the main topics of the papers in the corpus, and represented each document as a vector of topic proportions. In topic modeling approach there is consistency in the identification of features across the entire dataset, while author-defined keywords provide features relevant to the author of a single article in the corpus. In author defined approach, the authors have no idea what is in the other papers, but in topic modeling approach, the topics are based on what is in the other papers. On the other hand, author-defined keywords provide more distinguished features whereas we observed some topics extracted from topic modeling have some overlap. Thus, these are different concepts for building models of novelty. In our future study, we plan to study these models separately. In the future, we also plan to explore other approaches to representing the corpus of learning resources, including NLP and machine learning techniques, and extend our current computational models of learning resource novelty by applying these new representations.

We also developed three models for personalization and recommendation during the Pique project. The first model was just based on the student's stated interests (student's destination). The second model was based on directing them from what they already knew (student's origin) to their interests (student's destination again). The third model was based on a mixture of the origin–destination effect with similarity to the things they've recently explored. Each of these three models combined student preferences with our developed computational models of novelty to encourage curiosity in the learning process. We did not compare directly the personalization and sequence recommendation models, however, we believe, from the evidence of using them in the classroom, that both of the latter two models offer advantages over the former one.

This paper presents a proof of concept from the deployment of Pique, as a personalized curiosity engine and sequence generator in a recommender system for education. We have identified a number of areas for future research, as well as provide evidence of the well-known complexity and nuance of applying intelligent systems in education. One limitation for our study was that we did not gather data for the time the students spent on reading the papers before reflecting on them. We evaluated the experiences of students who used Pique as part of their courses, and found three aspects that made recommended learning materials interesting: how novel they were, how personally relevant they were, and the curiosity and further self-directed learning that they evoked. Our findings are evidence of how curiosity can be elicited from students as part of a course experience, when self-directed and open-ended engagement with learning resources is desirable. Our results from reflection surveys and written reports indicate that students were interested in the personalized papers recommended. We observed that students are eager to engage with educational recommender systems like Pique, and that their interests diversified as a result. While this study is limited by its lack of a control—educational controls are notoriously challenging both due to the difficulty of controlling for all possible confounds as well as the moral dubiousness of withholding the hypothesized "best" instruction from some students—it does show the promise of curiosity-driven recommendation. Developing educational systems like Pique can help students expand their knowledge by recommending novel scientific articles. While we cannot claim that student curiosity was entirely due to Pique, we conclude that the approach of encouraging curiosity Pique shows is promising for our future research on computational novelty in open-ended learning environments.

In summary, this paper makes 2 major contributions: a framework for structuring the AI component of educational recommender systems to encourage curiosity and the Pique model that integrates the AI component and its interaction with a learner model and course materials. We demonstrate how the framework is integrated in the Pique model, providing opportunities for future studies that leverage other models of novelty and personalization. The paper describes the students' experience with Pique demonstrating how their interests expanded over the period of a semester. Future studies that collect data from a larger number of students would allow an analysis of the relationships between students' expanding interests and the novelty score of the recommended learning materials.

