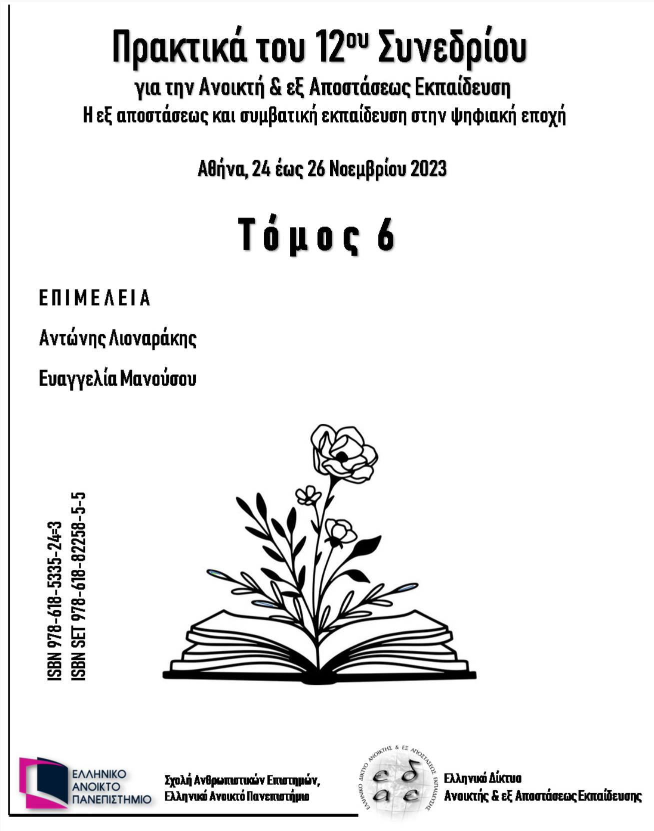
Διεθνές Συνέδριο για την Ανοικτή & εξ Αποστάσεως Εκπαίδευση

Τόμ. 12, Αρ. 6 (2023)

ICODL2023



Employing a Process Mining Approach to Recommend Personalized Adaptive Learning Paths in Blended-Learning Environments

Noura Joudieh, Nikleia Eteokleous, Ronan Champagnat, Mourad Rabah, Samuel Nowakowski

doi: [10.12681/icodl.5657](https://doi.org/10.12681/icodl.5657)

Εφαρμογή Προσέγγισης Εξόρυξης Διαδικασιών για Ανάπτυξη Εξατομικευμένων Προσαρμοστικών Διαδρομών Μάθησης σε Μικτά Περιβάλλοντα Μάθησης

Employing a Process Mining Approach to Recommend Personalized Adaptive Learning Paths in Blended-Learning Environments

Noura Joudieh
PhD Researcher
L3i Laboratory
La Rochelle University
La Rochelle, France
noura.joudieh@univ-lr.fr

Ronan Champagnat
Associate Professor
L3i Laboratory
La Rochelle University
La Rochelle, France
ronan.champagnat@univ-lr.fr

Mourad Rabah
Associate Professor
L3i Laboratory
La Rochelle University
La Rochelle, France
mourad.rabah@univ-lr.fr

Nikleia Eteokleous
Associate Professor
Frederick University
Limassol and Nicosia, Cyprus
n.eteokleous@frederick.ac.cy

Samuel Nowakowski
Associate Professor
Loria Laboratory
Lorraine University
Nancy, France
samuel.nowakowski@univ-lorraine.fr

Abstract

In education, e-learning is highly adopted to improve the learning experience and increase learning efficiency and engagement. Yet, an explosion of online learning materials has overwhelmed learners, especially when trying to achieve their learning goals. In this scope, recommender systems are used to guide learners in their learning process by filtering out the available resources to best match their needs, i.e. to offer more personalized content and learning paths. Concurrently, process mining has emerged as a valuable tool for comprehending learner behavior during the learning journey. To synergize these disciplines and optimize learning outcomes, our paper introduces an ontology-based framework that aims to recommend an adaptive learning path, driven by a learner's learning objective, personalized to his learning style and enriched by the past learning experience of other learners extracted via process mining. The learning path considers pedagogical standards by employing Bloom's taxonomy within its structure. The framework establishes an Ontological Foundation, to model the Learner, Domain Knowledge, and Learning Path. Choosing Computer Science as a domain, we construct a knowledge base using synthesized data to underpin the framework. For past learning experience, we analyze Moodle log data from 2018 to 2022, encompassing 471 students in the Computer Science and Engineering Department at Frederick University, Cyprus.

Keywords

E-Learning, Personalized Learning Paths, Ontology, Recommender Systems, Process Mining

1. Introduction

In a rising world of technology, education obtained a decent share of development opening the door to enhanced learning techniques, environments, and technologies. Traditional classrooms of a teacher-learner scenario in an academic setting have evolved to a wider concept, including various types of learning environments. Distance, blended, open, flexible and personalized learning environments will govern the university settings in the future (Kafa & Eteokleous, 2022). In the aforementioned environments, the exclusivity of education and learning materials is open to everyone, at anytime, anywhere (Colace et al., 2014). This scenario has changed a set of standards, starting with the identity of a learner that became anyone who wants to learn something rather than a student in an academic institution. Additionally, it provided learners with a vast scope of online available educational materials to nourish their knowledge curiosity at their own pace. Yet, this availability of learning resources can overwhelm a learner when trying to achieve his learning objective that may eventually lead him to a loss of motivation and a decline in the learning efficiency.

In this direction, Recommender Systems (RS), which are computer-based techniques that try to “intelligently” filter out the online content for a user based on his preferences, previous actions, needs, etc. Aggarwal (2016), gained interest in several domains, including the domain of e-learning (Pireva & Kefalas, 2018). Using various data about the learner as his preferences, learning objective, restrictions, background knowledge, etc. these systems can rank the available learning resources according to their compatibility with the given learner data and recommend him the “best matches”, thus offering a more personalized learning experience (Singh et al., 2021). This personalization went further to consider more aspects about the learner in the recommendation process, as his learning style, cognitive abilities, personality, etc. (Nabizadeh et al., 2020). Therefore, deviating from the “one-size-fits-all” models, RS in e-learning are used to tailor the learning experience to the uniqueness of each learner by personalized learning resources, and learning paths (Raj & Renumol, 2022).

