# Pedagogical AI-based Architecture for Encouraging Self-Regulated Learning Behavior in Students

**Geycy Dyany Oliveira Lima**

Federal University of Uberlândia
Faculty of Computer Science
Graduate Program in Computer Science

Uberlândia
2025

**Geycy Dyany Oliveira Lima**

# Pedagogical AI-based Architecture for Encouraging Self-Regulated Learning Behavior in Students

Ph.D. Thesis presented to the Graduate Program of the Faculty of Computer Science of the Federal University of Uberlândia as part of the requirements for obtaining the title of Doctor in Computer Science.

Area of concentration: Artificial Intelligence

Supervisor: Prof. Dr. Fabiano Azevedo Dorça
Co-supervisor: Prof. Dr. Rafael Dias Araújo

Uberlândia
2025

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

L732
2025

Lima, Geycy Dyany Oliveira, 1987-
Pedagogical AI-based architecture for encouraging self-regulated learning behavior in students [recurso eletrônico] / Geycy Dyany Oliveira Lima. - 2025.

Orientador: Fabiano Azevedo Dorça.
Coorientador: Rafael Dias Araújo.
Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-graduação em Ciência da Computação.
Modo de acesso: Internet.
DOI http://doi.org/10.14393/ufu.te.2025.709
Inclui bibliografia.
Inclui ilustrações.

1. Computação. I. Dorça, Fabiano Azevedo,1979-, (Orient.). II. Araújo, Rafael Dias,1986-, (Coorient.). III. Universidade Federal de Uberlândia. Pós-graduação em Ciência da Computação. IV. Título.

CDU: 681.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:
Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074

**UNIVERSIDADE FEDERAL DE UBERLÂNDIA**
Coordenação do Programa de Pós-Graduação em Computação
Av. João Naves de Ávila, 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpgfacom@ufu.br

# ATA DE DEFESA - PÓS-GRADUAÇÃO

| | | | | | |
|---|---|---|---|---|---|
| Programa de Pós-Graduação em: | Ciência da Computação | | | | |
| Defesa de: | Tese, 47/2025, PPGCO | | | | |
| Data: | 11 de Dezembro de 2025 | Hora de início: | 14:00 | Hora de encerramento: | 17:00 |
| Matrícula do Discente: | 12013CCP002 | | | | |
| Nome do Discente: | Geycy Dyany Oliveira Lima | | | | |
| Título do Trabalho: | Pedagogical AI-based Architecture for Encouraging Self-Regulated Learning Behavior in Students | | | | |
| Área de concentração: | Ciência da Computação | | | | |
| Linha de pesquisa: | Inteligência Artificial | | | | |
| Projeto de Pesquisa de vinculação: | ---------------- | | | | |

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Rafael Dias Araújo (Coorientador) - FACOM/UFU, Renan Gonçalves Cattelan - FACOM/UFU, Claudiney Ramos Tinoco - FACOM/UFU, Crediné Silva Menezes - FACED/UFRGS, Andrey Ricardo Pimentel - DInf/UFPR e Fabiano Azevedo Dorça - FACOM/UFU, orientador(a) do(a) candidato(a).

Os examinadores participaram desde as seguintes localidades: Claudiney Ramos Tinoco - Monte Carmelo/MG, Crediné Silva Menezes - Porto Alegre/RS, Andrey Ricardo Pimentel - Curitiba/PR. Os outros membros da banca e o aluno(a) participaram da cidade de Uberlândia.

Iniciando os trabalhos, o(a) presidente da mesa, Prof. Dr. Fabiano Azevedo Dorça, apresentou a Comissão Examinadora e o(a) candidato(a), agradeceu a presença do público, e concedeu ao(à) Discente a palavra para a exposição do seu trabalho. A duração da apresentação do(a) Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir, o(a) senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

**Aprovada**

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.

Documento assinado eletronicamente por **Andrey Ricardo Pimentel**, **Usuário Externo**, em 12/12/2025, às 18:25, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Claudiney Ramos Tinoco**, **Professor(a) do Magistério Superior**, em 13/12/2025, às 02:50, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Rafael Dias Araújo**, **Professor(a) do Magistério Superior**, em 13/12/2025, às 07:55, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Renan Gonçalves Cattelan**, **Professor(a) do Magistério Superior**, em 13/12/2025, às 10:14, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Fabiano Azevedo Dorça**, **Professor(a) do Magistério Superior**, em 15/12/2025, às 10:26, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Crediné Silva de Menezes**, **Usuário Externo**, em 26/12/2025, às 11:38, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **6894084** e o código CRC **E3928E40**.

**Referência:** Processo nº 23117.084753/2025-14 SEI nº 6894084

*À minha família, especialmente meus amados sobrinhos Pedro, Matheus e Isabella.*

# Acknowledgments

A jornada que culmina neste trabalho foi marcada por desafios, descobertas e intenso aprendizado. Nada disso, porém, teria sido possível sem o apoio, a confiança e o incentivo de pessoas especiais, às quais expresso minha mais sincera gratidão.

Agradeço, primeiramente, a Deus, pela força, luz e perseverança em todos os momentos desta caminhada. Aos meus familiares, pela presença constante, pelo amor incondicional e por acreditarem na minha capacidade, mesmo quando o caminho parecia difícil. Um agradecimento especial aos meus sobrinhos: Pedro, Matheus e Isabella, que com sua alegria e espontaneidade, tornaram mais leves os dias mais desafiadores.

Registro minha profunda gratidão ao meu orientador, Fabiano Dorça, pela paciência, pelas valiosas orientações e pelo exemplo de dedicação e profissionalismo. Mais do que um orientador, agradeço ao ser humano generoso, ético e inspirador que ele é, sempre disposto a ajudar, ensinar e acolher. Sua experiência e comprometimento foram fundamentais para o desenvolvimento deste trabalho e para meu crescimento acadêmico e pessoal.

Ao meu coorientador, Rafael Dias Araújo, deixo meu sincero e eterno agradecimento. Foi ele quem me ofereceu o incentivo inicial e acreditou no potencial deste trabalho desde o princípio. Profissional exemplar, dedicado e generoso, é uma das pessoas mais importantes nesta conquista. Suas contribuições foram essenciais em cada etapa da pesquisa.

Agradeço também à minha colega e amiga Juliete, que esteve ao meu lado em diversas fases deste projeto. Sua parceria, apoio, conversas, ideias e colaboração fizeram toda diferença ao longo dessa trajetória.

À banca de qualificação e aos avaliadores da defesa, agradeço pela disponibilidade, pelas observações rigorosas, críticas construtivas e sugestões que enriqueceram a qualidade científica deste trabalho.

Estendo meu agradecimento ao IFSULDEMINAS – Campus Inconfidentes e ao IFNMG – Campus Salinas, pelo apoio institucional e pela oportunidade de conciliar vida acadêmica, profissional e pesquisa científica. À equipe do Programa de Pós-Graduação e à FACOM, pelo ambiente acolhedor e pela dedicação em promover a pesquisa, a formação científica e o desenvolvimento pessoal. Agradeço também à CAPES pelo apoio que viabiliza a continuidade e fortalecimento da pesquisa no país.

A todos que contribuíram direta ou indiretamente para esta conquista, deixo meu mais sincero agradecimento.

# Abstract

With the advancement of educational technologies, Virtual Learning Environments have become essential for promoting new teaching and learning methods, especially in distance education and hybrid contexts. These environments allow students to access content, complete activities, and interact with peers and instructors in a flexible and personalized manner. In this scenario, Self-Regulated Learning stands out as a key competency, as it enables learners to autonomously manage, monitor, and direct their own learning process. This study proposes and validates an Artificial Intelligence - supported Pedagogical Architecture to foster SRL in VLEs, aiming to enhance students' autonomy and engagement. Initially, a systematic literature review was conducted, which identified research gaps and guided the PA design. Subsequently, Proofs of Concept were carried out using data from the Open University Learning Analytics Dataset and from Moodle at IFSULDEMINAS – Campus Carmo de Minas, applying Educational Data Mining techniques and clustering algorithms. These analyzes allowed the identification of behavioral patterns, SRL profiles, and significant correlations between engagement and academic performance. In the final stage, the PA was implemented and evaluated in the context of an online Introduction to Python Programming course. Among the resources integrated into the VLE, the Time Tracker SRL plugin stands out, developed to monitor the time dedicated to learning activities and provide automated feedback. Other plugins, such as Configure Reports, Completion Progress, Analytics Graphs, and OpenAI Chat, were also employed to support the self-regulation process. The results showed that the PA had a significant impact in promoting SRL, with a positive correlation between engagement and academic performance. The triangulation of evidence—based on VLE log analysis, self-regulation questionnaires, and focus group interviews—confirmed the effectiveness of the PA, validating its potential to develop SRL skills, foster autonomy, and improve student performance. Thus, the proposed approach constitutes an innovative, scalable, and adaptable solution to support and personalize learning in VLEs.

**Keywords:** Virtual Learning Environments. Educational Data Mining. Self-Regulated Learning. Pedagogical Architecture.

# Resumo


Com o avanço das tecnologias educacionais, os Ambientes Virtuais de Aprendizagem tornaram-se essenciais para a promoção de novos métodos de ensino e aprendizagem, especialmente na educação a distância e em contextos híbridos. Esses ambientes possibilitam que os estudantes acessem conteúdos, realizem atividades e interajam com colegas e professores de forma flexível e personalizada. Nesse cenário, a Aprendizagem Autorregulada, do Inglês *Self-Regulated Learning*, destaca-se como competência-chave, pois capacita os estudantes a gerenciar, monitorar e direcionar autonomamente seu próprio processo de aprendizagem. Este estudo propõe e valida uma Arquitetura Pedagógica apoiada por Inteligência Artificial para fomentar a SRL em AVAs, visando ampliar a autonomia e o engajamento dos estudantes. Inicialmente, foi conduzida uma revisão sistemática da literatura, que identificou lacunas de pesquisa e orientou a concepção da AP. Em seguida, realizaram-se Provas de Conceito com dados do *Open University Learning Analytics Dataset* e do Moodle do IFSULDEMINAS – Campus Carmo de Minas, aplicando técnicas de Mineração de Dados Educacionais e algoritmos de agrupamento. Essas análises permitiram identificar padrões de comportamento, perfis de SRL e correlações significativas entre engajamento e desempenho acadêmico. Na etapa final, a AP foi implementada e avaliada no contexto de um curso online de Introdução à Programação em Python. Entre os recursos integrados ao AVA, destaca-se o plugin *Time Tracker* SRL, desenvolvido para monitorar o tempo dedicado às atividades avaliativas e fornecer feedback automatizado. Outros plugins, como *Completion Progress*, *Analytics Graphs* e OpenAI Chat, também foram utilizados para apoiar o processo de autorregulação. Os resultados mostraram que a AP teve impacto significativo na promoção da SRL, com correlação positiva entre engajamento e desempenho acadêmico. A triangulação das evidências — baseada na análise dos logs do AVA, nos questionários de autorregulação e nas entrevistas com grupos focais — comprovou a eficácia da AP, validando seu potencial para desenvolver habilidades de SRL, promover autonomia e melhorar o desempenho dos estudantes. Assim, a proposta configura-se como uma solução inovadora, escalável e adaptável para apoiar e personalizar a aprendizagem em AVAs.

**Palavras-chave:** Ambientes Virtuais de Aprendizagem. Mineração de Dados Educacionais. Aprendizagem Autorregulada. Arquitetura Pedagógica.

# List of Figures

# List of Tables

# Acronyms list

**AI** Artificial Intelligence

**ANOVA** Analysis Of Variance

**DBSCAN** Density-Based Algorithm

**DM** Data Mining

**EDM** Educational Data Mining

**EM** Expectation Maximization

**GNU** General Public License

**HDBSCAN** Density-Based Clustering Based on Hierarchical Density

**HTML** HyperText Markup Language

**IFSULDEMINAS** Instituto Federal do Sul de Minas Gerais

**KDD** Knowledge Discovery in Databases

**LA** Learning Analytics

**LAD** Learning Analytics Dashboards

**MSLQ** Motivated Strategies for Learning Questionnaire

**MST** Minimum Spanning Tree

**OLM** Open Learner Model

**OLSQ** Online Self-Regulated Learning Questionnaire

**OULAD** Open University Learning Analytics Dataset

**PA** Pedagogical Architecture

**PoC** Proof of Concept

**SDGs** Sustainable Development Goals

**SDT** Self-Determination Theory

**SRL** Self-Regulated Learning

**ULE** Ubiquitous Learning Environment

**UML** Unified Modeling Language

**VLE** Virtual Learning Environment

# Contents

# Chapter 1

# Introduction

The increasing development of technological resources in intelligent learning environments has caused changes in the teaching and learning processes. Virtual Learning Environment (VLE) apply different resources to support educational processes, which are often implemented to improve academic performance and students' motivation, and to reduce dropout. VLEs are online computational systems used for educational purposes in different domains and levels.

VLE is a concept that emerged in the mid-1990s, right after the popularization of the Internet (O'LEARY; RAMSDEN, 2002). Such environments can be understood as Information Systems specific to the educational context that provide administrative and didactic support to the teaching and learning process by means of technological resources through the Internet (MUELLER; STROHMEIER, 2011). These resources can be diverse and have different natures. Furthermore, they generally include shared spaces for the distribution of educational materials and resources, communication between students and teachers/tutors, and mechanisms for evaluation, management, and monitoring of students (DILLENBOURG; SCHNEIDER; SYNTETA, 2002). In addition, a VLE can be used in different contexts, in formal or informal education, and in different modalities, whether face-to-face, distance, or hybrid.

Between 2020 and 2022 the COVID-19 pandemic (World Health Organization, 2020) has brought many significant changes to education around the world. Educational institutions have been forced to reinvent their teaching process, using VLEs to keep learning going while reducing the risk of exposure to the virus. With this, there has been a generation of many new data sets, which can be used to analyze the learning process of students.

Although distance learning has been an important solution to maintaining education during the pandemic, it has also presented significant challenges, especially for those who have learning difficulties. In addition, the lack of face-to-face interaction leads to emotional and social problems for students and teachers, which can negatively affect emotional well-being and motivation.

The datasets generated in VLEs can be used to understand how to improve learning techniques, such as Self-Regulation Learning (SRL) (COMAN *et al.*, 2020). Machine learning has great potential to exploit educational data collected from VLEs. By analyzing this kind of data, one can detect student profiles and provide interventions to improve students' performance. In online learning environments, students' behavior can be assessed from student interactions on the system, such as page views, access to materials, quiz solving, forum participation, number of clicks, activity submission deadlines, and others (COSTA; DORÇA; ARAÚJO, 2020).

Studies have consistently shown that self-regulation of the learning process is strongly associated with better academic performance (ZIMMERMAN; MARTINEZ-PONS, 1986). In SRL, students play an active role in their learning, developing cognitive, metacognitive, motivational, and affective strategies to regulate their progress (LIMA; ARAÚJO; DORÇA, 2020). Learners with higher levels of SRL skills are more capable of setting goals, monitoring their strategies, and achieving academic success (PINTRICH; GROOT, 1990).

Traditionally, the classification of SRL profiles in online learning has relied on data collected through self-report tools (BROADBENT; FULLER-TYSZKIEWICZ, 2018; YOT-DOMÍNGUEZ; MARCELO, 2017). Although easy to implement, these instruments present limitations, as students may not accurately recall the strategies they employ, as the questionnaires are usually applied before or after the learning process. Such constraints have led researchers to explore alternative or complementary methods that allow for a more comprehensive and reliable understanding of students' self-regulation processes.

Based on this perspective, Lima *et al.* (2024) conducted a literature review and synthesized the main technologies applied in the context of SRL. Figure 1 illustrates this mapping through a word cloud generated from the abstracts of the analyzed papers. The highlighted terms reveal the prominence of concepts such as Open Learner Model (OLM), *visualization*, *analytics*, *recommendation*, *personalized*, *adaptive*, *interactive*, *social*, *collaborative*, and *agents*. These keywords not only indicate the variety of technological approaches explored, but also demonstrate their direct association with SRL subprocesses, reinforcing the growing role of technology in supporting autonomous learning.

In this context, Educational Data Mining (EDM) techniques have emerged as a promising alternative, since they make use of real data collected directly from VLEs, such as log interactions, resource usage, and task submissions. By analyzing these behavioral traces, EDM enables more accurate measurement and profiling of students compared to self-report tools, thus offering a viable and reliable approach to detect learning behaviors and support the development of SRL.

Beyond these methodological and technological advances, it is important to emphasize that research on SRL in VLEs is not limited to detecting behavioral patterns, but also to understanding how these patterns can inform pedagogical practices and support tools. Integrating EDM with pedagogical architectures provides a promising pathway to


Figure 1 – Wordcloud considering selected papers' abstracts.
*Source:* Extracted from Lima *et al.* (2024)

designing adaptive interventions, offering students timely guidance and resources tailored to their needs. This intersection of Learning Analytics (LA), Artificial Intelligence (AI), and educational design highlights the growing relevance of interdisciplinary approaches to foster autonomy, engagement, and persistence in digital learning contexts.

## 1.1 Motivation

The motivation of this work arises from the need to understand and enhance SRL techniques in VLEs. With the rapid development of technological resources in intelligent learning environments, it has become significantly crucial to explore how these technologies can be used effectively to improve teaching and learning processes (LIMA; ARAÚJO; DORÇA, 2020).

VLEs have shown promise in collecting comprehensive datasets on student behavior (COSTA; DORÇA; ARAÚJO, 2020). These data can be leveraged to gain a better understanding of learning processes, identify patterns of student behavior, and provide personalized interventions that improve academic performance. However, it is important to go beyond student self-reporting tools and explore EDM techniques to obtain more precise insights into students' SRL.

SRL is a widely used and important theory to help the student's learning process. After several studies, we realized the need to develop a Pedagogical Architecture (PA) in a VLE that can help students learn more effectively and independently. This work can have a positive impact on the academic performance of many students. Although Moodle was utilized as the primary testing environment in this study due to its availability and flexibility, the proposed PA is designed to be compatible with any VLE that supports similar functionalities. This ensures that the architecture can be adapted to various contexts and platforms, providing a versatile framework to foster SRL.

## 1.2 Research Questions and Goals

This study is guided by the following research questions:

RQ1: How does the PA, through its various elements and resources, support and facilitate the development of SRL skills in the VLE?

This research question seeks to understand how the intentional design of PA can facilitate and enrich the student learning process, focusing specifically on the development of SRL skills. SRL is crucial for academic success as it empowers students to set goals, plan, monitor, and evaluate their own learning independently, skills that are essential both in academic contexts and in developing life competencies.