Acknowledgements The research reported in this paper was funded by NSF Award number #1618810: RI: Small: CompCog: Pique: A Cognitive Model of Curiosity for Personalizing Sequences of Learning Resources.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Imran H, Belghis-Zadeh M, Chang T-W, Graf S, et al. Plors: a personalized learning object recommender system. Vietnam J Comput Sci. 2016;3(1):3–13.
2. Niu X, Abbas F, Maher ML, Grace K. Surprise me if you can: Serendipity in health information. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; 2018. p. 1–12.
3. Wundt WM. Principles of physiological psychology. Leipzig: Wilhelm Engelmann; 1874.
4. Boyle GJ. Critical review of state-trait curiosity test development. Motiv Emotion. 1983;7(4):377–97.
5. Kashdan TB, Fincham FD. Facilitating curiosity: A social and self-regulatory perspective for scientifically based interventions. In: Alex Linley P, Joseph S, editors. Positive psychology in practice. Hoboken: Wiley; 2004.
6. Loewenstein G. The psychology of curiosity: A review and reinterpretation. Psychol Bull. 1994;116(1):75.
7. Kang MJ, Hsu M, Krajbich IM, Loewenstein G, McClure S.M, Wang J.T.-y, Camerer C.F. The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. Psychol Sci. 2009;20(8):963–73.
8. Vygotsky L. Interaction between learning and development. In: Readings on the development of children. New York: Scientific Academic Books; 1978. p. 34–41.
9. Schmidhuber J. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). IEEE Trans Auton Mental Dev. 2010;2(3):230–47.
10. Merrick KE, Maher ML. Motivated reinforcement learning: curious characters for multiuser games. Berlin: Springer; 2009.
11. Grace K, Maher ML. Specific curiosity as a cause and consequence of transformational creativity. In: ICCC; 2015. p. 260–267.
12. Grace K, Maher ML. Surprise and reformulation as meta-cognitive processes in creative design. In: Proceedings of the Third Annual Conference on Advances in Cognitive Systems ACS; 2015. p. 8.
13. Grace K, Maher ML, Mohseni M, Perez Y, Perez R. Encouraging p-creative behaviour with computational curiosity. ICCC; 2017.
14. Mohseni M, Maher ML, Grace K, Najjar N, Abbas F, Eltayeby O. Pique: Recommending a personalized sequence of research papers to engage student curiosity. In: International Conference on Artificial Intelligence in Education. Springer; 2019. p. 201–205.
15. Mohseni M, Maher ML. A framework for exploring computational models of novelty in unstructured text. In: Proceedings of the 6th International Conference on Information System and Data Mining. ICISDM '22. Association for Computing Machinery, New York, NY, USA; 2022. pp. 36–45.
16. Barto A, Mirolli M, Baldassarre G. Novelty or surprise? Front Psychol. 2013;4:907.
17. Adamopoulos P, Tuzhilin A. On unexpectedness in recommender systems: Or how to better expect the unexpected. ACM Trans Intell Syst Technol. 2014;5(4):1–32.
18. Kokotsaki D, Menzies V, Wiggins A. Project-based learning: A review of the literature. Improv Schools. 2016;19(3):267–77.
19. Guk I, Kellogg D. The zpd and whole class teaching: Teacher-led and student-led interactional mediation of tasks. Lang Teach Res. 2007;11(3):281–99.
20. Grace K, Maher ML, Davis N, Eltayeby O. Surprise walks: Encouraging users towards novel concepts with sequential suggestions. In: Proceedings of the 9th International Conference on Computational Creativity (ICCC 2018). Association of Computational Creativity; 2018.
21. Siddiqui S, Maher ML, Najjar N, Mohseni M, Grace K. Personalized curiosity engine (pique): A curiosity inspiring cognitive system for student directed learning. In: CSEDU (1); 2022. p. 17–28.
22. Baker T, Smith L, Anissa N. Educ-ai-tion rebooted? exploring the future of artificial intelligence in schools and colleges. NESTA; 2019.
23. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education-where are the educators? Int J Educ Technol Higher Educ. 2019;16(1):1–27.
24. Drachsler H, Verbert K, Santos OC, Manouselis N. Panorama of recommender systems to support learning. In: Ricci F, Rokach L, Shapira B, editors. Recommender systems handbook. New York: Springer; 2015. p. 421–51.
25. An S, Bates R, Hammock J, Rugaber S, Weigel E, Goel A. Scientific modeling using large scale knowledge. In: International Conference on Artificial Intelligence in Education. Springer; 2020. p. 20–24.
26. Ai F, Chen Y, Guo Y, Zhao Y, Wang Z, Fu G, Wang G. Concept-aware deep knowledge tracing and exercise recommendation in an online learning system. International Educational Data Mining Society; 2019.
27. Pozdniakov S, Posov I, Anton C. Interaction of human cognitive mechanisms and "computational intelligence" in systems that support teaching mathematics. In: International Conference on Intelligent Tutoring Systems. Springer; 2021. p. 259–266.
28. Kuo R, Krahn T, Chang M. Behaviour analytics-a moodle plug-in to visualize students' learning patterns. In: International Conference on Intelligent Tutoring Systems. Springer; 2021. p. 232–238.
29. Papamitsiou Z, Economides AA. Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. J Educ Technol Soc. 2014;17(4):49–64.
30. Del Solato T, Du Boulay B. Implementation of motivational tactics in tutoring systems. J Artif Intell Educ. 1995;6:337–78.