We define a learning path as a partially ordered sequence of learning objects that aims to achieve a learning objective. A learning object can be a learning resource, a course, an action or even a sub-process while a learning objective can vary among obtaining a degree, finishing a course, or acquiring a skill. For example, in Learning Management Systems (LMS) like Moodle, a student follows a sequence of learning resources provided by the instructor to finish an enrolled course of his.

In fact, learners’ studying traces in an LMS or an educational platform are captured by integrated logging systems and recorded as an “event log” containing information about every action performed by any user in the system at each point in time.

Process Mining (PM) (van der Aalst, 2016), a discipline that combines data mining, machine learning and process modeling, then mines from this event log, a process model that reveals the real process followed by users in the system. The former has paved the way for a rich line of research to discover the behaviors of students while performing different learning activities like attending a course, or taking a quiz (Nan Cenka, 2022). Also, the discovered process models can aid instructors and students to identify any bottlenecks, successful practices or patterns in the overall behavior, and thus making appropriate and informed decisions to enhance the learning process.

Using PM for a recommendation process is modestly used in e-learning and rarely harnessed to recommend a learning path (Yari Eili & Rezaeenour, 2022). Beyond existing literature on personalized learning path recommendations, this paper introduces an ontology-based framework that leverages PM to extract students' past learning experiences from event logs, integrating this insight into the personalized recommendation of an adaptive learning path tailored to each learner's profile. This approach can then be adopted in different learning scenarios to provide personalized learning paths. For example, it can serve distance or blended learning in enriching the experience, ameliorating the learning efficiency and emphasizing on the learners' engagement through personalized learning paths.

The subsequent sections offer an overview of related research in e-learning, encompassing recommendations, PM, and ontologies. This is followed by the articulation of the motivation and the main objective. The proposed framework and its ontological foundation are then explored, along with the knowledge base creation and data collection. Finally, the paper discusses some future perspectives.

2. Related Works

RS, a branch of information retrieval and filtering, are programs that try to understand the preferences and behavior of a user and predict him the most suitable item or service. In several domains like business, e-commerce social networks, finance, healthcare, and others, RS played an important role in improving the provided services and user experience by filtering the choices to the ones that mostly match the user's interest (Beheshti et al., 2020). In education, RS gained ground for personalizing the learning experience within the various emerging types of learning environments, as they have shown positive impacts on increasing academic achievements, helping students facing learning difficulties and emphasizing learning engagement (Nabizadeh et al., 2020). Several works addressed personalizing the recommendation of learning objects. In their work, Nafea et al. (2019) implemented a personalized learning object recommender to learn a course in a LMS, tailored by the learner's learning style. Course instructors set learning object (LO) profiles, while student profiles are collected through questionnaires. The process involves K-means clustering for LOs, repeated with each new addition of an LO. Recommendations are derived by calculating similarity between the student's learning style and cluster centroids, followed by ranking prediction and selection of top N learning objects. These are further presented to the learner in order of ranking. Tarus et al. (2017) introduced a novel hybrid knowledge-based methodology for personalized learning resource recommendations, combining collaborative filtering with ontologies and Sequence Pattern Mining (SPM). Learners assigned ratings to learning resources available on the LMS portal, assessing relevance to their courses on a 5-point scale. During the initial registration phase, they supplied fundamental information and rated some resources to enhance the ontology and their learner profiles. The RS subsequently predicts the learner's ratings of unseen learning resources and generates the list of top N learning items. Then, the learning patterns of the learner extracted from web logs by SPM are used to filter and generate the final list of learning items. Unlike PM that learns a model from a given log data, SPMs are used to discover frequent subsequences in a sequence dataset like web logs (Mabroukeh & Ezeife, 2010). Ontologies defined by Gruber (1993) as "a formal explicit specification of conceptualization" are a mechanism to model the knowledge where concepts and the relations among them are explicitly defined in a machine readable manner. They offer a reasoning mechanism that infers new knowledge from the already existing one. Being so, they are highly adopted by several works (Yarandi et al., 2013; Colace et al., 2009), to model knowledge and context about the learner and the domain. Sudhana et al. (2013) proposed an ontology-driven framework for context-aware adaptive e-learning, with the primary aim of tailoring learning material to individual learner contexts. This initiative focuses on seeking learning resources within a web-based e-learning environment. Personalization parameters encompass background, learning style, and learning objectives. While, Wu et al. (2020) leveraged the reasoning power of ontologies to personalize the learning material in a web-based environment via semantic rules applied on a knowledge base consisting the learner, learning resource and domain ontology.

When recommending a learning path, the problem is formulated while defining multiple aspects: (a) the structure / definition of the learning path, (b) the characteristics of the path (adaptive and/or personalized) (c) the personalization parameters if any, (d) the main context and involved actors, and (e) the recommendation technique. A learning path is a sequence of learning contents that guide the user to accomplish a learning goal. The learning content may vary among a course, a topic / a concept, a learning object, or an action, etc. (Costa et al., 2022; Nabizadeh et al., 2020). A learning object might be a video, a text file, an audio, or any other form of media. In a proposal of a RS for course learning, Cheng et al. (2018) define a path as a sequence of video content personalized by the learner's knowledge mastery. The approach is based on three ontologies, namely the Learner ontology, the Video ontology and the Domain ontology which models the course curriculum and is designed by the help of the course instructors. As Nitchot et al. (2019) considers the learning path as a sequence of web links in a web-based environment, while it is a sequence of learning content for Yu et al. (2007) who proposed a semantic-based approach to recommend a learning path personalized by the contextual information of the learner. The former includes the learner prior knowledge and learning goal. Ontologies are used to model the domain, the learning content and the learner. Selecting a target learning content by the learner, the learning path is constructed using the pre-requisite relation among learning content and the knowledge extracted from the domain ontology. The goal in the work of Nabizadeh et al. (2017) is to offer personalized adaptive learning paths for learning courses tailored by the learners' time constraints and knowledge background. Courses consist of lessons, each comprising multiple learning objects, and so is a learning path. A graph algorithm is employed to derive potential paths for studying a course which are then filtered to retain only those aligning with the student's time limitations while maximizing their score. Three approaches are employed to evaluate score and time for each path, with time and score estimated for individual learning objects. If the student falls short of the target score at a step, auxiliary learning objects are suggested to aid their progress. Finally, Bian (2019) proposes the use of graph theory and an improved immune algorithm to recommend a personalized learning path for each learner by their learning style and knowledge. A path is considered as a linear sequence of learning objects to finish a course. Yet, instructors are involved to construct the concept map of a course while modeling the prerequisites among the underlying concepts. The approach was tested in a flip classroom for a python programming course where students choose their learning goal from the concepts in the concept map to get a personalized path to their learning style.