Although the existing literature provides various examples of how technological resources can be used in isolation to encourage SRL behaviors, there is a significant gap in understanding how the integration of these resources into a cohesive PA can substantially and consistently amplify the development of these skills in VLEs. This gap highlights the need for more in-depth investigation into how a holistic approach, combining multiple resources in an integrated manner, can further enhance students' ability to self-regulate.

Exploring this question is crucial, as a well-structured online environment not only provides continuous and adaptive support but also promotes educational autonomy and improves student engagement and performance. By understanding how different elements and resources within a PA contribute to SRL, we can identify the most effective strategies and develop innovative educational practices that can be broadly applied in digital contexts.

RQ2: How does student engagement with PA resources correlate with their academic performance?

This research question aims to investigate how student engagement with PA resources correlates with their academic performance. Student engagement is widely recognized in the literature as a critical factor for academic success, directly influencing motivation, knowledge retention, and students' ability to apply what they have learned in practical contexts.

While there are studies that suggest a positive correlation between engagement with learning resources and academic performance, a deeper understanding of how different types of resources within a PA influence this relationship is needed. This research seeks to fill this gap by analyzing how the use and engagement of students with specific PA resources affect their academic performance, measured through metrics such as grades, course completion rates, and participation in activities.

Understanding this correlation is crucial for developing more effective pedagogical strategies and adapting educational resources to maximize student engagement and academic success. Additionally, the findings from this investigation can therefore also provide valuable insights for educators and educational technology developers, helping them create more student-centered learning environments that effectively meet individual needs.

RQ3: How do the interpretations of the results obtained from the analyses validate the proposed PA?

To address this research question, a multi-method analytical approach was adopted, combining three complementary sources of evidence: (i) EDM analyses of interaction logs, (ii) self-regulation questionnaires applied before and after the use of the PA resources, and (iii) focus group interviews with students. Together, these methods enable a comprehensive assessment of students' self-regulated learning behaviors and the effectiveness of the proposed PA.

EDM analyses provide deep insights into the patterns of student behavior in a VLE, allowing the identification of SRL profiles. With these techniques, it is possible to detect which behaviors and interactions with the PA resources are associated with better academic outcomes and the development of SRL skills.

The self-regulation questionnaires, administered before and after the use of PA resources, offer a direct measure of the evolution of students' SRL skills. These quantitative data provide a solid foundation for analyzing changes in self-regulatory competencies over time, helping to identify whether the resources and personalized interventions of the PA are indeed effective in promoting self-regulation.

Focus group interviews complement these analyses with qualitative data, providing a more detailed view of students' perceptions regarding their learning experience and satisfaction with the resources offered by the PA. Through these interviews, it is possible to explore more deeply the challenges encountered, the perceived benefits, and the impact of pedagogical interventions on student behavior and engagement. Therefore, the combination of these three methods — EDM, questionnaires, and interviews — offers a comprehensive view of the effectiveness of the PA.

The **main goal** of this study is to develop a Pedagogical Architecture (PA) supported by AI to foster SRL in students. This PA should include resources that help students set goals, plan, and monitor their progress. The specific goals of this project are:

- Provide a comprehensive literature review on technologies that facilitate SRL in VLEs, with the aim of identifying and analyzing key technological resources and methodologies that can be integrated into the proposed PA to promote self-regulation among students;
- Evaluate the correlation between student engagement with the resources of the PA, their SRL skills, and their academic performance, using EDM techniques and clustering algorithms to identify behavior patterns and SRL profiles;
- Develop and implement the Time Tracker SRL plugin for VLE, a time management tool that allows detailed monitoring and analysis of the time students dedicate to learning activities such as quizzes, assignments, and forums. The plugin also offers features for teachers to plan the time allocated for each activity, providing guidance for students' time planning and management. The goal is to assess the effectiveness of this plugin in providing detailed feedback on the time spent by students, aiming to systematically enhance their self-regulation skills and optimize their time management more efficiently.
- Evaluate the effectiveness of the PA in developing students' SRL skills by conducting empirical studies measuring the impact of the PA on specific skills such as planning, monitoring, and assessment, as well as on students' overall academic performance, through self-regulation questionnaires, focus group interviews, and performance analyses;

## 1.3 Hypothesis

The hypotheses of this study are formulated based on the assumption that the proposed AI-supported PA will have a significant positive impact on students' self-regulation skills and engagement. The specific hypotheses are as follows:

H1: The implementation of an AI-supported PA will result in a statistically significant improvement in students' self-regulation skills. This hypothesis is based on the premise that structured self-regulation support tools, when effectively integrated into a learning environment, improve students' abilities to set goals, plan, monitor, and evaluate their progress, thereby fostering greater SRL behaviors;

H2: Engagement with a PA will be positively correlated with enhanced academic performance among students. The correlation between the engagement with PA resources and academic performance will be analyzed to assess the effectiveness of the PA in contributing to better educational results;

H3: The proposed PA is validated as effective in developing students' SRL skills and in identifying SRL profiles in VLEs, based on EDM analyses, self-regulation questionnaires, and focus group interviews.

Development of an AI-based PA to encourage SRL in the VLE;

## 1.4 Impacts of Research on Society

The main objective of this research is to develop and validate an AI enhanced learning approach that contributes to improving definable metrics of learning and equity in VLEs. It is expected that, if replicated and scaled, the solution can support policies and practices that promote more inclusive and higher-quality education in contexts similar to those studied. In an increasingly globalized world, it is vital to tackle the urgent issues of sustainability, education, and equality by extending the reach of academic influence beyond traditional boundaries, thereby impacting society on a broader and more significant scale.

In this context, research, focusing on educational technologies and pedagogy, goes beyond the mere expansion of scientific knowledge. It is designed to contribute significantly to the achievement of several Sustainable Development Goals (SDGs) set by the United Nations (WENTROBA; VOGT; BOTELHO, 2023). This research aims to incorporate these global objectives, not only promoting the academic field but also creating lasting positive impacts for society in general, demonstrating a commitment to quality education and social equity.

**Quality Education (SDG 4):** This research contributes directly to SDG 4 (NATIONS, 2023), which seeks to ensure inclusive, equitable and quality education, and promote lifelong learning opportunities for all. By developing a PA that supports SRL in virtual environments, this research facilitates access to quality education.

**Decent Work and Economic Growth (SDG 8):** Training provided by a self-regulated learning environment prepares students with the skills needed for the modern job market, including self-management, critical thinking, and problem solving. This directly contributes to the creation of a more qualified workforce, driving sustainable economic growth (NATIONS, 2023).

**Reduced Inequalities (SDG 10):** By providing more accessible and adaptable educational resources, research contributes to reducing educational inequalities (NATIONS, 2023). It promotes inclusion by enabling students of diverse backgrounds and with different learning styles to have access to educational materials and methods that meet their specific needs.

## 1.5 Outline

This thesis is organized as follows: Chapter 2 presents a comprehensive discussion of the theoretical foundations underlying the research, covering PA, SRL, LA, EDM, and clustering techniques. Chapter 3 reviews related works, with emphasis on previous studies and technological tools associated with SRL and VLEs, highlighting investigations that address PA. This chapter outlines the main objectives, technologies, and results achieved by such studies, as well as the existing gaps, which underscore the originality of this research in proposing a new AI-based PA aimed at fostering SRL. Chapter 4 describes the methodology adopted in the development of the research, detailing the **PoCs!** (**PoCs!**) conducted throughout the study. Chapter 5 presents the Proposed Approach, explaining the development process of the PA, with emphasis on the implemented resources and strategies. Subsequently, Chapter 6 reports the results obtained and their discussion. Finally, Chapter 7 brings together the conclusions of the study, the identified limitations, the publications produced, and the directions for future work.

# Chapter 2

# Theoretical Foundations

This chapter presents a literature review on the main concepts that underpin the proposed approach. The key concepts described are: PA, SRL, LA, EDM, and Data Clustering.

## 2.1 Pedagogical Architecture

PA are approaches that integrate pedagogical and technological aspects to create innovative educational proposals (MENEZES; JUNIOR; ARAGÓN, 2020). They are based on the articulation between the constructivist conception of learning and the pedagogy of questioning, focusing on solving real problems, transforming information into knowledge, encouraging authorship and cooperation, and promoting metacognition. These architectures require an actively engaged stance from students, involving research, interactive activities, and the use of digital technologies, with the teacher playing a crucial role in guiding and creating a collaborative and reflective learning environment (SILVA; MENEZES; JUNIOR, 2023a).

When compared to traditional educational models, PA provide a more dynamic and interactive learning environment. Traditional models often rely on passive learning and rote memorization, whereas PA emphasize active participation and the application of knowledge in real-world contexts. This shift requires a fundamental change in both teaching practices and student attitudes towards learning.

Expanding on the necessity of an interactive and reflective learning environment, the authors Biancardi *et al.* (2021) highlight that the primary goal of a PA is to transform the teaching and learning environment, promoting a pedagogy that values creativity, reflection, critique, and innovation. In the same work, they assert that a PA aims to create a cognitive ecosystem where technology is not merely a support but a constitutive element of new relationships and ways of thinking.

In line with these transformative objectives, it is crucial to consider how PA can effectively integrate various pedagogical strategies and virtual environments to enhance the overall learning experience. By leveraging technology as a foundational element, PA can facilitate deeper collaboration and engagement among students and educators. This integration not only supports the development of critical thinking and problem-solving skills but also fosters a sense of community and shared purpose in the learning process.

In accordance with Behar *et al.* (2020), PA can be understood as a system of theoretical premises that underpins, explains, and directs the curriculum approach. This system materializes in pedagogical practices and the interactions between the teacher, the student, and the object of study or knowledge.

According to the authors Menezes (2009), PA act as essential frameworks for learning, composed of various components including pedagogical approaches, software, conceptions of space and time, and artificial intelligence, among other elements. Learning is understood as a construction based on experiences and reflections, as well as meta-reflections, derived from the individual's interaction with the socio-ecological environment.

Recent studies also reinforce the potential of PA to enhance active and meaningful learning. For instance, Santos *et al.* (2025) analyze the integration of digital technologies in educational contexts and highlight that pedagogical architectures should be conceived as flexible frameworks that adapt to students' needs and contexts, rather than rigid models. The authors emphasize that innovation emerges from the articulation between technological resources, collaborative practices, and reflective processes, which together foster student autonomy and critical engagement. This perspective aligns with the central principles of this research, in which the proposed PA seeks to combine technological tools with pedagogical strategies to support SRL and collaborative knowledge construction.

This study is based on the theoretical conceptions of Behar *et al.* (2020), Ribeiro (2019), and Sonego (2019), who propose the formation of a PA based on four main aspects. These aspects will be detailed as follows:

- **Organizational:** These elements serve as the cornerstone for the educational blackprint, encompassing the objectives of remote education, the structuring of time and space, and the anticipated outcomes for participant engagement, along with the responsibilities and privileges assigned to each individual;

- **Content:** The theme in focus and the resources employed, including learning objects, software, and a variety of educational tools such as quizzes, animations, and more form integral parts of this educational framework;

- **Methodological:** Concern the ways in which technologies are applied and how educational material is evolved over the course of the teaching and learning experience. This includes the types of activities conducted, methods of interaction and communication, assessment protocols, and the structured organization of these components into a coherent instructional sequence;

- **Technological:** These elements pertain to the digital resouces and tools utilized in the online course, including the VLE, its features, and communication tools like video and teleconferencing, among other technological resources.

While the aspects of PA often appear to be choices and delineations made by the teacher, it is essential to recognize that, in many cases, they are predetermined by the educational institutions that offer the courses (BEHAR *et al.*, 2020). In these situations, the institutional platform, the content to be covered, and even the methodology to be adopted are specified. Thus, there is a predefined PA, intended for different teachers who teach a specific educational activity.

Even with these restrictions, it is possible to customize and adjust a PA, taking into account social and affective aspects, to promote greater interaction among students (RIBEIRO, 2019). Moreover, the integration of elements that foster SRL can be particularly beneficial. The adaptation of a predefined PA can be carried out by making changes to one or more aspects, according to the specific objectives established. Incorporating strategies that encourage students to set goals, monitor their progress, and reflect on their learning can enhance the effectiveness of the PA, promoting greater autonomy and responsibility for the learning process.

A PA demands an active stance from students (SILVA; MENEZES; JUNIOR, 2023a), a key characteristic of SRL theory. SRL plays a central role in dynamic and interactive educational environments, such as those promoted by pedagogical architectures, by fostering greater autonomy, self-reflection, motivation, and responsibility in the learning process (ZIMMERMAN, 1986). Additionally, a PA should facilitate collaboration among students and promote the development of critical thinking by providing an environment that encourages the exchange of ideas and joint problem-solving (BIANCARDI *et al.*, 2021). In this context, SRL enables students to manage their activities more effectively, setting clear goals, planning strategies, and continuously evaluating their outcomes. Next, the core concepts of SRL will be discussed for the implementation of a PA aimed at developing more autonomous and collaborative learning skills.

## 2.2 Self-Regulated Learning

SRL is a research area of Educational Psychology that aims to study and analyze personal aspects of students that influence their self-guided learning process. SRL is a conceptual framework for understanding the cognitive, meta-cognitive, behavioral, motivational, and emotional/affective aspects of learning (PANADERO, 2017). In competitive and evaluative contexts, human achievements depend very much on the individual's ability to self-regulation (ZIMMERMAN; MARTINEZ-PONS, 1986).

However, employing SRL strategies and competencies is a complex process, as it involves the development of skills related to engagement, self-monitoring, self-assessment, perceived competencies, and contextual understanding (GARCIA; FALKNER; VIVIAN, 2018). In the work of Zimmerman and Martinez-Pons (1986), fourteen categories of self-regulatory strategies were proposed, in addition to an extra category termed Other, used to represent behaviors that are not self-regulated. Table 1 presents these categories along with their descriptions.

Table 1 – Self-Regulated Learning strategies.

| **Strategy categories** | **Definition** |
|---|---|
| Self-assessment | Assessing the quality or progress of student-initiated work. |
| Organization and transformation | Students rearrange materials to improve their learning. |
| Set of objectives and planning | Students establish a set of educational goals and sub-goals as well as their planning for completing the activities. |
| Information search | Searching for information in different media in order to perform a task. |
| Record keeping and monitoring | Records of events or results. |
| Structuring of the environment | Organization of the learning environment in order to improve performance. |
| Self-consequence | Punishment or praise for performing tasks. |
| Listen again and memorize | Memorizing the studied material through practice. |
| Seeking social assistance (peers) | Request help from colleagues. |
| Seeking social assistance (teachers) | Request help from the teacher. |
| Seeking social assistance (adults) | Request help from adults (family). |
| Testing review | Review tests. |
| Annotation review | Review annotations. |
| Textbook review | Review textbooks used during the learning process. |
| Other | Statements indicating learning behavior initiated by others, for example, parents or teachers. |

*Source:* Zimmerman and Martinez-Pons (1986)

The taxonomy proposed by the authors was developed based on interviews with high school students and considered the application of SRL strategies across different learning contexts, including face-to-face and non-face-to-face settings. Students were asked to report the methods they used to participate in classes, study, and complete academic tasks. The strategies defined by the authors have since been widely discussed and adopted in several self-regulated learning studies and surveys. The findings indicate a clear relationship between self-regulated learning and academic performance, showing that students with higher academic achievement employ a greater number of SRL strategies than those with lower performance.

In the review carried out by Panadero (2017), six models of SRL were presented and compared: Zimmerman (1986); Boekaerts (1988); Winne and Hadwin (1998); Pintrich and Groot (1990); Efklides (2011); and Hadwin, Järvelä and Miller (2011). According to Panadero (2017) and Puustinen and Pulkkinen (2001), SRL models can be defined as cyclical, and they have different phases and sub-processes of self-regulation. SRL phases and some sub-processes are described in Table 2. Although the models present different nomenclatures for the processes, their understanding allows them to be grouped into three major phases: a) Preparatory (or planning); b) Execution; and, c) Assessment.

Table 2 – Definition of phases and sub-processes.

| **Phases** | **Sub-processes** |
|---|---|
| Preparatory | Forecasting, task analysis, definition of objectives and goals. |
| Execution | Performance, monitoring, implementation of strategies. |
| Assessment | Feedback, regulation, adaptation and self-reflection. |

*Source:* Panadero (2017) and Puustinen and Pulkkinen (2001)

The preparatory phase comprises the analysis of tasks, the planning, the definition of objectives, and the establishment of goals (PANADERO, 2017). In this phase, one can use, among other technologies, administrative tools such as a calendar so that the student can plan the course development. Considering the preparatory phase, Kitsantas (2013) mentions two technologies that can be used: blogs/online newspapers and podcasts.

Blogs/newspapers allow students to provide and receive feedback from colleagues about the contents and to prepare a study guide. Because it is an open technology, students can post questions, interact, and create a collaborative environment. Podcasts are multimedia resources that students access at any time. Therefore, the recordings of study groups can be used to outline the student's learning objectives. In this phase, students who use SRL strategies demonstrate more self-efficacy, a greater expectation of results, and interest in tasks than other students (KITSANTAS, 2013).

Sharp and Sharp (2016) assert that the planning phase of SRL is characterized by students setting clear learning goals and effectively organizing their study time to meet these objectives. Tzimas and Demetriadis (2024) underscore the significance of metacognitive strategies, such as goal-setting, in equipping students for upcoming tasks. These strategies enable learners to anticipate potential challenges and devise structured plans to overcome them, a process that is especially advantageous in online learning environments where students are required to navigate their studies with greater autonomy. Both studies underscore the pivotal role of the planning phase in fostering learner independence and enhancing educational outcomes in SRL contexts.

The second phase presented in the SRL models is the execution phase, where tasks are performed while monitoring progress and performance (PANADERO, 2017). Various Web publishing tools to underline, highlight, and group teaching materials can be used at this stage. The work of Kitsantas (2013) describes several technologies that can help at this stage: social networks, virtual environments, administrative tools, testing tools, discussion forums, and bookmarks. Social networks are an important tool for student motivation. Most adolescents and adults currently access and dedicate part of their time to some social network. These interactions are important for connecting students and professionals. With the incorporation of social resources, students can self-monitor and define strategies for carrying out tasks.