31. Keller JM. Strategies for stimulating the motivation to learn. Perform Instr. 1987;26(8):1–7.
32. Jones A, Issroff K. Learning technologies: Affective and social issues in computer-supported collaborative learning. Comput Educ. 2005;44(4):395–408.
33. Hill JR, Land SM. Open-ended learning environments: A theoretical framework and model for design. In: Proceedings of Selected Research and Development Presentations at the National Convention of the Association for Educational Communications and Technology; 1998.
34. Kose U, Arslan A. Intelligent e-learning system for improving students' academic achievements in computer programming courses. Int J Eng Educ. 2016;32(1):185–98.
35. Hall OP Jr, Ko K. Customized content delivery for graduate management education: Application to business statistics. J Stat Educ. 2008;16(3).
36. Cobos C, Rodriguez O, Rivera J, Betancourt J, Mendoza M, León E, Herrera-Viedma E. A hybrid system of pedagogical pattern recommendations based on singular value decomposition and variable data attributes. Inf Process Manag. 2013;49(3):607–25.
37. Barria Pineda J, Brusilovsky P. Making educational recommendations transparent through a fine-grained open learner model. In: Proceedings of Workshop on Intelligent User Interfaces for Algorithmic Transparency in Emerging Technologies at the 24th ACM Conference on Intelligent User Interfaces, IUI 2019, Los Angeles, USA, March 20, 2019, vol. 2327; 2019.
38. Freyne J, Berkovsky S. Intelligent food planning: personalized recipe recommendation. In: Proceedings of the 15th International Conference on Intelligent User Interfaces; 2010. p. 321–324.
39. Maher ML, Grace K. Encouraging curiosity in case-based reasoning and recommender systems. In: International Conference on Case-Based Reasoning. Springer; 2017. p. 3–15.
40. Carayol N, Agenor L, Oscar L. The right job and the job right: Novelty, impact and journal stratification in science. Impact and Journal Stratification in Science (March 5, 2019); 2019.
41. Harris ZS. Distributional structure. WORD. 1954;10(2–3):146–62. https://doi.org/10.1080/00437956.1954.11659520. (**Accessed 2017-03-31**).
42. Salton G, McGill MJ. Introduction to modern information retrieval. New York: McGraw-Hill; 1983.
43. Mikolov T, Yih W-t, Zweig G. Linguistic regularities in continuous space word representations. In: Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; 2013. p. 746–751.
44. Pennington J, Socher R, Manning CD. Glove: Global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP); 2014. p. 1532–1543.
45. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res. 2003;3(Jan):993–1022.
46. Blei DM, Lafferty JD. A correlated topic model of science. Ann Appl Stat. 2007;1(1):17–35.
47. Blei DM. Probabilistic topic models. Commun ACM. 2012;55(4):77–84.
48. Grace K, Maher ML, Fisher D, Brady K. Data-intensive evaluation of design creativity using novelty, value, and surprise. Int J Design Creat Innov. 2015;3(3–4):125–47.
49. Maher ML, Fisher DH. Using ai to evaluate creative designs. In: DS 73-1 Proceedings of the 2nd International Conference on Design Creativity Volume 1; 2012.
50. Herlocker JL, Konstan JA, Terveen LG, Riedl JT. Evaluating collaborative filtering recommender systems. ACM Trans Inf Syst. 2004;22(1):5–53.
51. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948;27(3):379–423.
52. Baldi P, Itti L. Of bits and wows: A Bayesian theory of surprise with applications to attention. Neural Netw. 2010;23(5):649–66.
53. Pimentel MA, Clifton DA, Clifton L, Tarassenko L. A review of novelty detection. Signal Process. 2014;99:215–49.
54. Jumarie G. Relative information-what for? In: Relative information. Berlin: Springer; 1990. p. 1–11.
55. Kullback S. Information theory and statistics. New York: Dover Publication, Inc.; 1968.
56. Bouma G. Normalized (pointwise) mutual information in collocation extraction. Proc GSCL. 2009;30:31–40.
57. Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat. 1951;22(1):79–86.
58. Roberts ME, Stewart BM, Tingley D. Stm: An r package for structural topic models. J Stat Softw. 2019;91(1):1–40.
59. Berlyne DE. Curiosity and exploration. Science. 1966;153(3731):25–33.
60. Schön DA. The reflective practitioner: how professionals think in action. London: Routledge; 2017.
61. Kolb DA. Learning style inventory. Boston, MA: McBer and Company; 1999.
62. Cowan J. On becoming an innovative university teacher: reflection in action: reflection in action. London: Society for Research Into Higher Education & Open University Press; 2006.
63. Zhang J, Li C, Kosov S, Grzegorzek M, Shirahama K, Jiang T, Sun C, Li Z, Li H. Lcu-net: A novel low-cost u-net for environmental microorganism image segmentation. Pattern Recogn. 2021;115: 107885. https://doi.org/10.1016/j.patcog.2021.107885.
64. Zhang J, Li C, Yin Y, Zhang J, Grzegorze M. Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer. Artif Intell Rev. 2023;56:1013–70. https://doi.org/10.1007/s10462-022-10192-7.
65. Li X, Li C, Rahaman MM, Sun H, Li X, Wu J, Yao Y, Grzegorzek M. A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches. Artif Intell Rev. 2022;55:4809–78. https://doi.org/10.1007/s10462-021-10121-0.
66. Kulwa F, Li C, Zhang J, Shirahama K, Kosov K, Zhao X, Jiang T, Grzegorzek M. A new pairwise deep learning feature for environmental microorganism image analysis. Environ Sci Pollut Res. 2022;29:51909–26. https://doi.org/10.1007/s11356-022-18849-0.
67. Chen A, Li C, Zou S, Rahaman MM, Yao Y, Chen H, Yang H, Zhao P, Hu W, Liu W, Grzegorzek M. Svia dataset: A new dataset of microscopic videos and images for computer-aided sperm analysis. Biocybern Biomed Eng. 2022;42(1):204–14. https://doi.org/10.1016/j.bbe.2021.12.010.