On the other hand, information systems record the actions performed by their users in a form of an *Event Log*. The former records every activity taking place along with the time of its occurrence and a unique identifier that refers to the user of that activity, respectively known as the *activity*, the *timestamp* and the *case id*. Using the case id of a user, an ordered sequence of his performed activities can be extracted from the event log, namely *a trace*. As a single trace can tell about the behavior of one user in the system, from an event log, which is a set of traces, PM discovery algorithms are used to learn a process model that represents the general behavior of all the users in the system (van der Aalst, 2016). Trabelsi et al. (2019) applied PM to event logs from the Gallica digital library. Traces were initially grouped into three categories: lookup, borderline, and exploratory. Lookup traces represent users accessing documents with few manipulations, borderline traces pertain to those targeting a specific subject area, and exploratory traces involve wide-ranging document access. Subsequently, a process model was mined from each group's traces to comprehensively comprehend and analyze user journeys within the digital library. In e-learning, on the other hand, Juhaňák et al. (2019) employed PM to extract the patterns and interactions of students in an LMS while taking online-based quizzes. While, Bey and Champagnat (2022) conducted a study in a context of a programming course, to analyze the paths taken by students in 17 exercises. Students were first grouped into six clusters depending on their behavior. They were classified as “at-risk” or “good” using their final score on the course. Finally, the mined process model for each category revealed the trajectory patterns for high and low performing students while programming, thus determining which behaviors can contribute to the success or failure in a programming course. PM techniques are also used to discover and analyze learning strategies of learners in different learning environments as self-regulated learning (Bannert et al., 2014) or MOOCs (Rohani et al., 2023).

Learning environment logs store student learning history for performance evaluation, grouping based on similarities, and detecting undesirable behavior. Yet, it can also be used to recommend them appropriate learning material (Raj & Renumol, 2022). In this scope, Wang and Zaïane (2018) proposed diverse sequence-based strategies for a Course Recommender System (CRS), including a PM approach. The CRS aims to select the optimal next course for individual undergraduates in higher education, guided by their past course history. This history encompasses course records, performance (grades), and sequencing. To recommend the next course, a similarity measure identifies successful students akin to the current one, with their subsequent courses forming candidates. These are ranked considering criteria like being mandatory and agility (average graduation time). The top-ranked course is recommended. This CRS enhanced instructors' understanding of prerequisites, accelerated students' graduation, and boosted their performance. PM was used for recommendation activities in few works in other domains. In fact, in the works of Huber et al. (2015) and Schonenberg et al. (2008), PM tools and algorithms were used to mine the model of a business process to enable the recommendation of the possible next steps during the execution of the former. Terragni and Hassani (2018, 2019) employed web logs, transformed into event logs, to extract a process model encapsulating customer journeys on a website. Focusing on enhancing the "click product link" Key Performance Indicator (KPI), the process model analysis revealed that optimizing the "Visit Product Page" action positively influences this KPI. Leveraging user interactions with pages, they devised a RS that tailors links on the currently viewed page based on user behavior. This approach led to an increase in the "Visit Product Page" action, effectively optimizing the designated KPI. In healthcare, Pereira Detro et al. (2020), Yu et al. (2007), and Yang et al. (2017) used PM to discover the treatment process of patients and used this model to give recommendations on the next medical activity to perform in a treatment process. Despite the substantial progress in both fields of PM and RSs in different domains and research lines, using PM for a recommendation activity is yet modestly explored especially in e-learning. This paucity of exploration is particularly evident in the limited application of PM in recommending a learning path. Consequently, this work intends to propose an ontology-based framework that uses PM to extract and integrate learners' past educational experiences in the recommendation process for a personalized adaptive learning path for a learner aiming at a learning goal. The proposed approach can find ground in different types of learning environments like blended, distance, open learning, and others.

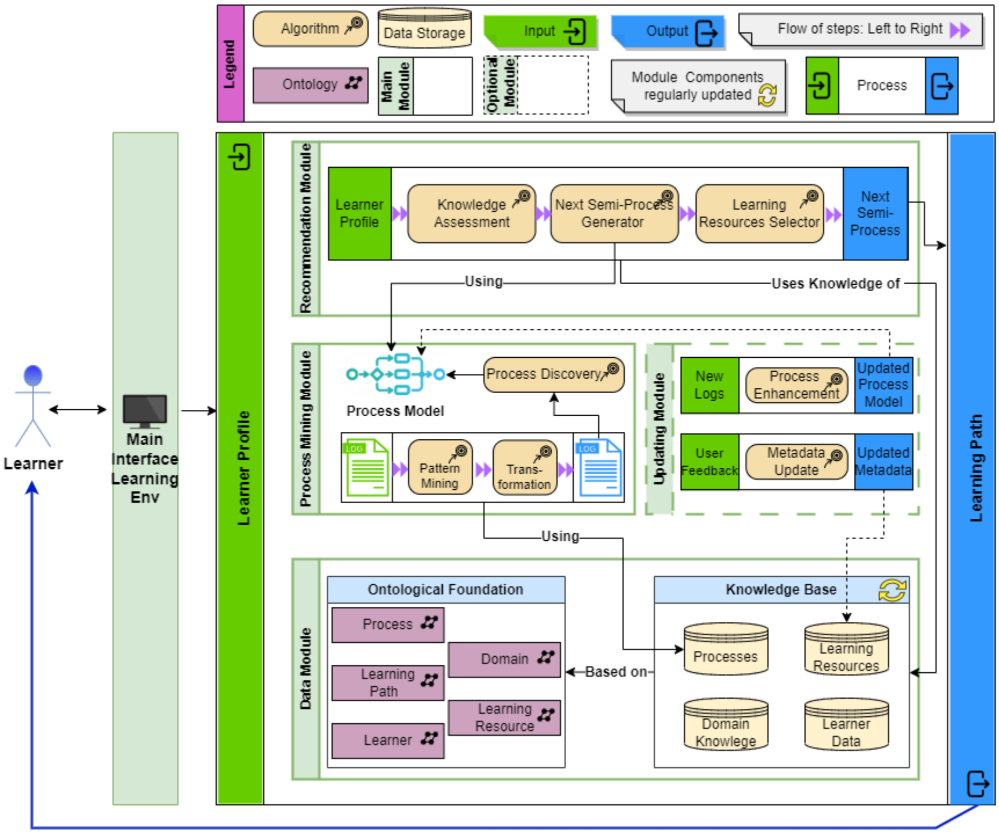
3. Motivation and Problem Statement

The literature extensively explores personalized strategies for presenting learning content and learning paths based on factors like learning style, time restrictions, knowledge level, and objectives. Predominantly, knowledge-based and ontology-based methods are favored, effectively addressing challenges such as RS cold-start issues. Ontologies, valued for their adeptness in knowledge representation and reasoning, have gained wide acceptance for their ability to infer new knowledge from existing one. In response, our framework revolves around ontologies as its core, integrating the less-explored in recommending a learning path, yet promising technique of PM to extract past learning experience.