In the execution phase, it is essential to develop and implement a well-structured study plan supported by clearly defined task strategies. Time management tools play a key role in assisting students with the self-control and organization of their learning activities. Virtual learning environments, in turn, provide a wide range of tools to support learning processes, allowing students to engage in simulations, modeling, training activities, and online meetings. Maintaining detailed records of these activities is fundamental for effective time management and monitoring of task completion. Collaborative tools, such as wikis, can also be employed to enable students to create, edit, and manage shared learning spaces, publish content, and receive feedback from peers and instructors. Overall, the self-regulation process encompasses self-monitoring, strategy definition, self-control, and peer modeling.

According to Sharp and Sharp (2016), time management plays a crucial role in fostering SRL in online environments. Effective time management strategies are essential for students to organize their learning activities, meet deadlines, and avoid procrastination, which is often prevalent in online settings. The use of tools such as online calendars, detailed course schedules, and checklists helps students plan their study time more effectively by breaking down larger tasks into manageable components with staggered due dates. These strategies promote higher engagement and academic performance by ensuring that learners have a structured approach to completing their assignments on time. By integrating these tools into course design, educators can support the development of SRL skills, ultimately leading to greater success in online learning contexts.

Effective time management allows these students to allocate sufficient time for their studies, thus optimizing their learning experience and minimizing the impact of external distractions. According to Hemmler and Ifenthaler (2024), time management serves as a crucial resource management strategy, ensuring that students can structure their study sessions efficiently and make use of other SRL strategies such as planning, monitoring, and self-assessment. This highlights the importance of integrating time management tools and practices into educational projects to support student autonomy and promote better learning outcomes.

According to Tzimas and Demetriadis (2024), personalized and structured feedback plays a crucial role in enhancing students' time management skills, a key component of SRL. In their study, students who received strong guidance through learning analytics interventions showed significant improvements in organizing their study schedules and making more effective use of their time. This targeted feedback allowed students to reflect on their performance and adjust their behaviors accordingly, demonstrating the impact of well-structured interventions on the development of SRL strategies, particularly time management. These findings highlight the importance of providing personalized support to foster better learning outcomes in online education.

Finally, there is the assessment phase, where the student reflects, regulates, and adapts his learning process for future executions (PANADERO, 2017). In this phase, blogs can also be allies to improve students' understanding of the subject, thus allowing the student to self-monitor and self-evaluate (KITSANTAS, 2013). Blogs and Wikis are important tools for collaborative learning and knowledge sharing. Students use these tools to monitor their evolutionary process within the platform and to receive feedback. The dissemination of bulletins with grades can help students in their self-assessments. Students with higher academic performances tend to self-assess more frequently than those with low performance (ZIMMERMAN; MARTINEZ-PONS, 1986). After evaluating himself, the student can reestablish new goals and outline new learning strategies to improve his academic performance.

Sharp and Sharp (2016) emphasize the importance of continuous assessment strategies that support SRL in online environments. They argue that formative assessments, such as quizzes, self-assessment exercises, and reflective journals, help students actively engage in their learning process. These assessments provide immediate feedback, allowing students to adjust their strategies, monitor their progress, and develop metacognitive skills. The authors highlight that regular, structured evaluation helps learners internalize feedback and apply it to future tasks, thus promoting deeper learning and sustained academic performance. This approach fosters an iterative learning cycle where assessment is integrated as an essential tool for improving learning outcomes.

SRL plays an important role in pedagogical environments that seek to promote students' autonomy, motivation, engagement, and responsibility in their own learning process (ZIMMERMAN; MARTINEZ-PONS, 1986). In this context, a well-structured PA provides the necessary support for students to effectively implement SRL strategies by offering an interactive, collaborative, and technology-rich learning environment. Through tools such as VLEs, automated feedback systems, time management, and progress monitoring, the PA facilitates planning, monitoring, and evaluation of learning, which are fundamental stages of SRL.

By incorporating these resources, the PA enables students to set clear goals, monitor their performance, and adjust their strategies as needed, fostering greater critical reflection and the development of autonomy. Thus, the combination of PA and SRL results in a more dynamic and efficient learning environment, where students are empowered to continuously manage and enhance their academic skills.

## 2.3 Learning Analytics

LA is a field that focuses on the measurement, collection, analysis, and dissemination of data about learners and their contexts to optimize learning and the environments in which it occurs, as defined in the 1st International Conference on Learning Analytics and Knowledge (LAK) (SIEMENS; LONG, 2011). LA examines how learning occurs in VLE by applying techniques to measure, collect, analyze, and present data on student interactions to improve the learning process (HAMDANE; MHOUTI; MASSAR, 2022).

This data can also be used to predict student performance and make informed decisions aimed at enhancing educational outcomes (ALMAAZMI *et al.*, 2022).

Authors such as Lockyer, Heathcote and Dawson (2013) have highlighted that LA enables the systematic collection of data on how students interact with learning resources and with other participants in the educational process, including teachers and tutors. By analyzing indicators such as time spent on activities, duration of online presence, interactions mediated through chats and forums, and preferences regarding activity formats, it becomes possible to obtain deeper insights into students' behaviors, engagement patterns, and learning trajectories within virtual learning environments.

The systematic review conducted by Alhazbi *et al.* (2024) examines empirical studies that use LA to measure SRL in higher education contexts. The primary objective of the review is to identify the indicators and metrics employed in previous investigations to assess students' self-regulation, with a focus on the digital tracking data generated by VLEs. According to the authors, the use of LA to measure SRL offers several advantages over traditional approaches, such as self-report questionnaires. By adopting LA, student behavior is captured non-intrusively by the VLE, allowing data to be collected without directly interfering with students' engagement in the course, ensuring greater accuracy and authenticity in the analysis of learning processes. The following indicators (construct) for measuring SRL were extracted: engagement, study regularity, anti-procrastination, help-seeking, monitoring, planning, motivation, and environmental structuring.

The engagement indicator comprises a wide range of metrics related to students' interaction with the learning environment, including the total time dedicated to accessing learning materials, the number of logins, participation in discussion forums, accesses to educational resources, views of instructional videos, submissions of formative assessments, total course views, and the number of days on which students accessed online materials (ALHAZBI *et al.*, 2024). Regularity, in turn, refers to the consistency and continuity of students' engagement over time, reflecting how steadily they interact with online learning resources throughout the course. Several studies rely on simple statistical measures to capture regularity, such as the average number of accesses during the course (CICCHINELLI *et al.*, 2018), the average number of hours dedicated and topics accessed per week (LI *et al.*, 2022), the average number of weekly logins (MONTGOMERY *et al.*, 2019), and the average number of formative assessment resolutions per session (CICCHINELLI *et al.*, 2018). In addition to these approaches, other studies employ more advanced statistical techniques to assess regularity, including the analysis of variance in weekly time devoted to learning tasks (GADELLA *et al.*, 2020).

The studies collected in Alhazbi *et al.* (2024) used different indicators to measure procrastination or anti-procrastination, based on the deadlines for submitting assessments or the completion of specific study units. In Li *et al.* (2018), the average number of days between deadlines, the day of questionnaire submission, and the number of questionnaires completed before the deadlines were used. Ye and Pennisi (2022) evaluated the number of late submissions, while Ilves, Leinonen and Hellas (2018) analyzed the number of days between the availability of tasks and the students' start of activities, as well as the average number of days of submission before the deadlines. The studies by Li, Baker and Warschauer (2020), Papamitsiou and Economides (2021), Rodriguez *et al.* (2021), and Feldman-Maggor, Blonder and Tuvi-Arad (2022) measured anti-procrastination based on the proportion of materials studied or tasks completed before the deadline.

One of the SRL strategies presented by Zimmerman and Martinez-Pons (1986) is help-seeking. Some studies have used this indicator to measure SRL ability based on participation in discussion forums. The work of Kim *et al.* (2018) analyzed the total time spent in the question-and-answer forum, the number of visits, and the number of contributions. In Ye and Pennisi (2022), the total number of questions asked to the instructor was used as an indicator of help-seeking. Regarding the construct of environment structuring, this refers to the strategies adopted by students for choosing and organizing their study environment, aiming to optimize the learning experience and minimize distractions. The study by Montgomery *et al.* (2019) considered factors such as location, day of the week, and time of access as indicators of this strategy.

The motivation indicator was measured in the studies by Yamada *et al.* (2017) and Tan *et al.* (2020). In the first, motivation was evaluated by the total number of pages read from online materials per minute and the total number of notes taken. In the second, student motivation was measured based on the time taken to complete tests. According to Winne and Hadwin (1998), planning refers to the ability to set goals before starting tasks. In the study by Cicchinelli *et al.* (2018), this ability was measured by the total and average number of accesses to organizational course resources, such as content objectives and assessment deadlines. The works by Raković *et al.* (2022) and Ye and Pennisi (2022) used the total number of accesses to syllabi, rubrics, calendars, and assessment instructions to measure this construct.

Different approaches are used to capture the monitoring ability. The studies by Pardo, Han and Ellis (2017) and Papamitsiou and Economides (2021) employed the number of times students accessed the dashboard analytics views as an indicator of their self-monitoring behavior. These dashboards serve as data visualization tools, providing students with insights into their performance and activity tracking.

In recent years, Learning Analytics Dashboards (LAD) as promising tools in supporting have emerged as promising tools for supporting SRL. As emphasized by Park and Jo (2019), the current challenge lies in developing Dashboards that are not only perceived by students as useful and effective but also positively influence their self-regulatory behaviors and, consequently, their teaching and learning process.

Complementing this perspective, various studies report that information visualization tools are effective in student self-regulation due to their ability to synthesize complex data in a manner that allows users to quickly understand and compare their results. These tools are also valued for assisting teachers in decision-making, as evidenced in the studies synthesized below.

Chen *et al.* (2019) describes the design of a Dashboard aimed at monitoring and organizing students' learning progress, divided into two main components: knowledge monitoring and SRL strategies. This tool allows students to view information about their own progress and compare it with that of their peers, fostering deep reflection on their learning process.

Manganello *et al.* (2021) explores the development of a Dashboard focused on monitoring SRL behaviors in an online platform. Utilizing LAD, the study highlights the analysis of behaviors categorized as "Consume", "Create", "Connect", and "Contribute", derived from a specific theoretical model, emphasizing the importance of data visualization in tracking student progress.

Additionally, Safsouf, Mansouri and Poirier (2022) introduces the TaBAT tool, an online Dashboard designed for virtual learning environments, which not only allows teachers to track students' progress but also provides students with the opportunity to view their own development, promoting self-regulation skills. The results indicate an improvement in student performance, enhancing their autonomy and academic success.

Farahmand, Dewan and Lin (2020) addresses the implementation of a LAD with dual functionalities: monitoring students' progress in comparison with their peers over time and providing automatic reminders to assist in meeting deadlines and managing tasks, highlighting the multifunctionality of Dashboards in supporting SRL.

Concluding, a study conducted by Botelho (2019) presents a comparative analysis of LAD plugins available for Moodle, focusing on their functionalities and how they facilitate the monitoring of the learning process by teachers and students. The study includes Table 3, which provides a detailed comparison of the examined plugins: Analytics Graphs (1); GISMO (2); Completion Progress (3); Heatmap (4); Forum Graph (5).

Table 3 – Plugins Comparison.

| **Comparison criteria** | **1** | **2** | **3** | **4** | **5** |
|---|---|---|---|---|---|
| Analysis and monitoring of student attendance | X | X | X | X | |
| Analysis and monitoring of activity delivery | X | X | | X | |
| Sending alerts or automatic messages | | | | | |
| Analysis and monitoring of interaction in synchronous and asynchronous communication tools | X | X | | | X |
| Sending messages by the teacher non-automatically | X | | | | |
| Information available for teachers | X | X | X | | X |
| Information available for students | | | X | X | |

*Source:* Botelho (2019)

The integration of LA and LAD into a PA aimed at encouraging SRL is essential, as these tools provide crucial data for personalizing and optimizing the learning process.

In a PA, a detailed understanding of student behavior, obtained through data collection and analysis by LA, allows teachers and automated systems to identify difficulties, monitor individual progress, and implement timely targeted interventions. This significantly enhances the ability to support students in their learning journey, adjusting pedagogical practices according to their specific needs. LAD complement this function by offering visual resources that enable students to monitor their own performance. This visualization provides students with a clear view of their progress and the areas needing improvement, promoting SRL, which is essential for autonomy in learning.

In addition to the fundamental role played by LA and LAD in promoting SRL, the application of EDM techniques stands out as a complementary and equally essential approach. While LA focuses on real-time data analysis, enabling immediate adjustments in the teaching-learning process, EDM delves deeper into extracting meaningful patterns from large volumes of educational data. These techniques not only allow the identification of latent trends and behaviors but also significantly enhance the capacity for pedagogical personalization and intervention, boosting the efficiency of a PA. Next, the key concepts of EDM and their relevance to the development of a PA that fosters SRL will be explored.

## 2.4 Educational Data Mining

Data Mining (DM) techniques are essential components of the Knowledge Discovery in Databases (KDD) process, which involves extracting knowledge from large volumes of data. This process consists of several stages, including selection, pre-processing, transformation, actual data mining, and the interpretation and evaluation of the extracted patterns. In the context of learning data, this practice is commonly referred to as EDM.

Educational systems used for online teaching generate a large amount of data, in particular, records of student interactions with the system. These data can be used to detect interesting insights about the learning process through the use of DM techniques. EDM is a specific area of data mining that focuses on analyzing data related to educational contexts (COSTA; DORÇA; ARAÚJO, 2020). The main goal of EDM is to improve the quality of the teaching and learning process by analyzing stored data about it.

In recent years, there has been an exponential growth in the use of VLEs, both for distance learning and in support of face-to-face or hybrid education, due to the Covid-19 Pandemic, large amounts of educational data have been generated. In this context, EDM techniques can be used to analyze the educational data in these learning environments.

According to Fischer *et al.* (2020), analyzing these data is crucial for gaining a deeper understanding of students, learning processes, the environment where learning occurs, as well as other elements that may impact learning. Furthermore, this exploration enables the identification of patterns and trends that can help customize and enhance educational strategies, resulting in more effective teaching tailored to the individual needs of students.

Specifically, data analysis can provide valuable insights into SRL, allowing educators to develop targeted interventions to support students in managing their own learning and in setting goals, planning, and monitoring their activities.

EDM utilizes a variety of advanced analytical techniques to extract meaningful information from large volumes of educational data. Among the techniques most frequently employed in EDM: Prediction, Clustering, Relationship Mining, Distillation of Data for Human Judgment, and Discovery with Models (BAKER; ISOTANI; CARVALHO, 2011).

In the Prediction, the goal is to develop models capable of inferring specific aspects of data, known as predicted variables, through the analysis and integration of various elements present in the data, called predictor variables. Prediction requires that a portion of the data be manually coded to enable the correct identification of one or more previously known predictive variables. There are three main types of prediction: classification, regression, and density estimation (BAKER; ISOTANI; CARVALHO, 2011).

On the other hand, Relationship Mining aims to identify possible connections between variables within databases. This task can include attempts to determine which variables are most strongly associated with a specific variable previously identified as important, or it may involve exploring the relationships between any variables present in the data. To identify these connections, four main approaches are used: (a) association rules, (b) correlations, (c) sequences, and (d) causal (BAKER; ISOTANI; CARVALHO, 2011).

Two other techniques presented in the work of Baker, Isotani and Carvalho (2011) are Distillation of Data for Human Judgment and Discovery with Models. The first technique aims to conduct research that presents complex data in a simplified manner, highlighting its most important characteristics. Through distillation, it is possible for the data to be used by people to infer aspects about them and thus make decisions that previously could not be made or automated using only traditional EDM methods. An example is the learning curve, which indicates the level of learning of a student or a group of students.

The EDM technique employed in this study is applied to data collected from students who made use of the proposed PA. According to Fischer *et al.* (2020), the primary objective of clustering—the central EDM technique adopted—is to identify natural groupings within a dataset, allowing instances to be classified into previously unknown groups or categories. Through the analysis of students' behavioral and performance-related data, these groups are automatically identified, enabling a more detailed understanding of learning profiles, engagement patterns, and behavioral characteristics within the SRL environment. The integration of EDM into the PA supports an in-depth and systematic analysis of educational data, contributing to the identification of potential academic difficulties and enabling the adaptation of pedagogical resources to address students' specific needs. In this context, the combined use of LA and EDM provides a robust foundation for the development of data-driven pedagogical practices aimed at fostering SRL, enhancing personalization, and promoting greater student autonomy. The following section presents a more detailed discussion of this technique and its application within the approach.

## 2.5 Data Clustering

Over the years, the main DM techniques employed in the EDM process have been classification and clustering (ALDOWAH; AL-SAMARRAIE; FAUZY, 2019). Classification is a supervised learning technique in which a predictive model is trained using a dataset composed of input attributes and corresponding output labels, enabling the prediction of predefined classes for new instances. Clustering, in contrast, is an unsupervised learning technique that does not require labeled data, as the output associated with each record is unknown, allowing patterns and natural groupings to emerge directly from the data.

In the educational context, the clustering technique can be employed to analyze students' behavioral data and identify patterns that facilitate the personalization of learning (RABELO *et al.*, 2024). This includes identifying groups of students with similar characteristics in terms of performance, engagement, and learning needs. By uncovering these groups, educators and intelligent systems can design targeted interventions and adaptive strategies that better support students' progress and improve overall learning outcomes.

Zorrilla, García and Álvarez (2010) assert that through mining tasks such as clustering and association, it is possible to identify student profiles and their respective learning patterns, providing teachers and tutors with valuable information about the teaching and learning process. Practical applications of clustering include creating student profiles, personalizing educational materials, and identifying interventions needed to improve student performance. By analyzing groups of students, educators can develop more effective teaching strategies that cater to the specific needs of each group (RABELO *et al.*, 2024).

In EDM, a variety of clustering techniques are employed, each selected based on the specific characteristics of the data and the objectives of the analysis (RABELO *et al.*, 2024). The K-Means algorithm is particularly popular due to its simplicity and effectiveness in handling large datasets (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Hierarchical algorithms are valued for their ability to produce a hierarchy of clusters, allowing similar data to be grouped together without the need to specify the number of clusters in advance (MURTAGH; CONTRERAS, 2011). Density-Based Clustering Based on Hierarchical Density (HDBSCAN) excels at identifying clusters with varying densities, which is advantageous for datasets with diverse groupings (MCINNES JOHN HEALY, 2024). Additionally, techniques such as Density-Based Algorithm (DBSCAN), Expectation Maximization (EM) and Spectral Clustering are also utilized. These algorithms provide valuable insights by grouping students based on their characteristics and learning behaviors, which can substantially aid in customizing learning strategies and interventions.