Thus, the main objective becomes to personalize the learning process through the recommendation of an adaptive learning path driven by a learner's learning goal while using PM to extract past learning experience.

4. Proposed Framework

Considering the stated objective and the state-of-the-art, the following framework is derived and illustrated in this section. As Picture 1 shows, the learner, who is the main actor, interacts with the learning environment through the user interface. The framework takes as an input his profile communicated from the main interface, and gradually returns the learning path as an output. Thus, from the learner's perspective, first he chooses his learning goal, then takes a questionnaire to detect his learning style and pre-knowledge about the goal, after he starts receiving some learning contents. At each state, he gets assessed to decide whether to expand the current state or move to the next. The cycle repeats until reaching the goal. In what follows, the structure of the input and output, the ontological foundation and the role of the underlying modules of the framework are detailed.



Picture 1: The Proposed Framework

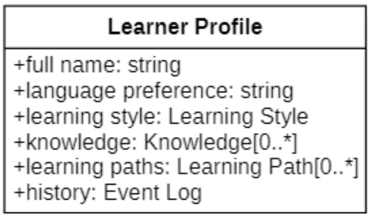
4.1. Input and Output

The structure of the learner profile, being the input and the learning path, being the output, is as follows.

Learner's Profile

There are several propositions to standardize a learner profile (Dolog & Nejdl, 2003) as PAPI – IEEE Personal and Private Information (Farance, 2000) or IMS LIP – Learner Information Package (IMS, 2001). In this work, the learner profile is inspired by the existing standards yet customized to the context as presented in Picture 2.

The profile has the personal information along with the *language preference*. The *knowledge* is a collection of the learner's pre-knowledge, captured via a questionnaire prior to the learning process, and his mastered ones recorded throughout the process.

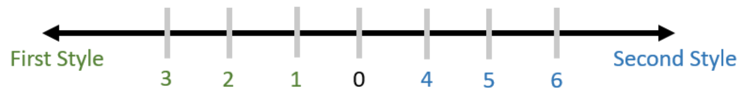


Picture 2: *Learner Profile*

The *learning paths* (elaborated in the following subsection) have all the information about the learning goals and the progress of each, while the history records all the actions done by the user in the learning environment in the form of an event log. We define the learning goal (LG) via pedagogical means, as *LG = {Knowledge, Learning Level}*.

The *Knowledge* is what the learner wants to learn chosen from the domain, while the *Learning Level* could be either "Surface" (short term usage and retaining of information), "Intermediate" (average term usage and retaining of information), or "Profound" (active processing of information on long term with the ability to elaborate on the acquired information rather than only memorizing). To translate the learning levels into our context, we used the revised version of Bloom's taxonomy (Fastiggi, 2019; Shabatura, 2022) that defines six successive levels of learning along with their associated verbs to represent the educational goals. Thus, we consider that the first two levels of "Remembering" and "Understanding" belong to the surface level, the next two levels of "Applying" and "Analyzing" belong to the intermediate level, and the last two of "Evaluating" and "Creating" belong to the profound level. For example, if the learning goal of a learner is "being able to apply clustering", it is translated to *LG = {Clustering, Intermediate}*.

For the *learning style*, being highly adopted and gaining a wide acceptability in literature, we chose Felder and Silverman's model (Felder & Silverman, 1988) captured by the ILS questionnaire (Felder & Silverman, 1996) (Index of Learning Style Questionnaire) that comprises 44 questions. This model expresses the preferred style of learning on four dimensions, each with two styles within: (a) processing information (active, reflective), (b) perceiving information (sensing, intuitive), which shows the preferred type of information to be received (c) receiving information (visual, verbal), that spots the sensory channel through which external information is most effectively perceived and (d) understanding information (sequential, global) that questions whether a learner progresses into understanding through sequential detailing of the information or a global overview. Following the explanation of the results of the ILS given by (Felder & Soloman, 1993), we define a learning style (S) as a vector of 4 values $S = \{d_1, d_2, d_3, d_4\}$, one for each dimension, where $d_x \epsilon\ [0,6]$ expresses to which style in this dimension is the learner more bias and to which extent as depicted in Picture 3. A zero value signifies no preference to any style. For example, a value 6 for the third dimension, means the learner has an extreme verbal style, thus it might be difficult for him to perceive information if it was not verbal.

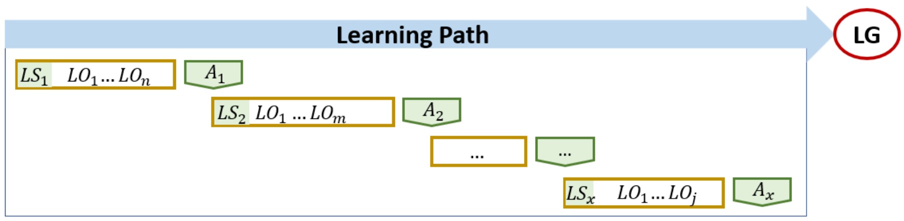


Picture 3: *Learning Style Encoding*

Learning Path Structure

The recommended learning path is expected to be adaptive and personalized, which is translated in its structure. Thus, we define a learning path (LP) to reach a learning goal as $LP = \{[LS_1, A_1], [LS_2, A_2], \dots [LS_n, A_n]\}$, where LS_x is a *Learning State* and A_x is its following *Assessment* phase. A *Learning State* (LS) is characterized by a *learning level* and a goal concept, and defined as the set of learning objects that helps achieve the goal concept at the specified learning level. Thus, $LS = \{ LO_1, LO_2, \dots, LO_n\}$ where LO_x, is a learning object. The following assessment is initially a questionnaire that asks if the goals of the learning state are satisfied given its learning level.