In this study, four clustering algorithms were used: K-Means, HDBSCAN, Agglomerative Clustering, and EM, to conduct a comparative analysis aimed at identifying the best clustering method regarding SRL skills in the data collected from VLEs. The selection of these algorithms was based on their distinct characteristics and their ability to detect relevant patterns in educational data. The comparative analysis allowed for an evaluation of which algorithm would be most effective in identifying groups of students with different SRL profiles, providing more accurate insights for the personalization of pedagogical interventions.

### 2.5.1 K-Means

A widely used clustering technique in EDM is the K-Means algorithm (AGGARWAL, 2015). This partitioning algorithm aims to group data into $k$ distinct clusters based on similarity measures among instances (TAN *et al.*, 2016). The clustering process requires the prior specification of the number of groups to be formed, which represents a critical parameter in the application of the algorithm. To support this decision, several evaluation metrics can be employed, such as the silhouette coefficient (COSTA; DORÇA; ARAÚJO, 2020), which is obtained by executing the K-Means algorithm with different values of $k$ and assessing the quality of the resulting cluster structures.

The K-Means algorithm follows a simple and efficient procedure, consisting of the following steps: initially, $k$ starting points, known as centroids, are chosen, which can be selected randomly or through specific initialization methods. Subsequently, each point in the dataset is assigned to the nearest centroid, forming $k$ clusters, typically using Euclidean Distance to measure proximity. The steps of cluster assignment and centroid recalculations are repeated iteratively until the centroids stabilize or a maximum number of iterations is reached. Convergence is achieved when there are no significant changes in the centroids between iterations. The final result is a set of $k$ clusters, each with its own centroid and associated data points (TAN *et al.*, 2016).

### 2.5.2 HDBSCAN

The HDBSCAN is a divisive hierarchical clustering algorithm based on the Minimum Spanning Tree (MST) obtained from the Accessibility Distance Graph Mutual. This new algorithm transforms the DBSCAN into a hierarchical form (CAMPELLO; MOULAVI; SANDER, 2013). The HDBSCAN operates by progressively removing edges with the highest weights (mutual reachability distance) from the MST in descending order, thus establishing each level of the hierarchy with connected objects (clusters) and isolated objects (outliers).

The HDBSCAN algorithm is particularly effective in scenarios where the shape, size, and density of clusters are not known a priori and may vary considerably across the dataset. Unlike many learning methods, HDBSCAN does not rely on a target variable, which characterizes it as an unsupervised learning technique. Rather than predicting predefined outcomes, the algorithm focuses on identifying patterns and underlying structures in the data by analyzing the density distribution of data points, enabling the detection of clusters with varying densities as well as noise (MCINNES JOHN HEALY, 2024).

The flexibility of HDBSCAN in handling clusters of arbitrary shapes and its ability to identify outliers make it a powerful tool for educational data analysis, where student behavior patterns can vary widely. This method provides valuable insights by identifying groups of students with similar learning characteristics and behaviors, which can assist in customizing educational strategies and pedagogical interventions.

### 2.5.3 Agglomerative Clustering

Agglomerative Clustering is part of the family of hierarchical algorithms and uses a bottom-up approach to cluster execution, i.e., each element in the dataset is started in a group and at each step pairs of elements are merged according to their proximity (TAN *et al.*, 2016).

The algorithm is implemented within the scikit-learn library [1], which offers the flexibility to select from various linkage criteria for merging clusters. The available options include *single linkage*, *complete linkage*, *average linkage*, and *Ward*. These methods provide diverse approaches to calculating distances between clusters, thereby accommodating different dataset characteristics and analysis requirements.

One of the primary advantages of Agglomerative Clustering lies in its ability to construct a hierarchical organization of clusters. This characteristic is particularly valuable in applications where hierarchical or hereditary relationships between clusters are relevant, or when the exact number of clusters is not known in advance (ACKERMANN *et al.*, 2014). The resulting hierarchical structure enables analysts to examine cluster formations at multiple levels of granularity, supporting both the customization of clustering outcomes to specific analytical goals and the in-depth exploration of data structures, which is especially beneficial in exploratory data analysis and other contexts that require flexible clustering approaches.

### 2.5.4 Expectation Maximization

EM algorithm is an iterative technique used for estimating the parameters of probabilistic models, particularly when the data are incomplete. Rather than assigning a single most likely outcome for the missing variables at each iteration, the EM algorithm calculates the probabilities of all possible outcomes of the missing data using the current parameters. These probabilities are then used to create a weighted training set, consisting of all possible completions of the data. A modified version of maximum likelihood estimation, which handles weighted training examples, provides new parameter estimates. In

[1] https://scikit-learn.org/stable/

this way, the EM algorithm takes into account the model's confidence in each data conclusion, leading to a more robust estimation of parameters (DO; BATZOGLOU, 2008).

Furthermore, the EM algorithm is designed to estimate the maximum likelihood parameters of a statistical model in a wide range of situations, particularly when the underlying equations cannot be solved analytically. This iterative technique is composed of two fundamental steps: the E step (Expectation) and the M step (Maximization). During the E step, the algorithm computes the expected value of the conditional likelihood function of the parameters, given the observed data and the current parameter estimates. Subsequently, in the M step, this conditional likelihood function is maximized with respect to the parameters, resulting in updated estimates for the model parameters. These two steps are repeated iteratively until the parameter values converge to a stable solution, indicating that the maximum likelihood estimates have been reached (MOON, 1996).

The clustering algorithms presented above play a crucial role within the proposed PA, as they enable an in-depth analysis of the educational data collected and facilitate the identification of behavioral patterns and distinct SRL profiles. A comparative analysis among these algorithms is conducted in order to determine the most suitable technique for the characteristics of the collected data, after which the clustering process is performed based on students' behavioral, engagement, and performance-related attributes. Through this process, the PA can support the personalization of pedagogical interventions by identifying groups of students with similar characteristics, learning behaviors, and educational needs, as well as detecting atypical patterns, such as students at risk of poor academic performance, thus enabling timely and preventive interventions. In addition, these algorithms contribute to monitoring the evolution of SRL skills over time, allowing changes in students' learning behaviors to be systematically observed. Furthermore, data collected from the resources provided by the PA are analyzed to assess whether such resources effectively contribute to the development of SRL. By integrating multiple analytical dimensions, the proposed PA becomes increasingly adaptive and data-driven, thereby promoting SRL development and supporting students' academic success.

The development, analysis, and validation of the resources provided in the proposed PA are grounded in the theories discussed throughout this chapter, including the concepts of PA, SRL, LA, and EDM. These approaches provide a solid foundation for a deeper understanding of the factors that impact the development of SRL, while also facilitating the personalization of pedagogical interventions. The proposal evolves by integrating these theories and techniques into a single PA, which aims to both identify SRL profiles and examine the correlation between academic performance, engagement, and satisfaction in the learning process. Although innovative, this approach is supported by evidence from previous studies that have already explored the potential of these theories and technologies. In the next chapter, related works addressing these topics will be discussed, highlighting how previous research has employed similar technologies and strategies to foster SRL and optimize academic performance.

# Chapter 3

# Related Work

In this chapter, we present the main related works in the areas of PA, SRL, and EDM. Initially, we will discuss the key studies that address PA, highlighting their objectives, technologies used, and contributions. Then, we will explore the works that employ EDM techniques to profile students' SRL, emphasizing the types of SRL profiles identified, the data sources used, and the techniques applied in each study. Additionally, we will present a summary of the main technological strategies used to stimulate SRL.

Menezes *et al.* (2013) introduced forward-thinking ideas for distance education, emphasizing networked learning through the framework of Pedagogical Architectures. The study delves into the strategic use of cooperative evaluation mechanisms and technological aids to advance networked learning in remote education settings. These pioneering concepts were applied and tested within the specific milieu of the Integrator Seminar. The paper outlines three unique models of pedagogical architecture: Thesis Debates, Collaborative Concept Maps, and Reflective Portfolios.

Following this line of innovation, Sonego *et al.* (2018) presents an PA for M-learning considering the use of applications and interactive content for mobile devices. This architecture expands the possibilities for teachers who plan to use mobile devices in schools. In this way, it helps to guide and guide the teacher in this process, enabling the architecture to adapt according to the level of education and context of the target audience.

Another significant study, presented by Sonego (2019), identified the elements that comprise the aspects of a pedagogical architecture (organizational, content, methodological, and technological), in addition to the strategies focused on mobile learning. Using the Design Science Research methodology, a pedagogical architecture called "ArqPed-Mobile" was developed, applied in an extension course to guide teachers on the effective use of mobile technology in the classroom. The results indicate that ArqPed-Mobile, although it needs to be adapted according to the context, has the potential to revolutionize pedagogical practices, encouraging new learning methods through mobile devices.

Ribeiro (2019) focuses on the development of a Pedagogical Model centered on social interactions in Virtual Learning Environments (MP SocioAVA), based on Piaget's theory.

The research emphasizes the importance of interactions to keep students engaged, challenging the common perception that distance education is predominantly solitary. Using the concept of the "social subject" and the ROODA Virtual Environment, the study proposes a pedagogical architecture and versions of the MP SocioAVA. The methodology used was qualitative, with multiple case studies in various courses. The analysis highlights the need for pedagogical models that encourage interaction between students and between students and teachers for effective learning.

Silva, Menezes and Junior (2023b) proposed a PA based on Computational Thinking for problem understanding and the construction of a Learning Portfolio, aimed at teaching programming. Meanwhile, the study conducted by Jacaúna *et al.* (2022) combined a PA with digital games, aiming to offer new learning situations focused on education about waste disposal, recycling, and reuse of materials. In Biancardi, Menezes and Vilhagra (2020), a PA is presented for the Cooperative Construction of Reflective Reviews in the Context of Distance Education. Table 4 summarizes the aforementioned works and finally describes the PA proposal that will be developed in the current study.

Several studies found in the literature indicate that there are different SRL profiles among students in VLEs. These profiles can be defined using EDM techniques on data collected in VLEs, through self-report or trace data. Table 5 presents a summary of the works that identified the students' SRL profiles, using self-report questionnaires (SR) and trace data (TD).

In Ye and Pennisi (2022), the authors investigate the correlation between students' self-reported SRL data and digital tracking data collected in VLEs. By applying cluster analysis, students were classified into different levels of self-regulatory skill, and the characteristics associated with each group were examined in detail. The results indicate that digital tracking data predict student performance more accurately than self-reported SRL measures, thereby challenging the validity and reliability of traditional SRL instruments. Furthermore, the study discusses important practical implications for online education, highlighting that the use of digital behavioral data can enhance the personalization of the educational process and provide more precise insights into students' learning behaviors.

Barnard-Brak, Paton and Lan (2010) proposed a study to examine whether there are profiles for self-regulated learning skills and strategies among students. They performed two studies with two different samples. The Online Self-Regulated Learning Questionnaire (OLSQ) was applied. They used latent class analysis to identify SRL profiles, resulting in the presence of five distinct SRL profiles replicated in both study samples: super self-regulators, competent self-regulators, premeditated self-regulators, performance/reflection self-regulators, and no or minimal self-regulators.

The study Cao *et al.* (2023) applies LA techniques to analyze differences in time management among students with different levels of academic performance. The article highlights that students with better academic performance tend to invest more time in

Table 4 – Summary of PA From Previous Studies.

| N | Goals | Technologies | Results |
|---|---|---|---|
| 1 | Develop innovative proposals with an emphasis on networked learning, in the context of the Integrative Seminar, through the use of PA. | Blog, Wiki, and VLE. | Three PA were developed and applied with a focus on networked learning: Thesis Debates, Learning Projects and Learning Portfolio. The results indicate that the PA developed and applied in the course provided a collaborative learning environment, with an emphasis on the collective construction of knowledge, meaningful interactions and a new approach to teacher training. |
| 2 | Identify the elements that constitute the aspects of a PA with a focus on mobile learning. | Mobile devices, Applications, Interactive content, and VLE. | Presentation of ArqPed-Mobile with indicators for each aspect of PA and pedagogical strategies for implementation. |
| 3 | Analyze how MP_SocioAVA can contribute to the development of social interactions of distance learning students in a VLE. | Youtube, Weebly, Support tools, and VLE. | MP_SocioAVA presented, in its elements, ways to contribute to the development of social interactions in VLE. |
| 4 | Propose and evaluate a PA for M-Learning with a focus on teacher training. | Mobile devices, Applications, Interactive content, and VLE. | The PA developed was effective in providing significant training for teachers regarding the use of mobile devices in education, promoting reflection on pedagogical practices and encouraging the adoption of new technologies for learning. |
| 5 | Propose a PA based on computational thinking to understand the problem and build a learning portfolio. | Learning Portfolio, Google Meet, Google Docs, Loom, and VLE. | The PA proposal allowed students to develop new ways of understanding and solving problems through computational thinking, promoting interaction and cooperation between students. |
| 6 | Combine PA with digital games to offer new learning situations, focusing on learning about disposal, recycling, and reuse of materials. | Digital Games, Mconf, Google Docs, and VLE. | The results indicated that the combination of PA with digital games is an effective approach to promote collaborative learning and knowledge construction, as well as providing a more comprehensive and innovative training for the educators involved. |
| 7 | Propose a PA for the cooperative construction of reflective reviews in the context of Distance Education. | Chat, forums, Google Docs, Google Sheets, Google Forms, and VLE. | The proposed PA promoted a significant improvement in cooperation and the quality of student productions, with an average increase of 9% in the grades of collective reviews and a positive perception from students about the activity. |
| 8 | Develop a PA supported by AI to encouraging SRL behavior in students. This PA should include resources that help students set goals, plan, and monitor their progress. | Learning Analyctis, Time Tracker, Generative AI, Plugins Moodle and VLE. | The proposed PA is expected to effectively support students through the three phases of SRL, thereby enhancing their engagement and motivation. |

(1)(MENEZES *et al.*, 2013), (2)(SONEGO, 2019), (3)(RIBEIRO, 2019), (4)(SONEGO *et al.*, 2018), (5)(SILVA; MENEZES; JUNIOR, 2023b), (6)(JACAúNA *et al.*, 2022), (7)(BIANCARDI; MENEZES; VILHAGRA, 2020), and (8)Current study

SRL activities and exhibit more consistent time-use patterns throughout the academic term. The results suggest that clickstream data analysis provides valuable insights into SRL strategies and can help design personalized pedagogical interventions to optimize online learning.

Li *et al.* (2018) analyzed the student tracking data in VLE collected throughout the course. They used records related to access to learning materials, completion of questionnaires, and response records to profile the SRL. The K-Means clustering algorithm was applied and four distinct groups were identified: 1) early graduates, 2) late graduates, 3) early dropouts, and 4) late dropouts.

A mixed approach was used in the work of Ainscough *et al.* (2019). Trace and self-report data were used to define the SRL profiles. They were divided into three groups: high self-regulators, medium self-regulators, and low self-regulators. A two-step cluster analysis was used to group students. The first step was the formation of the pre-cluster. In the second step, the hierarchical clustering algorithm was used to merge the pre-clusters, leading to the three different clusters.

Costa, Dorça and Araújo (2020) analyzed data from a Ubiquitous Learning Environment (ULE) using data clustering techniques to observe student behavior in learning sessions. The authors applied the K-Means algorithm to perform data clustering. Two groups were found, and one of these groups showed strong evidence of students' self-regulation capabilities.

The review presented by ElSayed *et al.* (2019) showed that there is a lack of studies to define which EDM algorithm has a better performance in identifying SRL profiles through tracking data collected in VLEs. The present study aims to identify students with self-regulation and non-regulation profiles through data collected from the proposed PA. EDM techniques will be used to classify students' profiles based on their learning behaviors. Additionally, self-regulation questionnaires will be administered to complement the analysis and provide a broader understanding of students' self-regulation strategies. These profiles will be correlated with students' academic performance, allowing for a deeper understanding of how self-regulation strategies impact educational outcomes.

In the systematic literature review conducted as part of this research, the technologies utilized to support and encourage SRL in VLEs were meticulously mapped and synthesized (LIMA *et al.*, 2024). This review focused on works published between 2011 and 2020, aiming to evaluate the effectiveness of these technologies across different phases of the self-regulation process, predominantly in higher education contexts. Additionally, the article provides an overview of the state of the art, offering guidelines on the various technologies employed to foster self-regulatory skills in VLEs.

Strategies for visualizing personalized information have been used in the three phases of the self-regulatory process. Molenaar *et al.* (2020) utilized log data to help students establish learning goals and monitor progress. Barria-Pineda, Guerra-Hollstein

Table 5 – Summary of SRL Profiles Identified From Previous Studies.

| **Article** | **SRL profiles identified** | **SD** | **Technique applied** |
|---|---|---|---|
| (YE; PENNISI, 2022) | High SRL level<br>Moderate SRL level<br>Low SRL level | SR<br>TD | K-Means |
| (BARNARD-BRAK; PATON; LAN, 2010) | Super self-regulators<br>Competent self-regulators<br>Forethought-endorsing self-regulators<br>Performance/reflection self-regulators<br>Non/minimal self-regulator | SR | Latent class analysis |
| (YOT-DOMÍNGUEZ; MARCELO, 2017) | High-level regulators<br>Low-level regulators | SR | Stepwise cluster analysis<br>Hierarchy analysis<br>Ward method<br>K-Means |
| (CAO *et al.*, 2023) | High SRL<br>Medium SRL<br>Low SRL | TD | K-Means |
| (LI *et al.*, 2018) | Early completers<br>Late completers<br>Early dropouts<br>Late dropouts | TD | K-Means |
| (KIM *et al.*, 2018) | Self-regulation<br>Partial self-regulation<br>Non-self-regulation | SR<br>TD | K-Means |
| (AINSCOUGH *et al.*, 2019) | High self-regulators<br>Medium self-regulators<br>Low self-regulator | SR<br>TD | Two-step cluster analysis |
| (COSTA; DORÇA; ARAÚJO, 2020) | Cluster 0: Non self-regulation<br>Cluster 1: Evidence self-regulation | TD | K-Means |
| (PEACH *et al.*, 2019) | Early birds<br>On time<br>Low engagers<br>Crammers<br>Sporadic outliers (unclustered learners) | TD | Mathematical framework (based on dynamic time warping kernel and clustering algorithm) |
| (ÇEBI; GÜYER, 2020) | Cluster 1: Students with least interaction<br>Cluster 2: Intense interaction with video, example, and forum activities<br>Cluster 3: Students who spend more time on tutorial, exercises, concept map, summary, and highlight activities | SR<br>TD | Hierarchical clustering<br>K-Means |
| This study | Evidence self-regulation<br>Non self-regulation | SR<br>TD | EM<br>Agglomerative clustering<br>K-Means<br>HDBSCAN |

Source of Data (SD), Self-Report (SR) and Trace Data (TD).

and Brusilovsky (2018) presented an open model for visualizing progress in course topics. Su (2020) and Kia *et al.* (2020) employed visualizations for personalized learning and monitoring goals. Ilves, Leinonen and Hellas (2018) studied the impact of different visualizations on academic performance.