As seen in Picture 4, the learning path is delivered one learning state at a time, and according to the assessment results, the next state is created thus being adaptive. While, the personalization takes place at the level of the learning objects chosen, where a compatibility value is computed for a learning object considering its learning style, covered concepts, and learning level with the learning style of the learner and the characteristics of the learning state.



Picture 4: *Learning Path Structure*

The proposed framework of the study has several important implications for the instructors, instructional designers, universities as well as learners. The instructors and instructional designers that are responsible for designing, developing and delivering courses are expected to take into consideration the important information to be provided in regards to learners' profile, style and expectations. Designing a personalized learning path involves tailoring the educational experience to the unique needs, preferences, and abilities of each learner. This approach requires a shift in the role of instructional designers and instructors, focusing on customization, flexibility, and active involvement. Implications for instructional designers when creating a personalized learning path for learners can be summarized as follows: understand learner profiles, customize learning content, develop flexible learning pathways, create adaptive assessment strategies, integrate technology, and monitor and analyze data. Implications for instructors when creating a personalized learning path for learners, differ at some level from the instructional designers, and can be summarized as follows: facilitate personalized learning, provide individualized support and feedback, build a learning community, promote self-regulation and reflection, continuous professional development trainings, and cultivate empathy and understanding towards diverse needs and backgrounds.

When learners have a personalized learning path, it means their educational experience is tailored to their specific needs, interests, learning styles, and pace. This approach has significant implications for learners, which can positively impact their engagement, motivation, understanding, and overall academic success. Some key implications of personalized learning paths for learners are the following: increased engagement and motivation, better understanding and retention, customized learning pace, addressing diverse learning styles, individualized support and feedback, empowerment and autonomy, enhanced critical thinking and problemsolving skills, flexibility and adaptability, long-term retention, and lifelong learning mindset.

When universities implement personalized learning paths, it signifies a shift towards a more student-centered and tailored educational experience. This approach has several implications for universities, affecting various aspects of their structure, processes, and outcomes, such as: restructure curriculum in order to be customizable, invest and integrate new, emergent and adaptive technologies, provide continuous development training for the faculty members, promote accessibility and inclusivity, develop appropriate quality assurance and accreditation mechanisms, etc. Implementing personalized learning paths in universities requires a comprehensive approach involving technology, training, student support, assessment strategies, and an emphasis on flexibility and individualization. Universities should strive to create an educational environment that empowers students to take charge of their learning while maintaining academic standards and supporting their holistic development.

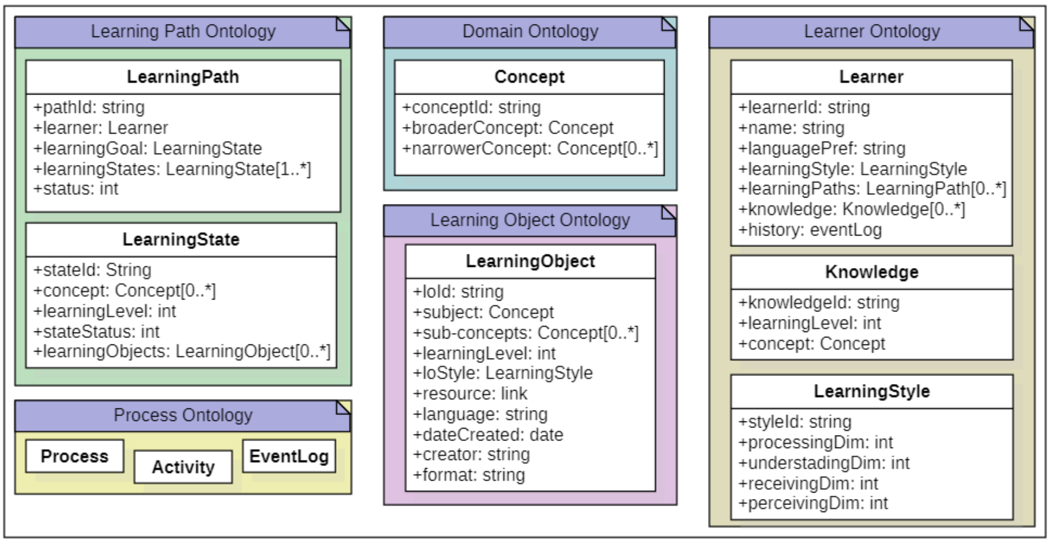
Finally, a personalized learning path can lead to a more engaging, effective, and fulfilling educational experience for learners, addressing their unique needs and preparing them for success in both academic and real-world settings.

4.2. The Ontological Foundation

The framework that results in the learning path starting from the learner's profile, constitutes three main modules and another optional. The main entities in the whole framework are the Learner, Learning Path, Learning Objects, Learning Process and a chosen Domain. To represent each of these entities with their metadata, an ontology is designed using Protégé. The created ontologies formed the Ontological Foundation of the framework, presented in Picture 5. The Learner and Learning Path ontologies represent the structure of the input and output as previously explained. The domain ontology, used to express the knowledge about the chosen domain, is SKOS-based rendering it understandable and easily reused.

Also, inspired by the LOM standard (*IEEE Standard for Learning Object Metadata*, 2020) in the Learning Object ontology the chosen metadata to record for a learning object are specified. We added the learning level and learning style to serve the context. Finally, the Process Ontology is used to represent the PM related entities. This foundation is then used as a model to create the knowledge base of the framework, in which the data is saved. The Ontological Foundation together with the Knowledge Base make the Data Module of the framework.

The Process Mining Module is responsible for extracting the past learning experience from the event logs and saving them in the Processes data storage. The procedure starts by preprocessing the event logs, then transforming them into high-level traces from which a process model is discovered. The model and the extracted information are then saved in the knowledge base and later used for recommendation.



***Picture 5**: UML Class Diagram of the Ontological Foundation*

Following, the Recommender Module is responsible for the core functionality of constructing the adaptive personalized learning path. Starting with the learner profile, the learning goal indicates the goal knowledge and the target learning level. At first, the concepts to be covered are extracted from the domain ontology, then the plan of each learning state is created using the information about the past experience and the knowledge of the learner at the moment. After, the learning objects with the highest compatibility values with the state are chosen. An assessment then takes place before the next state is created. Finally, this cycle repeats until the learning goal is reached.

An optional “Updating Module” might be added to the framework where the feedback on the recommended learning objects and path is recorded and considered in future recommendations. Also, where the history of the learner is used to enhance the process model initially built.