The Doubtfire++ tool (LAW *et al.*, 2017) supports SRL with visualizations of learner models. Robal *et al.* (2018) proposed real-time attention loss detection in videos. Gamification strategies, such as digital badges (MORRIS *et al.*, 2019) and adaptive educational games (LEONARDOU; RIGOU; GAROFALAKIS, 2019), have been used to motivate students and assess self-regulation. Al-Hatem, Masood and Al-Samarraie (2018) utilized the game Second Life for nursing education.

Moccozet and Tardy (2014) proposed a social learning platform with peer feedback and gamification, while Tang and Fan (2011) described an SRL platform enhanced with Web 2.0 technologies. Gaeta *et al.* (2011) developed a web-based metacognitive environment using Semantic and Social Web methods.

Activity records (log data) and interactive learning resources are also utilized to support SRL (MANSO-VÁZQUEZ; CAEIRO-RODRÍGUEZ; LLAMAS-NISTAL, 2015; HUANG *et al.*, 2014; JANSEN *et al.*, 2020) used short videos for SRL instruction, and Wong *et al.* (2019) and Liu, Zheng and Jiang (2019) added interactive elements to videos for better engagement.

Tan *et al.* (2018) proposed collaborative video annotation with real-time feedback. Mentari, Subchan and Supeno (2020) developed an online learning environment supporting all SRL phases. Lawrie *et al.* (2016) designed online modules for chemistry education to foster SRL through visual representations and feedback.

Wang (2011) described a web-based assessment system that incorporates SRL strategies such as peer feedback, while Menezes (2017) employed content recommendation and peer review mechanisms to support SRL. In a complementary perspective, Fung, Abdullah and Hashim (2019) and Broadbent, Panadero and Fuller-Tyszkiewicz (2020) suggested the use of personalized journals and mobile diaries to support self-reflection activities.

Romero *et al.* (2019) presented a conceptual model using e-portfolios for SRL. França and Tedesco (2014) proposed a collaborative model for computational thinking with self-regulation elements. The Self Regulation System (SRS) incorporates diagnostic, continuous, and monitoring modules (LIMA; PIMENTEL, 2013).

Yun, Fortenbacher and Pinkwart (2017) used sensor technology to promote SRL, detecting behaviors and providing feedback. Ruipérez-Valiente *et al.* (2022) developed Vi-Tracker for tracking student progress using Bayesian Networks. Chen and Huang (2014) proposed a web-based reading annotation system with attention-based SRL mechanisms.

Several recent studies have suggested that ChatGPT, as an AI tool, has the capacity to enhance the SRL process (CHIU *et al.*, 2023; MOLENAAR *et al.*, 2023; WU *et al.*, 2024; XIA *et al.*, 2023). These investigations propose that ChatGPT can potentially foster SRL through various mechanisms.

Building on this idea, Chiu (2024) developed a classification tool for learning activities utilizing ChatGPT, grounded in Self-Determination Theory (SDT) and aimed at encouraging SRL. The study engaged a panel of 36 expert teachers who, using a Delphi method, identified and refined 20 learning activities designed to fulfill students' needs for autonomy, competence, and relatedness, as described by SDT. The findings indicate that ChatGPT can provide immediate and personalized feedback, thereby enhancing students' competence and supporting all stages of SRL.

Throughout the study, newly developed technologies were mapped and are presented in the current work. These advances reflect the ongoing innovation in educational technology, aiming to provide more effective and personalized learning experiences. The comprehensive list of technologies identified both in the systematic review and in additional recent studies is summarized in Table 6, along with their respective references.

This research innovates in the field of PA by integrating various theories, such as SRL, LA, and EDM. The proposed approach not only personalizes pedagogical interventions based on students' profiles but also promotes the continuous development of self-regulation skills, creating a more adaptive and effective learning environment. Thus, this research expands the field of Pedagogical Architectures by incorporating a data-driven approach grounded in established theories, providing valuable insights for enhancing SRL and fostering students' academic success.

Similar to the studies by Cao *et al.* (2023) and Ye and Pennisi (2022), which use digital tracking data (clickstream) and LA and EDM techniques to analyze students' behavior and SRL strategies, the present research also relies on the collection and analysis of educational data to identify SRL profiles. However, the approach adopted here goes further by integrating self-regulation questionnaires and focus group interviews, combining quantitative and qualitative data to provide a more comprehensive view of the development of SRL skills and the effectiveness of the proposed PA. Additionally, the Time Tracker SRL plugin, specifically developed for this research, will enable detailed monitoring of the time students spend on learning activities in the virtual environment, offering a new strategy for time management.

The proposed PA also differs from related work by holistically integrating organizational, content, methodological, and technological aspects to encourage SRL. By utilizing a combination of technologies and interactive methodologies, our proposal aims not just to facilitate knowledge transmission, but also to engage students in an active and reflective learning process, promoting autonomy and educational effectiveness. Specifically, the PA incorporates various plugins in VLE, such as Analytics Graphs, Configurable Reports, OpenAI Chat (Athena), Completion Progress, and the Time Tracker SRL. These plugins provide essential resources for students to set goals, plan their activities, monitor their progress, and reflect on their performance.

Table 6 – Summary of technological strategies to stimulate SRL in the articles analyzed.

| **Technologies** | **Articles** |
|---|---|
| Adaptation | (MENEZES, 2017), (FUNG; ABDULLAH; HASHIM, 2019), (BROADBENT; PANADERO; FULLER-TYSZKIEWICZ, 2020), (NEITZEL; RENSING; BELLHÄUSER, 2017), (SPILIOTOPOULOS *et al.*, 2019), (KHIAT, 2019), (LEE; BARKER; KUMAR, 2011), and (SELVI; PANNEERSELVAM, 2012) |
| Peer evaluation | (WANG, 2011), (MENEZES, 2017), and (FRANÇA; TEDESCO, 2014) |
| e-portfolios | (MANSO-VÁZQUEZ; CAEIRO-RODRÍGUEZ; LLAMAS-NISTAL, 2015), (ROMERO *et al.*, 2019), (KARAMI *et al.*, 2019), (LAW *et al.*, 2017), and (HSU, 2020) |
| Linked Data / Semantic Web | (GAETA *et al.*, 2011) and (ROMERO *et al.*, 2019) |
| Feedback strategies | (MOCCOZET; TARDY, 2014), (HUANG *et al.*, 2014), (SIROTHEAU *et al.*, 2011), (TAN *et al.*, 2018), (MENTARI; SUBCHAN; SUPENO, 2020), (LAWRIE *et al.*, 2016), (LEE; BARKER; KUMAR, 2011), (KINNARI-KORPELA; SUHONEN, 2020), (SINATRA, 2014), (PÉREZ-ÁLVAREZ; MALDONADO-MAHAUAD; PÉREZ-SANAGUSTÍN, 2018), (KAPP; BRAUN; HARA, 2016), (DELEN; LIEW; WILLSON, 2014), and (OGATA *et al.*, 2017) |
| Open student model | (BARRIA-PINEDA; GUERRA-HOLLSTEIN; BRUSILOVSKY, 2018), (LAW *et al.*, 2017), (LEONARDOU; RIGOU; GAROFALAKIS, 2019), (MENEZES, 2017), and (MOLENAAR *et al.*, 2020) |
| Interactive learning resources | (HUANG *et al.*, 2014), (JANSEN *et al.*, 2020), (WONG *et al.*, 2019), (KAUFFMAN; ZHAO; YANG, 2011), (LIU; ZHENG; JIANG, 2019), (DELEN; LIEW; WILLSON, 2014), and (BAHRI *et al.*, 2021) |
| Social / collaborative features | (MOCCOZET; TARDY, 2014), (GAETA *et al.*, 2011), (TAN *et al.*, 2018), (FRANÇA; TEDESCO, 2014), (TANG; FAN, 2011), (JUNUS; SANIOSO; SADITA, 2014), (LEE; BARKER; KUMAR, 2011), (MORGAN; PETTERSEN; DAALHUIZEN, 2020), (PARASKEVA *et al.*, 2017), and (BAHRI *et al.*, 2021) |
| Information visualization | (MOLENAAR *et al.*, 2020), (BARRIA-PINEDA; GUERRA-HOLLSTEIN; BRUSILOVSKY, 2018), (SU, 2020), (KIA *et al.*, 2020), (ILVES; LEINONEN; HELLAS, 2018), (MANSO-VÁZQUEZ; CAEIRO-RODRÍGUEZ; LLAMAS-NISTAL, 2015), (LAWRIE *et al.*, 2016), (FRANÇA; TEDESCO, 2014), (PÉREZ-ÁLVAREZ; MALDONADO-MAHAUAD; PÉREZ-SANAGUSTÍN, 2018), (CHEN; HUANG, 2014), (HAYNES, 2020), (HE *et al.*, 2019), (FARAHMAND; DEWAN; LIN, 2020), (PÉREZ-ÁLVAREZ *et al.*, 2017), (PHODONG; SUPNITHI; KONGKACHANDRA, 2019), (KHIAT, 2019), (OGATA *et al.*, 2017), (JR. *et al.*, 2020), and (CHEN *et al.*, 2019) |
| Formative assessment | (WANG, 2011), (LIMA; PIMENTEL, 2013), (KAPP; BRAUN; HARA, 2016), and (SELVI; PANNEERSELVAM, 2012) |
| Sensors | (ROBAL *et al.*, 2018), (YUN; FORTENBACHER; PINKWART, 2017), (RUIPÉREZ-VALIENTE *et al.*, 2022), (SINATRA, 2014), and (CHEN; HUANG, 2014) |
| Gamification | (MORRIS *et al.*, 2019), (LEONARDOU; RIGOU; GAROFALAKIS, 2019),(AL-HATEM; MASOOD; AL-SAMARRAIE, 2018), (MOCCOZET; TARDY, 2014), (SPILIOTOPOULOS *et al.*, 2019), and (TANG; KAY, 2014) |
| Generative AI | (CHIU, 2024), (CHIU *et al.*, 2023),(MOLENAAR *et al.*, 2023), (WU *et al.*, 2024), and (XIA *et al.*, 2023), |

# Chapter 4

# Method

The methodology adopted in this work follows an adapted Sequential Explanatory Mixed-Methods Design, which integrates quantitative and qualitative approaches into a single, coherent, and articulated research process. This design was selected to capture the complexity of educational phenomena through the complementary use of both approaches, enhancing analytical depth, validity, and the applicability of the findings (BORREGO; DOUGLAS; AMELINK, 2009).

As shown in Figure 2, the methodological framework was structured into multiple interrelated stages, beginning with the definition of the research problem and questions, a systematic literature review, and an analysis of related works. These initial procedures established the theoretical and empirical foundations of the study and guided the subsequent development of the proposed PA. The research design encompassed a sequence of complementary stages, including two quantitative phases — Proof of Concept 1 (OULAD) and Proof of Concept 2 (Moodle IFCDM)— aimed at identifying patterns of engagement and self-regulated learning behaviors through educational data mining techniques; the development of an artificial intelligence–supported PA; its implementation and evaluation in a real learning environment; and a qualitative phase focused on students' perceptions through focus group interviews. The integration of these results was conducted through a triangulation process that combined quantitative and qualitative evidence. The following sections describe each of these stages in detail.

## 4.1 Preliminary Stage

The initial stage of this research was dedicated to identifying and defining the investigated problem, as well as formulating the research questions that guide the development of the study. To this end, a systematic literature review and a critical analysis of related works were conducted, aiming to deepen the understanding of the challenges faced by students in VLEs and to identify gaps in existing approaches aimed at promoting SRL.

The literature review revealed that SRL is a fundamental process for academic success

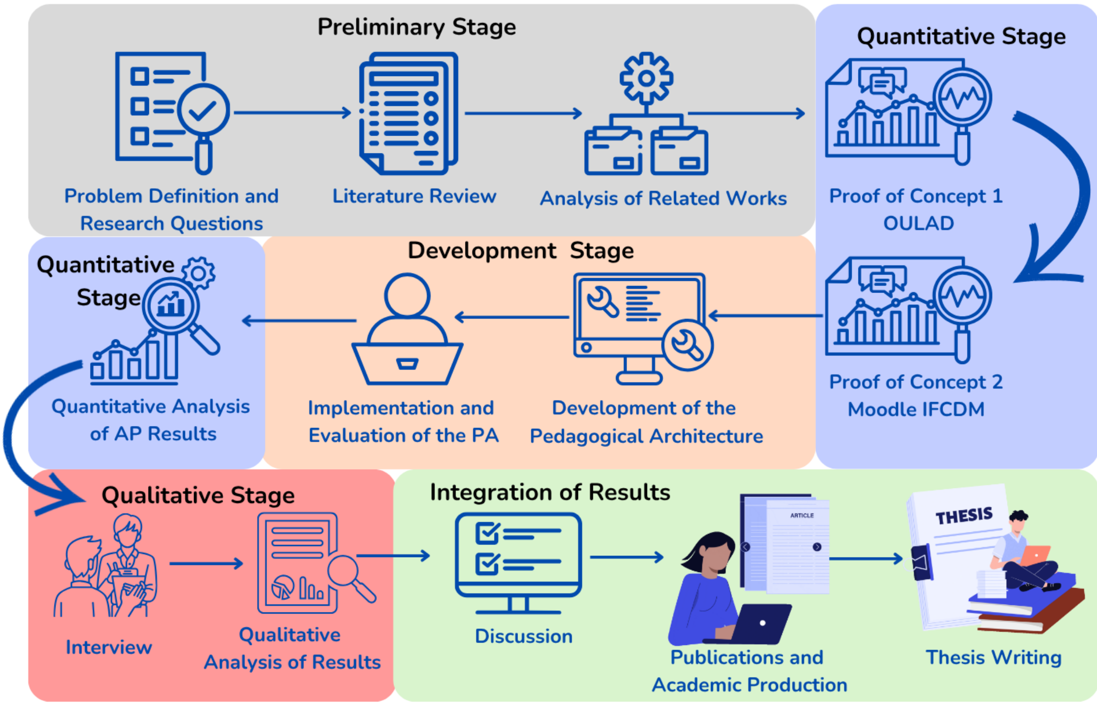



Figure 2 – Research Method
Source: Adapted of Ivankova, Creswell and Stick (2006).

in digital contexts, especially in courses offered through distance education or mediated by educational technologies. However, several studies indicate that many students struggle to plan, monitor, and evaluate their own learning, which negatively impacts their performance and may contribute to dropout rates. Although VLEs store large volumes of data on student behavior, this information is still underutilized in effectively supporting self-regulation processes, thereby limiting the potential of virtual environments as tools for personalized pedagogical support.

Additionally, the analysis of related works revealed that, although there are proposals employing LA and EDM techniques to identify patterns of student engagement, integrated approaches that connect these analyses to the design of pedagogical interventions aimed at fostering self-regulated learning are still scarce. It is observed that a significant portion of the studies focuses on predictive models of academic performance or dropout, without meaningfully exploring the potential for personalized pedagogical support grounded in educational data.

In light of this scenario, this research set out to address three central questions: (i) How does the PA, through its various elements and resources, contribute to the development of SRL skills within the context of the VLE? (ii) How does students' engagement with the resources made available in the PA correlate with their academic performance? (iii)

To what extent do the interpretations of the results obtained from the analyses validate the effectiveness of the proposed PA? The formulation of these questions, grounded in the specialized literature and previous studies in the field, guided the subsequent stages of the investigation, supporting methodological decisions and directing the design of the proposed PA.

Building on the theoretical foundation established through the problem definition, literature review, and related works analysis, the next stage of this research involved conducting a Proof of Concept (PoC) to empirically explore the formulated research questions. This transition from conceptual investigation to practical experimentation enabled the validation of hypotheses regarding the role of SRL in VLEs, providing initial evidence to guide the development of the proposed PA. The following section details the procedures and findings of this PoC phase, which served as a critical step in aligning theory with real-world educational data.

## 4.2 Quantitative Stage - Proof of Concept

The implementation of a PoC was essential to provide both theoretical and empirical support for the design of the proposed PA. A PoC involves the development of a preliminary version of a system or model with the aim of demonstrating its functional feasibility and potential value in a real-world application context (LACHHEB *et al.*, 2025). Within the scope of this research, this stage allowed us to investigate the student behavior patterns in VLE, focusing on identifying signs of the use of SRL strategies, and to conduct an initial assessment of the PA potential to foster greater autonomy and effectiveness in learning processes.

The adoption of the PoC approach provided two key benefits to the research: (i) Theoretical validation, by conceptually supporting the hypotheses related to the use of educational data to foster SRL; and (ii) Empirical grounding, through the application of clustering algorithms and engagement analyses, ensuring that the proposed architecture was based on concrete evidence even prior to its full implementation. In this way, the PoC served as a strategic step in the structuring of the PA, ensuring greater methodological robustness for the study.

The first PoC was conducted using the OULAD, a dataset widely recognized in the literature for its richness and organization. Made available by the Open University in the United Kingdom, it contains data from over 30,000 students and is frequently used in LA research (KUZILEK; HLOSTA; ZDRÁHAL, 2017). This analysis aimed to examine the correlation between students' engagement with educational resources, their academic performance, and the presence of SRL strategies. The results, presented at the Brazilian Symposium on Informatics in Education (LIMA, 2023), revealed two distinct profiles: one group without clear evidence of self-regulation and another exhibiting behaviors consistent with self-regulated learners. Furthermore, a positive correlation was identified between the presence of SRL strategies and academic performance, evidenced by a significantly higher pass rate among students classified as more self-regulated.

In order to validate the findings in a different educational context one more closely aligned with the national reality a second PoC was carried out using data extracted from the Moodle platform of a public educational institution, specifically a post-secondary technical course offered by the Federal Institute of Education, Science and Technology of Southern Minas Gerais. These data were subjected to a rigorous pre-processing procedure to ensure the quality and consistency of the information analyzed. This stage focused on identifying patterns of student engagement and self-regulation in a real teaching environment, contributing to a deeper understanding of the applicability of the PA across different educational scenarios.