5. Data Collection and Primary Results

Following the objectives of the work (cf. section 3), at this point, the proposed framework was designed, with the learning path structure, the learner's model and the Ontological Foundation. To validate the former, Computer Science was chosen as a domain in which a synthesized data of learners, learning objects, domain knowledge and learning goals is created to build a knowledge base. Also, to initiate with the process mining module, a Moodle log data is collected and preprocessed. These two data-related phases are covered in this section along with a glimpse of the resulting knowledge base, followed by the next phases in the work.

5.1. Data-Related Phase

The data related phase has a part connected to the event logs collection, i.e. related to process mining module, and two other parts related to the knowledge base.

Process Mining Related: The Event Log Data

As the source data for process discovery algorithms is system logs, in this study, we needed to collect log data that records the behavior and paths taken by students while studying courses, i.e. while trying to achieve a learning objective. In what follows, the data collection and preprocessing processes are explained.

Event Logs Collection Process

Moodle, a well-known learning management system, widely used in universities and educational institutions, has a logging system that records the actions taken by any user in the system at each point in time. In regards to the current study, the event logs from Moodle were collected as an input source for the PM module. Thus, the Moodle event logs of 471 students taking courses in the Department of Computer Science and Engineering of Frederick University in Cyprus for the period of 2018 until 2022, were used. The log data records all the actions taken by students on Moodle while learning, including taking courses, exams, and assignments. The conventional and blended learning modes of delivery were employed.

The structure of a single log file is illustrated by Table 1. The *"Regnum"* is the registration number, used as a unique identifier to track the path of a student throughout different courses and different years, i.e. it is used as "case id". The

"Timestamp" records the exact time of each event taken by the students, used to order the events. While the *"Event Name"* is used as the activity and the *"Event Context"* gives information about the concerned learning resource (file, assignment, folder, etc.) affected by the event. Finally, the *"Description"* explains the event in a more detailed manner.

Table 1: *Event Log Structure*

Regnum	Timestamp	Event Context	Event Name	Description

Data Preprocessing

The initial extracted logs included the actions taken by all users in the system (students, instructors, assisting instructors, manager, etc.). The goal is to transform the raw log data into abstract informative traces that will be then used as a source data to the process discovery step. Thus, two main steps were taken:

1. Anonymizing and cleaning the logs to keep only student logs using the roles given to users in the system.
2. Understanding Moodle events names to enrich the logs with semantics that serve the main goal

The initial number of event names was 65, including events related to course actions, quiz taking, assignments submissions, chats and discussions, profile viewing, and others. Only 14 events are kept, as shown in Table 2.

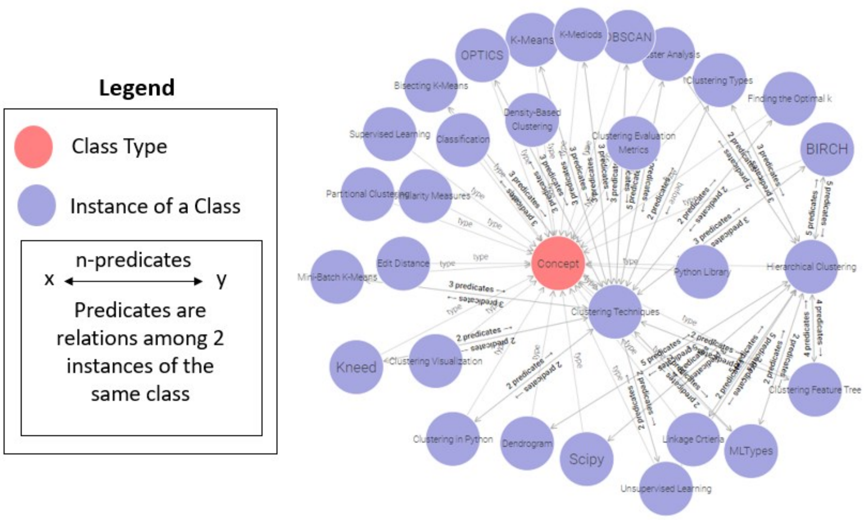
Table 2*: Chosen Event Names*

Event Name		
"A submission has been submitted."	"Quiz attempt submitted"	"Lesson started"
"Course activity completion updated"	"Course module viewed"	"Lesson resumed"
"Zip archive of folder downloaded"	"Content page viewed"	"Feedback viewed"
"Clicked join meeting button "	"Course summary viewed"	"Course viewed"
"Course module instance list viewed"	"Sessions viewed"	

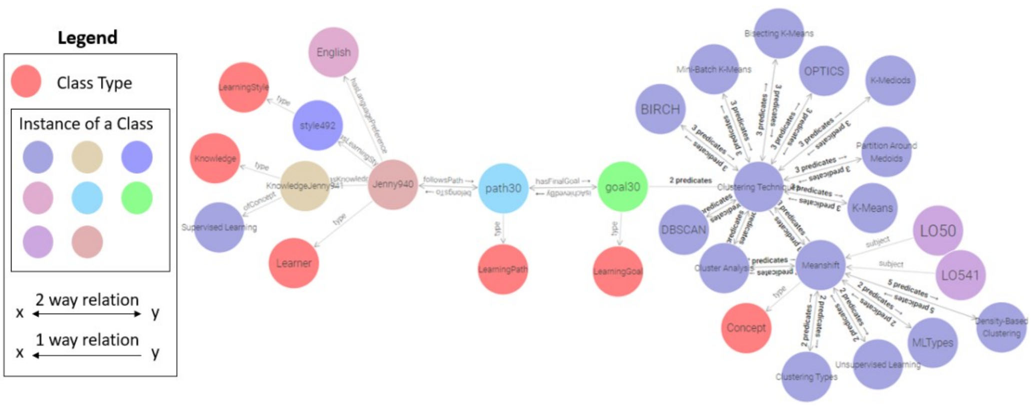
These events were chosen as they particularly show actions like completion of an assignment or a learning resource, assessment, taking feedback, studying, and exploring. After choosing the events to keep, the log is cleaned again to exclude other recorded events to finally end with 471 students with a total of 4291 traces for the period of 2018 until 2022, where a trace is an ordered sequence of events taken by a student in one course. The trace is ordered by the timestamp in an ascending chronological order. At this step, the logs are ready for transformation.

The Knowledge Base Related: the synthetic data

The knowledge base is constructed by initializing data on top of the ontological foundation. As an initial approach, we built a synthetic knowledge base that includes a domain knowledge, learning objects with different metadata fields and learners with different learning styles and learning goals. In the domain ontology, we created 90 connected concepts related to the machine learning domain, some of which are previewed in Picture 6 using GraphDB.



Picture 6: *Part of the Domain Ontology*



Picture 7: *Zoom into the Knowledge Base*

Picture 7, shows a complete example of the linked synthetic data from multiple ontologies in the knowledge base. From the Learner ontology, there is the learner "Jenny940" along with the data about her, including her pre-knowledge and learning style. Jenny is following a path (Learning Path ontology) that has a learning goal with an existing concept in the Domain ontology. For this concept, related concepts are shown with the linked learning objects for some, from the Learning Object ontology.

5.2. Next Phases

Trace transformation is a non-trivial endeavor, yet it is performed to alleviate informational value and significance of the raw traces (Ho et al. 2018; Settouti, 2009; Clauzel et al., 2011)). Following, the transformed traces undergo analysis, then a PM discovery one. The resulting process model is modeled within the Process Ontology framework and incorporated into the knowledge base to become a part of the framework and serve as an input source for the recommendation process. Finally, the RS is programmed to construct the adaptive learning path based on the current context of the learner, that encompasses information about his current knowledge level, past learning experience, preferences and learning style and his learning goal.

6. Conclusion and Future Perspectives

In the realm of e-learning, where a plethora of learning resources exists, the imperative to recommend suitable content becomes essential to amplify learner engagement and performance. In response, our paper proposes an ontology-based framework to recommend an adaptive personalized learning path, tailored to learners' goals, learning style and knowledge state. Yet, we intend to integrate the insights extracted from other past learners' experiences via PM. Within the scope of this study, we elucidate the ontological foundation that constitutes the framework's core, validated against a knowledge base through the deployment of synthesized data of learners, learning paths, learning goals, and learning objects in the chosen domain of Computer Science. As we embark on the PM facet, we have gathered and preprocessed Moodle logs of 471 students which stand primed for transformation.

References

Aggarwal, C. C. (2016). *Recommender Systems*. Springer International Publishing.

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process Mining Techniques for Analyzing Patterns and Strategies in Students' Self-Reg. Learning. *Metacognition and Learning*, *9*(2), 161–185.

Beheshti, A., Yakhchi, S., Mousaeirad, S., Ghafari, S. M., Goluguri, S. R., & Edrisi, M. A. (2020). Towards Cognitive Recommender Systems. *Algorithms*, *13*(8), 176.

Bey, A., & Champagnat, R. (2022). Analyzing Student Programming Paths using Clustering and Process Mining: *Proc. of the 14th International Conference on Computer Supported Education*, 76–84.

Bian, C.-L. (2019). Adaptive Learning Path Recommendation based on Graph Theory and an Improved Immune Algorithm. *KSII Transactions on Internet and Information Systems*, *13*(5).

Cheng, B., Zhang, Y., & Shi, D. (2018). Ontology-Based Personalized Learning Path Recommendation for Course Learning. *2018 9th Int. Conf. on Inf. Tech. in Medicine and Education (ITME)*, 531–535.

Clauzel, D., Sehaba, K., & Prié, Y. (2011). Enhancing synchronous collaboration by using interactive visualisation of modelled traces. *Simulation Modelling Practice and Theory*, *19*(1), 84–97.

Colace, F., De Santo, M., & Gaeta, M. (2009). Ontology for e-learning: A case study. *Interactive Technology and Smart Education*, *6*(1), 6–22. https:

Colace, F., Santo, M. D., & Greco, L. (2014). E-Learning and Personalized Learning Path: A Proposal Based on the Adaptive Educational Hypermedia System. *Int. Jour. of Emerging Tech. in Learning (IJET)*, *9*(2), 9–16.

Costa, N. T., José De Almeida, D., Oliveira, G. P., & Fernandes, M. A. (2022). Customized Pedagogical Recommendation Using Automated Planning for Sequencing Based on Bloom's Taxonomy: *International Journal of Distance Education Technologies*, *20*(1), 1–19.

Dolog, P., & Nejdl, W. (2003). Challenges and Benefits of the Semantic Web for User Modelling. *Int. Workshop on Adaptive Hypermedia and Adaptive Web Based Systems (AH 2003)*, 99–111.

Farance, F. (2000). *Draft standard for learning technology. Public and private information (PAPI) for learners (PAPI Learner)* (6.0). Inc. http://ltsc. ieee. org/wg2/papi_learner_07_main.

Fastiggi, W. (2019). *Applying Bloom's Taxonomy to the Classroom—Technology for Learners*. https://technologyforlearners.com/applying-blooms-taxonomy-to-the-classroom/

Felder, R. M., & Silverman, L. K. (1996). *Index of Learning Styles Questionnaire*. https://www.webtools.ncsu.edu/learningstyles/

Felder, R. M., & Silverman, L. K. (1988). Learning and Teaching Styles in Engineering Education. *Journal of Engineering Education*, *78*(7), 674–681.

Felder, R. M., & Soloman, B. A. (1993). *Learning Styles and Strategies*. 4.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220.

Ho, H. N., Rabah, M., Nowakowski, S., & Estraillier, P. (2018). Trace-Based Multi- Cristeria Preselection Approach for Decision Making in Interactive Applications like Video Games. In D. Kergel, B.

Heidkamp, P. K. Telléus, T. Rachwal, & S. Nowakowski (Eds.), *The Digital Turn in Higher Education* (pp. 211–234). Springer Fachmedien Wiesbaden.

Huber, S., Fietta, M., & Hof, S. (2015). Next step recommendation and prediction based on process mining in adaptive case management. *Proceedings of the 7th International Conference on Subject-Oriented Business Process Management*, 1–9.

IEEE Std 1484.12.1-2020, IEEE Standard for Learning Object Metadata. (2020). IEEE.

IMS. (2001). *Learner Information Package v1.0.1 | IMS Global Learning Consortium*. https://www.imsglobal.org/content/learner-information-package-v101

Juhaňák, L., Zounek, J., & Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Comp. in Human Behavior*, *92*, 496–506.

Kafa, A., & Eteokleous, N. (2022). Adapting to an Online Learning Environment in the Midst of the Global Pandemic: Insights from a Private Higher Institution in Cyprus. In H. Burgsteiner & G. Krammer (Eds.), *Impacts of COVID-19 Pandemic's Distance Learning on Students and Teachers in Schools and in Higher Education – International Perspectives.* (pp. 468–488). Leykam Buchverlag.