Both Proofs of Concept made substantial contributions to the definition of the pedagogical and functional requirements of the PA, ensuring that its features were guided by empirical evidence and aimed at effectively supporting SRL. The learning profiles identified through the analyses served as the foundation for designing personalized pedagogical interventions, including tools for progress monitoring, adaptation of educational resources, and encouragement of the conscious use of SRL strategies. The following sections present the methodological procedures in each of the analyses conducted, which provided the basis for the development of the proposed PA.

### 4.2.1 Proof of Concept 1 - OULAD

In this stage is comprised of five steps, as depicted in Figure 3. Initially, the data was extracted from the OULAD dataset, where all 7 files were downloaded and analyzed. Upon examination of the files, we selected the files that were most relevant to our study: *studentInfo*, *studentVle*, and *vle*. Then, the data was pre-processed to select the relevant attributes for the dataset to be used in the next step. The pre-processed data was subjected to clustering algorithms using Agglomerative Clustering, K-Means, and Expectation Maximization libraries. Finally, the results were analyzed to identify the SRL profiles of the students.

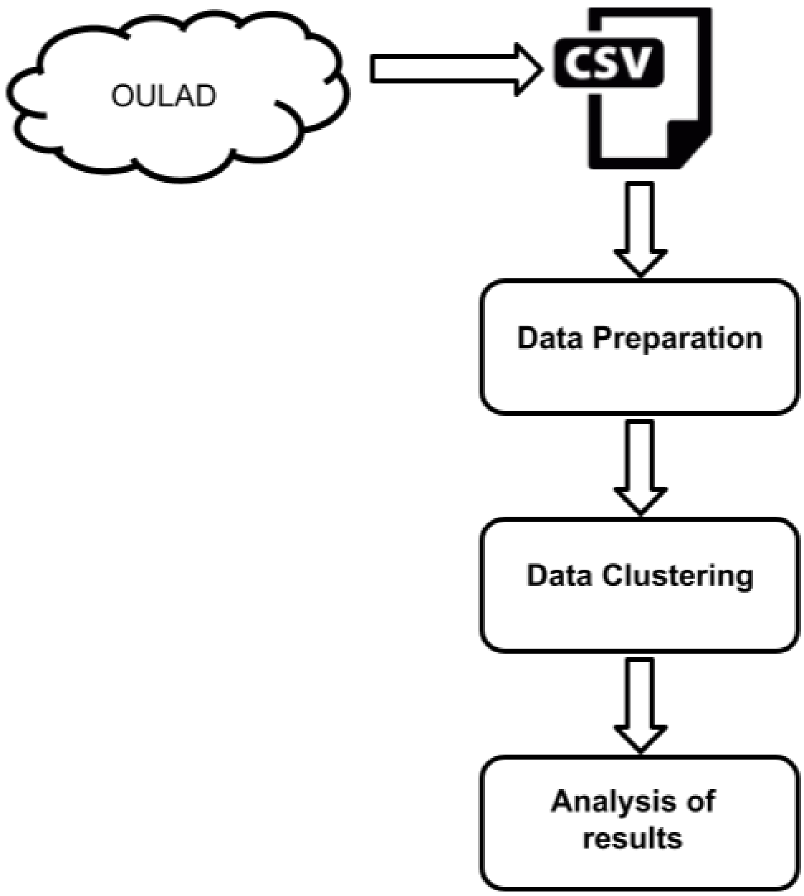



Figure 3 – Workflow using OULAD.

Following a comprehensive analysis of multiple databases, including those referenced in Costa, Dorça and Araújo (2020), CodeBench (LIMA *et al.*, 2021), Khan Academy (LIMA *et al.*, 2021), and edX (COBOS; WILDE; ZALUSKA, 2017), we have determined that the OULAD dataset, provided by the Open University UK, is the most appropriate for the context that we want to analyze in this work since it has more data on different types of interaction and learning resources.

According to Kuzilek, Hlosta and Zdráhal (2017), the dataset is categorized into three sections: demographic, evaluation, and VLE interaction (also referred to as student activities). The module presentation section provides details on the courses, assessments, and materials available in the VLE. The student demographics include attributes such as gender, region, registered module, and final results. The student activities encompass students' interactions with course materials, participation in discussion forums, and performance on quizzes and assignments. This information is particularly relevant for research related to SRL, as it enables the investigation of strategies such as goal-setting, self-monitoring, and self-evaluation. The dataset contains 7 CSV files, presented in Table 7, covering data from 22 courses and 32,593 students. To be used in research, it requires pre-processing and transformation to extract meaningful features for building forecasting models.

Table 7 – Tables of the Open University Learning Analytics Dataset.

| Table name | Records | Description | Table attributes |
|---|---|---|---|
| **studentInfo** | 32593 | Demographic information about the students | code_module, code_presentation, id_student, gender, region, highest_education, imd_band, age_band, num_of_prev_attempts, studied_credits, disability, final_result |
| **vle** | 6365 | Online learning resources and materials | id_site, code_module, code_presentation, activity_type, week_from, week_to |
| **studentVle** | 1.048574 | Student interaction with the VLE resources | code_module, code_presentation, id_student, id_site, date, sum_click |
| **courses** | 22 | Information about the courses | code_module, code_presentation, module_presentation_length |
| **studentRegistration** | 32593 | Registration of the student for a course presentation | code_module, code_presentation, id_student, date_registration, date_unregistration |
| **assessments** | 196 | Assessments for every course presentation | code_module, code_presentation, id_assessment, assessment_type, data, weight |
| **studentAssessments** | 173740 | Assessments submitted by the students | id_assessment, id_student, score, date_submitted, is_banked |

The current study used the dataset extracted from the *studentInfo*, *studentVle*, and *vle*. The interactions within the OULAD dataset can be classified into distinct types based on the learning activities and resources utilized by the students. The interactions are represented by the number of clicks in specific learning resources and activities, such as course notes in the form of HyperText Markup Language (HTML) pages and pdf files, and learning activities in the form of discussion forums and quizzes. The logs capture data on the number of clicks on each resource, this data can be used to understand which resources are most frequently accessed by students and how they engage with different types of content.

Many articles found in the literature have used OULAD to perform student performance prediction (YE; BISWAS, 2014) (HLOSTA; ZDRAHAL; ZENDULKA, 2017) (HUANG *et al.*, 2022). The predictive model presented in Gamie, El-Seoud and Salama (2019) showed that student engagement with digital material has a significant impact on their success throughout the course. In other work, the OULAD dataset has been used to predict the course result (RIAZY; SIMBECK; SCHRECK, 2020) (JHA; GHERGULESCU; MOLDOVAN, 2019) (KARIMI *et al.*, 2020). Araka *et al.* (2022) analyzed students engagement to identify SRL profiles. The profiles collaborate to conclude on final students performances, where self-regulatory profiles tend to be more successful. Zhao *et al.* (2019) grouped the students based on their use of learning materials. The dataset at hand offers a notably wider and diverse range of attributes than other available datasets. This characteristic renders it well-suited for the execution of the proposed research and the investigation of the initial inquiries that were raised.

#### 4.2.1.1 Data Preparation

In this subsection, we present the details of the data pre-processing. After feature extraction, the resulting table was summarized in Table 8. We used the three files mentioned in Table 7, and extracted the data using the unique identification of each student in the entire database, i.e., *id_student*, as the key. The pre-processing of the data and the construction of the final dataset were performed using the Pandas library in Python, which offers user-friendly data structures and data analysis tools for handling and manipulating large datasets (MCKINNEY, 2010).

Table 8 – Summary of the dataset after feature engineering.

| Category | N. Attrib. | Attributes | Type |
|---|---|---|---|
| **Sum click for each VLE activity_type** | 20 | Sum_Clicks_{resources, oucontent, url, homepage, subpage, glossary, forumng, oucollaborate, dataplus, quiz, ouelluminate, sharedsubpage, questionnaire, page, externalquiz, ouwiki, dualpane, repeatactivity, folder, htmlactivity} | Numeric |

The *vle* table contains information about 20 different types of online resources (attributes). Table 9 displays a description of each *vle* attribute. A new feature, the "sum for clicks each activity", was created for each unique resource type based on the *"studentVl"* file. Each row in the table presents the sum of clicks made on each attribute in the VLE by a student. The table ultimately contains 29,741 rows (students) and 20 columns (attributes).

In order to reduce the complexity of the data, the attributes were aggregated. Similar attributes were grouped together, resulting in the reduction of the original dataset from 20 attributes to 5 numerical attributes, as described in Table 10. The attributes *folder*, *sharedsubpage*, and *repeatactivity* were excluded from the dataset due to their low frequency and lack of significance for the current study's context. This resulted in a final dataset with 29,741 rows (students) and 5 columns (attributes).

At this stage, several statistical analyses were conducted. Firstly, the Anderson-Darling test was performed in order to determine if the data followed a normal distribution (ANDERSON; DARLING, 1952). Based on the test results, it was determined that the data did not follow a normal distribution, leading to the use of non-parametric

Table 9 – Description of attributes in the OULAD.

| **Activity** | **Description** |
|---|---|
| **homepage** | Visit the main course page |
| **oucontent** | View course content page |
| **forumng** | Discussion forum usage |
| **subpage** | Manage/view course activities on a page other than the homepage |
| **resource** | Download a document from the course |
| **url** | Click a link to an external site |
| **quiz** | Take a quiz |
| **ouwiki** | Access the course wiki |
| **page** | Non-interactive information page |
| **oucollaborate** | Audio/video conferencing |
| **externalquiz** | Externally-hosted quiz |
| **glossary** | View course glossary |
| **questionnaire** | Access survey form |
| **ouelluminate** | Audio-only conferencing |
| **dualpane** | Side-by-side view of instructions and related content |
| **dataplus** | Interact with a toy SQLite database |
| **htmlactivity** | Interactive HTML page |
| **folder** | View folder containing related activities |
| **sharedsubpage** | View page shared from another course |
| **repeatactivity** | Activities repeated from earlier in the course |

*Source:* Jha, Ghergulescu and Moldovan (2019)

Table 10 – Aggregation of attributes.

| **Activity** | **Grouping** |
|---|---|
| **collaborative** | forumng, outcollaborate and ouelluminate |
| **activities** | quiz, exam and questionnarie |
| **access** | homepage, resource, ouwiki, page, and htmlactivity |
| **resource_ext** | externalquiz and url |
| **views** | glossary, outcontent, dualpane, subpage, and dataplus |

variance tests. In order to identify how one variable behaves when another variable is varying, a correlation matrix of the variables was constructed using the non-parametric Spearman's correlation coefficient. This coefficient measures the monotonic relationship between two variables (SPEARMAN, 1904).

#### 4.2.1.2 Data Clustering

The performance of three clustering algorithms was investigated to track the student's SRL profile, considering clicks on the resources available in the VLE: K-Means, EM, and Agglomerative Clustering. We chose from different categories of algorithms: partitional, hierarchical, and model-based. The algorithms were applied with resources from scikit-learn libraries[1].

K-Means is a partitioning algorithm that divides a set of $X$ of $n$ samples into $K$ disjoint groups, each described by an average $u$ of the samples in the group. This mean is called the centroid of the group (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Agglomerative Clustering is part of the family of hierarchical algorithms and uses a bottom-up approach to perform the clustering, that is, each element of the dataset starts in a group and, at each step, the pairs of elements merge according to their proximity (KAUFMAN; ROUSSEEUW, 2009).

EM is designed to estimate the maximum likelihood parameters of a statistical model in situations where the equations cannot be solved directly. It is an iterative technique composed of two main steps: the E step (Expectation) and the M step (Maximization). During the E step, the algorithm computes the expectation of the conditional likelihood function of the parameters given the observations and the current parameter estimates. In the M step, this conditional likelihood function is maximized with respect to the parameters, yielding updated estimates for the model. These two steps are performed iteratively until the parameters converge to stable values (MOON, 1996).

To summarize, the first PoC comprised the complete process of data extraction, preprocessing, feature engineering, and clustering using three algorithms K-Means, Agglomerative Clustering, and EM. This stage aimed to identify patterns of engagement and SRL behaviors among students in the OULAD dataset, providing the empirical basis for the following stages of this research. The outcomes and comparative analysis of the clustering techniques are presented and discussed below.

#### 4.2.1.3 Results

The experimental results obtained from the application of three clustering algorithms are presented and analyzed in this part of the study. The primary objective was to identify the most appropriate algorithm and determine the optimal number of clusters. To support this evaluation, internal validation metrics were employed, namely the Silhouette Coefficient, Dunn's Index, Calinski–Harabasz Score, and Davies–Bouldin Index.

The Silhouette Coefficient is a metric that measures how well each data point fits into its assigned cluster based on the distance between the data point and other points within its cluster, as well as the distance between the data point and the points in other clusters (ROUSSEEUW, 1987). Dunn's Index measures the distance between the closest clusters relative to the average size of the clusters (KAUFMAN; ROUSSEEUW, 2009). Calinski-Harabasz is a measure of the density and separability between groups, while Davies-Bouldin measures the similarity between the group and its closest group (FURLANETTO *et al.*, 2022).

[1] https://scikit-learn.org/stable/

Table 11 – Optimal Algorithm and Cluster Evaluation Results.

| | | **Clusters** | | | |
|---|---|---|---|---|---|
| **Algorithm** | **Validation measure** | **2** | **3** | **4** | **5** |
| **K-Means** | Silhouette Coefficient | 0.64 | 0.62 | 0.52 | 0.54 |
| | Dunn Index | 0.0021 | 0.0025 | 0.0018 | 0.0015 |
| | Calinski-Harabasz | **20419** | 17165 | 16465 | 15101 |
| | Davies-Bouldin | **0.936** | 0.954 | 1.01 | 1.002 |
| **Agglomerative Clustering** | Silhouette Coefficient | **0.65** | 0.64 | 0.44 | 0.45 |
| | Dunn Index | **0.0038** | 0.0035 | 0.0012 | 0.0012 |
| | Calinski-Harabasz | 18170 | 14837 | 13929 | 12571 |
| | Davies-Bouldin | 0.945 | 0.878 | 1.112 | 1.145 |
| **Expectation Maximization** | Silhouette Coefficient | 0.36 | 0.21 | 0.10 | 0.09 |
| | Dunn Index | 0.0007 | 0.0003 | 0.00016 | 0.00011 |
| | Calinski-Harabasz | 10040 | 8382 | 6433 | 5840 |
| | Davies-Bouldin | 1.120 | 1.376 | 1.500 | 1.660 |

Table 12 – Statistics per Cluster Generated.

| | **Agglomerative** | | **K-Means** | |
|---|---|---|---|---|
| **Attribute** | **Cluster0 (N=29685)** | **Cluster1 (N=56)** | **Cluster0 (N=25165)** | **Cluster1 (N=4576)** |
| **collaborative** | $311.24 \pm 576.52$ | $8442.64 \pm 1576.74$ | $191.26 \pm 295.85$ | $1070.57 \pm 1363.12$ |
| **activities** | $298.20 \pm 555.62$ | $884.71 \pm 1240.22$ | $160.36 \pm 273.14$ | $1063.38 \pm 961.80$ |
| **access** | $367.86 \pm 447.73$ | $3760.67 \pm 2294.64$ | $248.17 \pm 237.33$ | $1067.59 \pm 791.44$ |
| **resource_ext** | $25.72 \pm 43.23$ | $179.30 \pm 165.51$ | $18.62 \pm 237.33$ | $66.66 \pm 77.52$ |
| **views** | $592.34 \pm 844.06$ | $1744.51 \pm 1875.97$ | $322.12 \pm 342.17$ | $2092.45 \pm 1176.59$ |

Table 11 presents the cluster validation measures for the different clustering algorithms used in this study. The results show that Agglomerative Clustering had the best performance in terms of the Silhouette Coefficient (0.65) and Dunn Index (0.0038). On the other hand, K-Means had the best performance in terms of Calinski-Harabasz (20419) and Davies-Bouldin (0.936) measures. Therefore, further analysis of the grouping data is necessary to determine the algorithm with the best performance.

We performed a descriptive analysis of the clusters, as shown in Table 12, which provides the mean and standard deviation of each attribute. After analyzing the clusters generated by each algorithm, we chose to use K-Means because the sizes of the generated clusters were more balanced than those obtained with Agglomerative. Uneven cluster sizes can indicate that the data is not well-clustered.

Figure 4 highlight the mean values of each attribute in the clusters generated using the K-Means algorithm. To determine if there were statistically significant differences between the means of the attributes, a non-parametric Kruskal-Wallis test was conducted, as the data did not follow a normal distribution, which was confirmed by the Anderson-Darling test. The Kruskal-Wallis test confirmed the statistical significance of the differences between the clusters.

The clusters depicted in Figure 4 reveal the discrepancies between the resources utilized

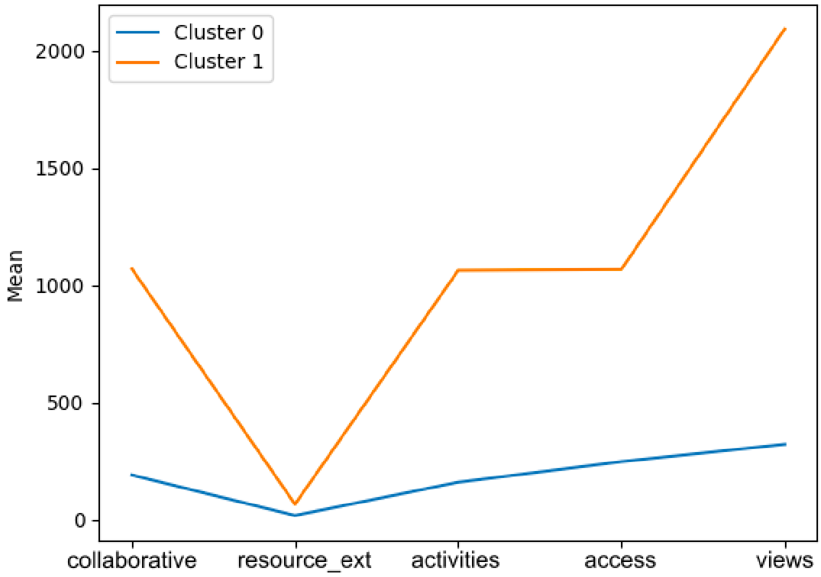



Figure 4 – K-Means - mean in each attribute.

by the groups. Specifically, **Cluster 1** exhibits higher mean values across all attributes, with *"views"* and *"access"* standing out. This suggests that the students in **Cluster 1** were more actively involved in utilizing the resources available on the VLE. This finding is consistent with Zimmerman and Martinez-Pons (1986), which indicates that SRL involves proactive students who take responsibility for their own learning process. Moreover, students who engage in more interactions on the educational platform tend to achieve better academic outcomes.