Mabroukeh, N. R., & Ezeife, C. I. (2010). A taxonomy of sequential pattern mining algorithms. *ACM Computing Surveys*, *43*(1), 1–41.

Nabizadeh, A. H., Leal, J. P., Rafsanjani, H. N., & Shah, R. R. (2020). Learning path personalization and recommendation methods: A survey of the state-of-the-art. *Exp. Sys. with Appli. 159*, 113596.

Nabizadeh, A. H., Mário Jorge, A., & Paulo Leal, J. (2017). RUTICO: Recommending Successful Learning Paths Under Time Constraints. *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*, 153–158.

Nafea, S. M., Siewe, F., & He, Y. (2019). A Novel Algorithm for Course Learning Object Recommendation Based on Student Learning Styles. *2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*, 192–201.

Nan Cenka, B. A. (2022). Analysing student behaviour in a learning management system using a process mining approach. *Knowledge Management & E-Learning: An Int. Journal*, *14*(1), 62–80.

Nitchot, A., Wettayaprasit, W., & Gilbert, L. (2019). Personalized learning system for visualizing knowledge structures and recommending study materials links. *E-Learning and Digital Media*, *16*(1), 77–91.

Pereira Detro, S., Santos, E. A. P., Panetto, H., Loures, E. D., Lezoche, M., & Cabral Moro Barra, C. (2020). Applying process mining and semantic reasoning for process model customisation in healthcare. *Enterprise Information Systems*, *14*(7), 983–1009.

Pireva, K., & Kefalas, P. (2018). A Recommender System Based on Hierarchical Clustering for Cloud e-Learning. In M. Ivanović, C. Bădică, J. Dix, Z. Jovanović, M. Malgeri, & M. Savić (Eds.), *Intelligent Distributed Computing XI* (Vol. 737, pp. 235–245). Springer International Publishing.

Raj, N. S., & Renumol, V. G. (2022). A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020. *Jour. of Comp. in Ed.*, *9*(1), 113–148.

Rohani, N., Gal, K., Gallagher, M., & Manataki, A. (2023). Discovering Students' Learning Strategies in a Visual Programming MOOC Through Process Mining Techniques. In M. Montali, A. Senderovich, & M. Weidlich (Eds.), *Process Mining Workshops* (Vol. 468, pp. 539–551). Springer Nature Switzerland.

Schonenberg, H., Weber, B., van Dongen, B., & van der Aalst, W. (2008). Supporting Flexible Processes through Recommendations Based on History. In M. Dumas, M. Reichert, & M.-C. Shan (Eds.), *Business Process Management* (Vol. 5240, pp. 51–66). Springer Berlin Heidelberg.

Settouti, L. S. (2009). A Trace-based system for TEL systems personalization. *2009 Ninth IEEE International Conference on Advanced Learning Technologies*. 2009 Ninth IEEE International Conference on Advanced Learning Technologies (ICALT), Riga, Latvia.

Settouti, L. S., Prie, Y., Champin, P.-A., Marty, J.-C., & Mille, A. (2009). *A Trace-Based Systems Framework: Models, Languages and Semantics*.

Shabatura, J. (2022). *Using Bloom's Taxonomy to Write Effective Learning Outcomes | Teaching Innovation and Pedagogical Support*. Using Bloom's Taxonomy to Write Effective Learning Outcomes. https://tips.uark.edu/using-blooms-taxonomy/

Singh, P. K., Pramanik, P. K. D., Dey, A. K., & Choudhury, P. (2021). Recommender systems: An overview, research trends, and future directions. *Recommender Systems*, *15*(1), 14–52.

Sudhana, Kalla. M., Raj, V. C., & Suresh, R. M. (2013). An ontology-based framework for context-aware adaptive e-learning system. *2013 Int.l Conf. on Computer Comm. and Informatics*, 1–6.

Tarus, J. K., Niu, Z., & Yousif, A. (2017). A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. *Future Generation Comp. Sys.*, *72*, 37–48.

Terragni, A., & Hassani, M. (2018). Analyzing Customer Journey with Process Mining: From Discovery to Recommendations. *2018 IEEE 6th Int. Conf. on Future IoT and Cloud (FiCloud)*, 224–229.

Terragni, A., & Hassani, M. (2019). Optimizing customer journey using process mining and sequence-aware recommendation. *Proc. of the 34th ACM/SIGAPP Symp. on Applied Computing*, 57–65.

Trabelsi, M., Suire, C., Morcos, J., & Champagnat, R. (2019). User's Behavior in Digital Libraries: Process Mining Exploration. In A. Doucet, A. Isaac, K. Golub, T. Aalberg, & A. Jatowt (Eds.), *Digital Libraries for Open Knowledge* (Vol. 11799, pp. 388–392). Springer International Publishing.

van der Aalst, W. (2016). *Process Mining*. Springer Berlin Heidelberg.

Wang, R., & Zaïane, O. R. (2018). Sequence-Based Approaches to Course Recommender Systems. In S. Hartmann, H. Ma, A. Hameurlain, G. Pernul, & R. R. Wagner (Eds.), *Database and Expert Systems Applications* (Vol. 11029, pp. 35–50). Springer International Publishing.

Wu, L., Liu, Q., Zhou, W., Mao, G., Huang, J., & Huang, H. (2020). A Semantic Web-Based Recommendation Framework of Educational Resources in E-Learning. *Technology, Knowledge and Learning*, *25*(4), 811–833.

Yang, S. (2019). *APPLIED PROCESS MINING, RECOMMENDATION, AND VISUAL ANALYTICS*.

Yang, S., Dong, X., Sun, L., Zhou, Y., Farneth, R. A., Xiong, H., Burd, R. S., & Marsic, I. (2017). A Data-driven Process Recommender Framework. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2111–2120.

Yarandi, M., Jahankhani, H., & Tawil, A.-R. H. (2013). *A personalized adaptive e-learning approach based on semantic web technology*. *10*(2).

Yari Eili, M., & Rezaeenour, J. (2022). A survey on recommendation in process mining. *Concurrency and Computation: Practice and Experience*, *34*(26).

Yu, Z., Nakamura, Y., Jang, S., Kajita, S., & Mase, K. (2007). Ontology-Based Semantic Recommendation for Context-Aware E-Learning. In J. Indulska, J. Ma, L. T. Yang, T. Ungerer, & J. Cao (Eds.), *Ubiquitous Intelligence and Computing* (Vol. 4611, pp. 898–907). Springer Berlin Heidelberg.