One can also highlight that collaborative resources, such as forums, had a higher average in **Cluster 1**. These resources are important for fostering student-teacher and student-student interaction. By using these resources, students can engage in discussions, which can help them self-monitor and define strategies for performing tasks (KITSANTAS, 2013). Interactions (*"activities"*) were grouped with clicks in quizzes and questionnaires. Students with self-regulation characteristics tend to self-evaluate more often (KITSANTAS, 2013). This result further indicates evidence of students' self-regulated behavior in the present analysis.

The *"collaborative"* attribute represents three resources in the VLE: *"forumng"* , *"outcollaborate"*, and *"ovelluminate"*. These resources promote collaboration among students by providing spaces for discussions, group projects, and online learning sessions. We grouped these resources into a single variable to investigate the relationship between collaborative interaction and academic performance. Collaboration can help students develop cognitive, metacognitive, and motivational skills by allowing them to share knowledge, learn from others, and receive feedback. This strategy also supports students in reflecting on their learning process and monitoring their performance. As a result, it enables adjustments to learning strategies when necessary.

The *"resource_ext"* attribute was created by counting student interactions with two specific VLE resources: *"external_quiz"* and *"url"*. The former allows students to answer quizzes from sources outside the VLE, while the latter allows access to external websites relevant to the course content. The attribute was used to investigate whether higher interaction with external resources corresponded to better academic performance. Greater interaction with external resources may indicate a more active and autonomous learning profile among students.

The *"activities"* was derived from three assessment resources: *"quiz"*, *"exam"*, and *"questionnaire"*, and measures the extent of student engagement with these activities within the VLE. This variable holds potential in identifying students' SRL profiles, as students who participate more in assessment activities may possess a proactive learning profile, driven by a desire to continually improve their academic performance. These students may seek frequent feedback, self-monitor their learning progress, and adjust their study strategies to achieve better results.

The *"access"* attribute is a measure of the number of times students access resources within the VLE, including *"homepage"*, *"resource"*, *"ouwiki"*, *"page"*, and *"htmlactivity"*. This attribute may serve as an indicator of students' level of engagement with the course content and can be utilized to evidence diverse learning profiles. Furthermore, access to resources can be related to the SRL profile, which encompasses active information seeking behaviors that are pertinent to learning. As such, the "access" variable can provide insights into how students' SRL strategies influence their engagement within the VLE.

Lastly, the *"views"* attribute depicts the quantity of clicks initiated by each student on five particular resources of the VLE, namely *"glossary"*, *"outcontent"*, *"dualpane"*, *"subpage"*, and *"dataplus"*. This attribute serves as a gauge of students' involvement with the VLE and can facilitate the identification of patterns of SRL behavior among students.

Finally, correlation analysis was used to identify the association between the SRL profiles and students' academic performance. The chi-square test was carried out to establish the correlation between the SRL profiles formed by K-Means grouping and students' final results. After calculating the ($p\_value = 0.00$), we can conclude that there is a significant relationship between the SRL profiles and the students' final results. Figure 5 shows the final average of approved and failed in each cluster. Therefore, we can conclude that students with more interactions evidence a self-regulated profile that tends to produce better performance results.

The findings presented in Figure 5 suggest that there is a positive correlation between student engagement and academic performance. Students who had higher levels of interaction with the learning resources evidence a SRL profile that was associated with better academic outcomes. The higher proportion of successful students in **Cluster 1**, as compared to **Cluster 0**, suggests that greater engagement with the VLE resources may have contributed to their academic success. Therefore, it can be concluded that student engagement is an important factor in promoting SRL and achieving academic success.

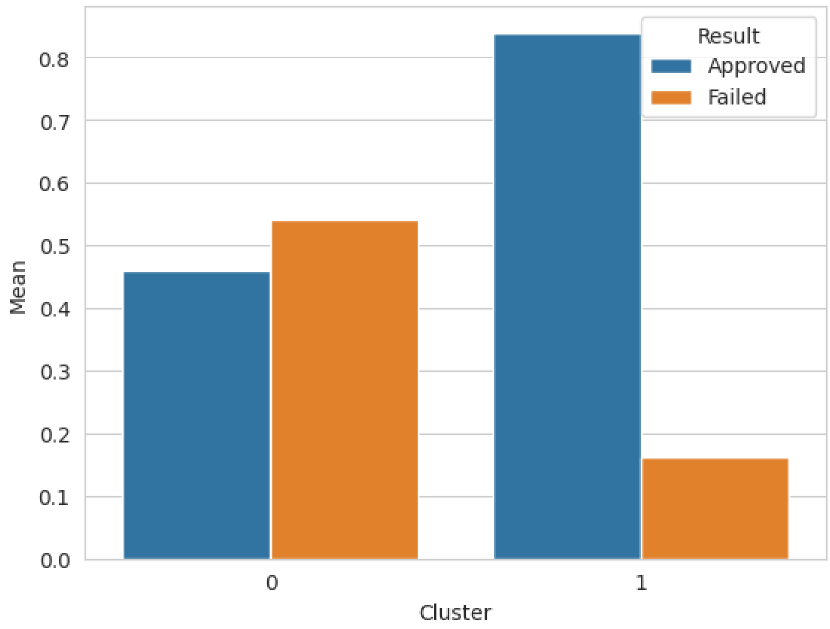



Figure 5 – Relationship Profile SRL based on Final Results.

### 4.2.2 Proof of Concept 2 - Moodle IFCDM

The second PoC was conducted using data from the Moodle of the Federal Institute of Education, Science and Technology of Southern Minas Gerais - Campus Carmo de Minas (IFCDM). This phase was essential to validate the findings of the first analysis in a new educational context, ensuring that the behavior patterns and SRL strategies identified in the OULAD dataset were also observable in a Brazilian public educational institution. This study was approved by the Human Research Ethics Committee[2].

This second PoC is justified by the need to evaluate the applicability of the results in a different learning environment, with distinct characteristics and student profiles. The use of Moodle data at IFCDM allowed the exploration of student behavior in a subsequent technical course, providing a broader perspective on student interactions and engagement in a professional and technological education setting. Furthermore, by analyzing data from a Brazilian public institution, it was possible to identify potential variations in SRL strategies that may not have been evident in the context of the Open University, a predominantly distance learning environment in the United Kingdom. Figure 6 presents the workflow applied during this PoC.

Firstly, user interaction data logs with Moodle were extracted, resulting in three main files: 1) Grade Report, which contains students' grades for each activity and the final grade in the course; 2) Logs Report, which records all events performed by users in the course, including the use of available resources; and 3) Report from the plugin Configurable Reports[3], which details the amount of time each user spent on each course.

[2] CAAE number: 78890524.5.0000.8158
[3] https://moodle.org/plugins/block_configurable_reports

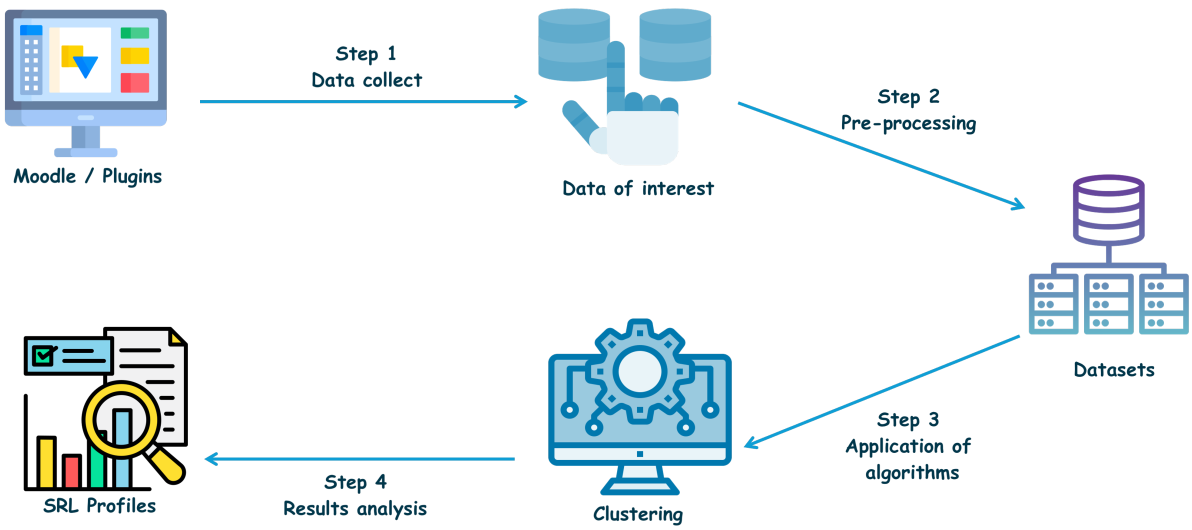



Figure 6 – Workflow using Moodle IFCDM.

Subsequently, the data were pre-processed to select the relevant attributes and compose the dataset to be used in the following stage. After pre-processing, the data were subjected to clustering algorithms, specifically Agglomerative Clustering, K-Means, and HDBSCAN, using appropriate libraries for this purpose. Finally, the obtained results were analyzed to evidence the SRL profiles of the students, providing a deeper understanding of interaction patterns and performance in the VLE.

#### 4.2.2.1 Moodle IFCDM

Moodle is a widely used VLE designed for academic activities. This system provides a comprehensive platform for the administration and facilitation of educational experiences, offering a variety of synchronous and asynchronous Web tools that promote interaction and communication among users. Moodle stands out as a powerful solution for educational institutions, enabling teachers and students to explore resources such as discussion forums, chats, assignment submissions, quizzes, and more. Its flexibility and adaptability make it a popular choice for efficient teaching and learning management, facilitating access to learning resources anywhere and anytime.

Moodle is an open-source project, built under the General Public License (GNU), representing an open, free, and cost-free proposition that can be installed and used by any educational institution. Its official documentation is readily available on the website [4]. Furthermore, platform users also have the opportunity to become developers themselves, creating plugins to collaborate with the project and further enhance its functionalities.

In this phase, we used real data from Moodle of a subsequent technical course at a public educational institution during the 2021/2 semester, with the aim of mapping the

[4] https://moodle.org

SRL profiles of students. To achieve this, we employed EDM techniques that were already applied in the first phase of the research.

The data collected from Moodle correspond to eight subjects from the first module of a Subsequent Technical Course in Administration offered through distance education. The analyzed subjects were: Customer Service and Consumer Rights (*COURSE_1*), Entrepreneurship (*COURSE_2*), Computing and Spreadsheets (*COURSE_3*), Introduction to Administration (*COURSE_4*), Labor and Social Legislation (*COURSE_5*), Business Model Canvas (*COURSE_6*), Sustainable Business (*COURSE_7*), and Recruitment and Selection (*COURSE_8*).

For each subject, three .csv files were obtained and combined using the unique identifier of each student. This resulted in a consolidated file containing information on all active students in the subjects, including the quantification of all log events and the access time to the platform for each user. Table 13 a) presents a description of the collected data, covering the analyzed subjects, the number of logs generated, and the number of students enrolled in each subject.

Table 13 – Data collected and creation of datasets.

| **a) Data collected** | | | **b) Creation of datasets** | | | |
|---|---|---|---|---|---|---|
| **Subject Tag** | **Logs** | **T-E** | **T-A** | **Attributes excluded** | **Attributes aggregates** | **Final Dataset** |
| *COURSE_1* | 62564 | 413 | 25 | 6 | 3 | 413 records × 5 attributes |
| *COURSE_2* | 88717 | 456 | 27 | 9 | 3 | 456 records × 5 attributes |
| *COURSE_3* | 120021 | 445 | 40 | 9 | 4 | 445 records × 6 attributes |
| *COURSE_4* | 64521 | 409 | 29 | 1 | 4 | 409 records × 6 attributes |
| *COURSE_5* | 79869 | 418 | 27 | 6 | 3 | 418 records × 5 attributes |
| *COURSE_6* | 64518 | 405 | 35 | 7 | 4 | 405 records × 6 attributes |
| *COURSE_7* | 60972 | 413 | 25 | 6 | 3 | 413 records × 5 attributes |
| *COURSE_8* | 111919 | 430 | 25 | 5 | 3 | 430 records × 5 attributes |

Legend: T-E: Total Students; T-A: Total Attributes.

#### 4.2.2.2 Pre-Processing

This phase details the data pre-processing process. The pre-processing and construction of the final dataset were carried out using the Pandas library in Python, which provides intuitive data structures and robust analysis tools for handling and manipulating large volumes of data (MCKINNEY, 2010).

Moodle collects a variety of attributes in its activity logs, the quantity and types of which can vary depending on the specific settings and additional plugins installed. In the Moodle used for this study, 16 common attributes are generated across all courses, as shown in Table 14. Attributes 1, 2, 3, and 16 are not attributes that represent activity logs in the system. Initially, for the construction of the final dataset, the common attributes and the specific features of the resources offered in each course were considered.

Table 14 – Description of Common Attributes.

| **Attribute** | **Description** |
|---|---|
| [1]:mat | attribute with the student's registration number. |
| [2]:nome | the student's name. |
| [3]:nota | attribute with the student's final grade in the course. |
| [4]:algum_conteudo_publicado | indicates whether the student has published any content in a course activity. |
| [5]:post_criado | attribute shows whether the student created a post in a specific activity, such as a forum, blog, or other discussion area. |
| [6]:curso_visto | demonstrates whether the student viewed a specific course. |
| [7]:discussao_visualizada | indicates whether the student viewed a specific discussion, such as in a forum or discussion group. |
| [8]:modulo_curso_visualizado | reveals whether the student has viewed a specific module within a course. |
| [9]:resumo_tentativa_questionario_visualizada | shows whether the student has viewed the summary of a quiz attempt. |
| [10]:tentativa_questionario_visualizada | indicates whether the student viewed a specific attempt at a quiz. |
| [11]:conclusao_atividade_curso | indicates whether a student has completed a specific activity within a course. |
| [12]:relatorio_de_notas_visualizado | reveals whether the student has viewed the grade report in a course. |
| [13]:tentativa_questionario_entregue | shows whether the student turned in a specific attempt at a quiz. |
| [14]:tentativa_questionario_iniciada | shows whether the student has initiated a specific attempt at a quiz. |
| [15]:tentativa_questionario_revisada | reveals whether the student reviewed a specific attempt at a quiz. |
| [16]:tempo | total time in seconds that the student dedicated to the course. |

Subsequently, the files were analyzed separately to identify the distinct events in each course and evaluate their impact on the creation of the final dataset. Each course offered in the program has different configurations for activities or resources. Therefore, in addition to the common attributes presented in Table 14, other attributes were listed. Statistical techniques were employed to analyze these additional attributes, resulting in the exclusion or combination of attributes as necessary.

To identify how one variable behaves when another variable is varying, a correlation matrix of the variables was constructed using Spearman's non-parametric correlation coefficient, since the data were previously analyzed with the Kolmogorov-Smirnov normality test and it was observed that the data do not follow a normal distribution across all courses with $p\_value < 0.05$.

Spearman's coefficient measures the monotonic relationship between two variables (SPEARMAN, 1904). Spearman's correlation was performed on all datasets, and those attributes with a strong correlation ($> 0.7$) were aggregated (ZAR, 2005). Table 13b) highlights the total number of attributes in each dataset, where each attribute represents a type of log in the system, the number of attributes that were excluded, the number of attributes present in the final dataset that were constructed through the aggregation criterion of other similar attributes and/or those with a strong correlation, and the final configuration of these.

Table 15 presents the final attributes of the datasets and their respective descriptions. The pre-processing resulted in refined datasets ready for the analysis stage. With these data prepared, the next step involves the data clustering process, where clustering techniques will be applied to identify patterns and usage profiles among the students.

Table 15 – Description of the attributes of the final datasets.

| **Attributes** | **Description** |
|---|---|
| *id* | user identifier attribute. |
| *publicacao* | aggregated attribute representing the number of posts made by the student throughout the course, including forum posts and comments on activities. |
| *visualizacoes* | aggregated attribute representing the number of times the student views any resource during the course, including views of modules, quizzes, materials, activities, courses, comments, grades, and others. |
| *atividades_concluídas* | quantidade de atividades do curso concluídas e/ou atualizadas. |
| *questionario* | aggregated attribute representing the number of times the student performed an action related to quizzes in the course, including starting, reviewing, and submitting. |
| *submissao* | aggregated attribute present only in subjects 3, 4, and 6. It represents the number of times the student submitted files or online texts to the teacher. |
| *tempo* | time in seconds that the student spent on the course using the platform. |

#### 4.2.2.3 Data Clustering

The effectiveness of three clustering algorithms is analyzed in highlighting the SRL profile of students using the logs provided by Moodle: K-Means, HDBSCAN, and Agglomerative Clustering. K-Means was selected for being an efficient partitioning algorithm that divides the data into groups based on the mean of the points, facilitating the identification of well-defined clusters. Agglomerative Clustering, on the other hand, is a hierarchical method that builds a clustering tree based on data proximity, allowing for a detailed analysis of the relationships between data points. HDBSCAN was chosen for its ability to handle data with varying densities and identify outliers, which is essential for capturing the diversity of student behaviors in the VLE. The combination of these algorithms enabled a robust and comprehensive analysis of SRL profiles, providing valuable *insights* into students' engagement patterns and academic performance.

The selection of clustering algorithms was based on the nature of the data and the research objectives. The algorithms were implemented using the scikit-learn libraries[5].The results of the clustering algorithms revealed different perspectives on the SRL profiles of students. K-Means was efficient in dividing the data into distinct groups based on general engagement patterns, while Agglomerative Clustering provided a hierarchical view that highlighted the proximity between subgroups of students with similar behaviors. HDBSCAN, on the other hand, was effective in identifying both the main groups and the outliers, offering a more detailed analysis of students whose learning patterns deviated from the norm.

To conclude, this second PoC systematically encompassed data extraction from Moodle (logs, grades, and time), integration and preprocessing with attribute aggregation, and the application of three clustering techniques K-Means, Agglomerative Clustering, and HDBSCAN chosen to capture complementary structures (partitional, hierarchical, and density-based) in students' behaviors. This phase was designed to validate and extend the findings from OULAD in a Brazilian public-education context, providing a consistent empirical basis to characterize SRL related engagement in a real institutional setting. The comparative outcomes of the algorithms and the characterization of the resulting profiles are presented and discussed below.

#### 4.2.2.4 Results

The experimental results generated by the three clustering algorithms were analyzed to identify the most appropriate algorithm and determine the optimal number of clusters. For this evaluation, the validation metrics used were the Silhouette Coefficient, Calinski-Harabasz index, and Davies-Bouldin index. In the case of the HDBSCAN algorithm, the number of identified outliers was also considered. Results from this algorithm that showed similar validation indices to those of the other algorithms but identified a high number of records as outliers were excluded from the analysis.

The Silhouette Coefficient is a metric that evaluates how well each data point fits within its assigned cluster, taking into account the distance between the data point and

[5] https://scikit-learn.org/stable/

other points within the same cluster, as well as the distance between the data point and points in different clusters (ROUSSEEUW, 1987). This metric was chosen because it provides a comprehensive assessment of the quality of the clustering, considering both internal cohesion and separation between neighboring clusters.

The Calinski-Harabasz (C-H) index measures the density and separability between groups, helping to identify well-defined and densely compacted clusters with clear separations. On the other hand, the Davies-Bouldin (D-B) index measures the similarity between a cluster and its nearest cluster (FURLANETTO *et al.*, 2022), and was selected for its ability to highlight distinct and well-separated clusters.

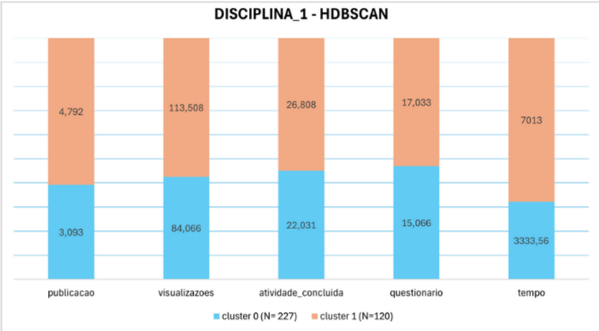
Table 16 presents the best clustering validation metrics for the different clustering algorithms used in this study. Now, Table 17 shows the average differences and standard deviations for each attribute of the clusters identified by these algorithms in each subject, supported by the non-parametric Kruskal-Wallis significance test to verify statistical significance in the observed means. Figure 7 presents the mean differences of each attribute between the clusters for each subject, highlighting how the mean values vary among the identified groups.

Table 16 – Clustering validation metrics for different algorithms.

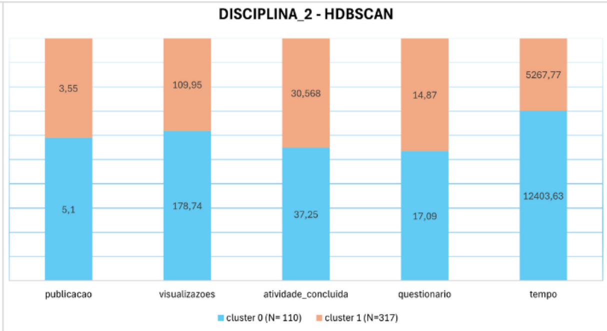
| **Subject Tag** | **Best algorithm** | **Silhouette** | **C-H** | **D-B** | **Generated clusters** |
|---|---|---|---|---|---|
| *COURSE_1* | HDBSCAN | 0.656 | 892.35 | 0.378 | cluster0(N=227)<br>cluster1(N=120)<br>Outliers(N=66) |
| *COURSE_2* | HDBSCAN | 0.695 | 1043.02 | 0.495 | cluster0(N=110)<br>cluster1(N=317)<br>Outliers(N=29) |
| *COURSE_3* | K-Means | 0.685 | 946.39 | 0.494 | cluster0(N=355)<br>cluster1(N=90) |
| *COURSE_4* | HDBSCAN | 0.691 | 1230.68 | 0.400 | cluster0(N=179)<br>cluster1(N=143)<br>Outliers(N=87) |
| *COURSE_5* | K-Means | 0.680 | 787.64 | 0.563 | cluster0(N=308)<br>cluster1(N=110) |
| *COURSE_6* | K-Means | 0.685 | 742.29 | 0.521 | cluster0(N=335)<br>cluster1(N=70) |
| *COURSE_7* | K-Means | 0.699 | 710.37 | 0.487 | cluster0(N=361)<br>cluster1(N=52) |
| *COURSE_8* | HDBSCAN | 0.701 | 948.99 | 0.517 | cluster0(N=131)<br>cluster1(N=219)<br>Outliers(N=80) |

C-H: Calinski-Harabasz Index; D-B: Davies-Bouldin Index.

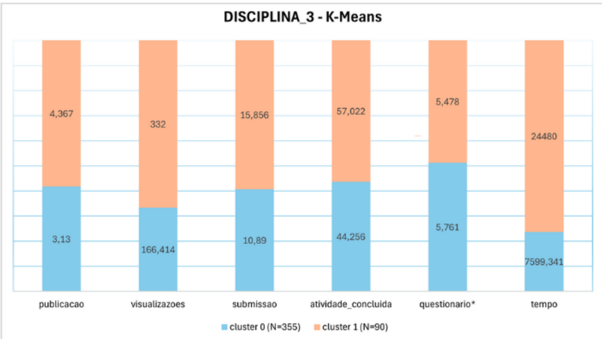
The best result for *COURSE_1* was obtained by the HDBSCAN algorithm. Both HDBSCAN and K-Means achieved the same Silhouette measure, but HDBSCAN was able to better separate the data by identifying 66 records as outliers, resulting in higher values for the C-H and D-B indices. Additionally, the mean differences between the identified clusters were statistically significant across all attributes. For *COURSE_2*, HDB-



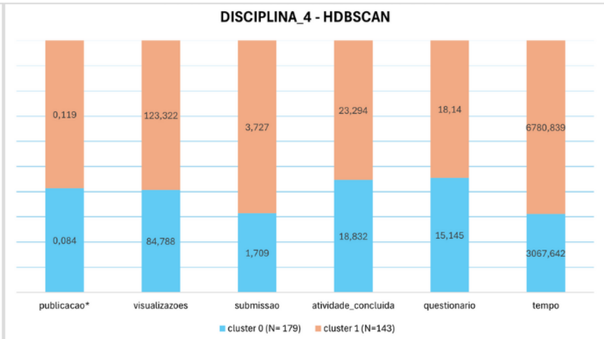

a) Best result *dataset1*



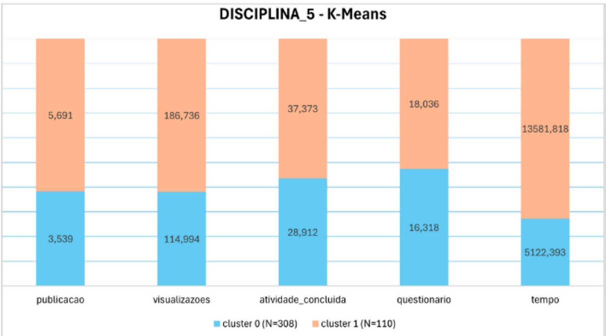

b) Best result *dataset2*



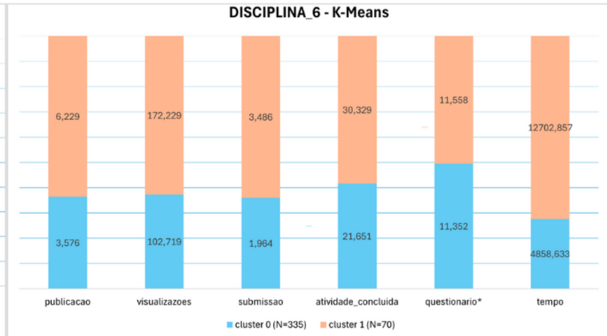

c) Best result *dataset3*



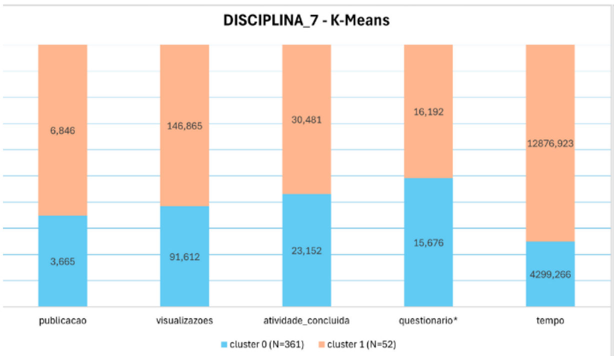

d) Best result *dataset4*



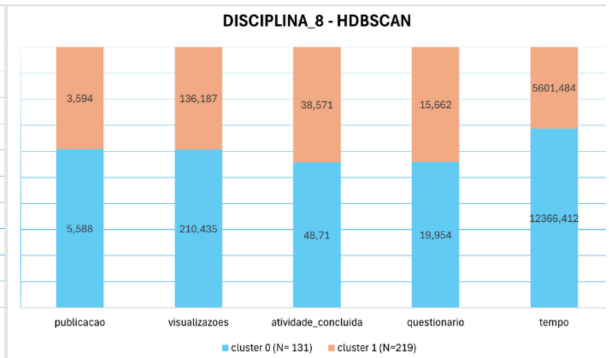

e) Best result *dataset5*




f) Best result *dataset6*




g) Best result *dataset7*




h) Best result *dataset8*

* atribute with *p_value>0.05 by* Kruskal Wallis test

Figure 7 – Comparison of the mean values of each attribute of the clusters generated in the course subject datasets.

SCAN also emerged as the best option, with superior validation measures and statistically significant cluster separation across all attributes.

Table 17 – Comparison of the mean values and standard deviations of each attribute of the clusters generated in the datasets.

| **COURSE 1 - HDBSCAN** | | | **COURSE 2 - HDBSCAN** | | |
|---|---|---|---|---|---|
| **Att.** | Cluster 0 N=227 | Cluster 1 N=120 | **Att.** | Cluster 0 N=110 | Cluster 1 N=317 |
| *publi.* | 3.09 ± 2.63 | 4.79 ± 2.854 | *publi.* | 5.1 ± 2.78 | 3.55 ± 2.38 |
| *visua.* | 84.06 ± 27.85 | 113.50 ± 27.75 | *visua..* | 178.74 ± 40.38 | 109.95 ± 42.88 |
| *ativ.* | 22.03 ± 17.98 | 26.80 ± 16.55 | *ativ.* | 37.25 ± 19.08 | 30.56 ± 19.02 |
| *quest.* | 15.06 ± 5.36 | 17.03 ± 4.43 | *quest.* | 17.09 ± 5.09 | 14.87 ± 5.96 |
| *tempo* | 3333.56 ± 1322.69 | 7013 ± 333.99 | *tempo* | 12403.63 ± 1789.25 | 5267.77 ± 2057.11 |
| **COURSE 3 - K-Means** | | | **COURSE 4 - HDBSCAN** | | |
| **Att.** | Cluster 0 N=355 | Cluster 1 N=90 | **Att.** | Cluster 0 N=179 | Cluster 1 N=143 |
| *publi.* | 3.13 ± 2.23 | 4.36 ± 2.57 | *publi.** | 0.08 ± 0.66 | 0.11 ± 0.76 |
| *visua.* | 166.41 ± 80.46 | 332 ± 110.66 | *visua.* | 84.78 ± 28.55 | 123.32 ± 31.54 |
| *ativ.* | 44.25 ± 28.09 | 57.02 ± 23.02 | *ativ.* | 18.83 ± 16.76 | 23.29 ± 14.88 |
| *quest.** | 5.76 ± 4.41 | 5.47 ± 3.18 | *quest.* | 15.14 ± 5.40 | 18.14 ± 4.26 |
| *tempo* | 7599.34 ± 3870.41 | 24480 ± 6876.74 | *tempo* | 3067.64 ± 112.20 | 6780.83 ± 643.52 |
| *sub.* | 10.89 ± 5.94 | 15.85 ± 7.27 | *sub.* | 1.70 ± 2.18 | 3.72 ± 2.96 |
| **COURSE 5 - K-Means** | | | **COURSE 6 - K-Means** | | |
| **Att.** | Cluster 0 N=308 | Cluster 1 N=110 | **Att.** | Cluster 0 N=335 | Cluster 1 N=70 |
| *publi.* | 3.53 ± 2.86 | 5.69 ± 3.59 | *publi.* | 3.57 ± 2.80 | 6.22 ± 2.92 |
| *visua.* | 114.99 ± 42.68 | 186.73 ± 55.28 | *visua.* | 102.71 ± 41.07 | 172.22 ± 50.52 |
| *ativ.* | 28.91 ± 20.81 | 37.37 ± 17.89 | *ativ.* | 21.65 ± 16.91 | 30.32 ± 12.66 |
| *quest.* | 16.31 ± 6.13 | 18.03 ± 4.55 | *quest.** | 11.35 ± 4.42 | 11.55 ± 3.18 |
| *tempo* | 5122.39 ± 2040.92 | 13581.81 ± 4022.26 | *tempo* | 4858.63 ± 2079.25 | 1702.85 ± 2632.92 |
| | | | *sub.* | 1.96 ± 1.82 | 3.48 ± 2.25 |
| **COURSE 7 - K-Means** | | | **COURSE 8 - HDBSCAN** | | |
| **Att.** | Cluster 0 N=361 | Cluster 1 N=52 | **Att.** | Cluster 0 N=131 | Cluster 1 N=219 |
| *publi.* | 3.66 ± 2.90 | 6.84 ± 4.23 | *publi.* | 5.58 ± 3.60 | 3.59 ± 2.83 |
| *visua.* | 91.61 ± 32.34 | 146.86 ± 43.43 | *visua.* | 210.43 ± 51.81 | 136.18 ± 53.47 |
| *ativ.* | 23.15 ± 17.10 | 30.48 ± 14.20 | *ativ.* | 48.71 ± 26.07 | 38.57 ± 29.47 |
| *quest.** | 15.67 ± 5.11 | 16.19 ± 4.84 | *quest.* | 19.95 ± 6.24 | 15.66 ± 6.34 |
| *tempo* | 4299.26 ± 2063.11 | 12876.92 ± 2765.76 | *tempo* | 12366.41 ± 1784.77 | 5601.48 ± 2090.63 |

*publi.: publicacao; visua.: visualizacao; ativ.: atividade_concluida; quest.: questionario; sub.: submissao*
* attribute with *p_value* > 0.05 by Kruskal-Wallis test

In *COURSE_3*, the K-Means and HDBSCAN algorithms yielded similar results, with K-Means showing superior values across the three validation measures. As shown in Table 17, the attribute *"questionario"* did not achieve statistical significance between *cluster0* and *cluster1*. However, in this course, the teacher opted to use other assessment resources, such as online text submissions and files, which were aggregated in the attribute *"publicacao"*. This may have contributed to this result. In *COURSE_4*, the K-Means and HDBSCAN algorithms were close in the Silhouette Coefficient, but HDBSCAN achieved better results on the C-H and D-B indices. Additionally, the mean values observed in *cluster1* are higher than those in *cluster0*, with statistical significance across all attributes.

*COURSE_5*, *COURSE_6*, and *COURSE_7* showed superior results using the K-Means algorithm. In the cases of *COURSE_5* and *COURSE_6*, the Silhouette measures were similar for K-Means and HDBSCAN, while the C-H and D-B indices were better for HDBSCAN. However, the presence of a significant number of outliers in HDBSCAN compromised the analysis. In *COURSE_7*, both K-Means and HDBSCAN presented similar results in terms of separation and statistical significance, but K-Means achieved better results in the Silhouette and C-H indices.

It is noted that for *COURSE_6*, the attribute *"questionario"* again does not show statistical significance, but the attribute *"submissao"* included in the *dataset* indicates that the teacher also used other assessment resources besides the quiz. In *COURSE_7*, it can be observed that the attribute *"questionario"* did not have a statistically significant difference; however, in the other attributes, *cluster1* is superior to *cluster0*, even with a group size of 52 students.

Finally, in *COURSE_8*, the K-Means and HDBSCAN algorithms presented similar results in terms of cluster separation and statistical significance. Nevertheless, HDBSCAN outperformed K-Means across all three validation measures analyzed. It was also observed that *cluster1* exhibited higher mean values than *cluster0*, with statistically significant differences across all attributes.

It can be observed that, in all subjects, the clusters with fewer students show higher average results across all evaluated attributes. In each subject, this cluster was classified as "SRL Profile", as it includes students who were more active in the system during the course. These students made more posts in forums and/or activities, viewed more course materials and activities, completed more activity submissions and quizzes, and consequently dedicated more study time during the course period. Additionally, it was noted that in only four subjects, the mean differences of the attributes did not achieve statistical significance in at least one attribute, with the majority being in the "questionnaire" attribute. This suggests that all students generally respond to and interact with quizzes, but the number of views, posts, completed activities, and the time dedicated to the subjects were superior across all datasets. This finding emphasizes that, beyond completing quiz activities, students with self-regulated profiles engage more extensively with other system resources.

In each cluster, the percentage of students with **Grade C** (students who failed the course), **Grade B** (students who passed with an average between 6 and 8), and **Grade A** (students who passed with an average above 8) was analyzed. The panel presented in Figure 8 highlights the analysis of SRL profiles in the clusters identified for each subject. The Mann-Whitney $U$ statistical significance test was applied (URDAN, 2010) to determine whether the grades within the cluster labeled "SRL Profile" are statistically significantly different from the grades in the other cluster identified by the algorithm. For this analysis, it was initially verified that the grades of students assigned to each cluster did not follow a normal distribution, as confirmed by the Kolmogorov-Smirnov test.

The Mann-Whitney $U$ test is a non-parametric test ideal for verifying whether two independent samples have statistically significant differences. It tests two hypotheses: the Null Hypothesis ($H_0$), which states that there is no difference in the distribution of the two populations, and the Alternative Hypothesis ($H_1$), which asserts that there is a difference in the distribution of the two populations. Based on these hypotheses, a comparison of the students' grades in each cluster was conducted. The test was applied to each subject, and it was found that in all subjects, the *p-value* was less than 0.05, leading to the rejection of the null hypothesis. This indicates that there is a statistically significant difference between the grades of the two clusters identified by the algorithm.

In Figure 8, it can be observed that, in all the subjects analyzed, the cluster with the "SRL Profile" identified by the clustering algorithms has a higher percentage of students with **Grade A** and a lower percentage of students with **Grade C**. These results demonstrate that, in the subjects analyzed, groups of students with self-regulated behaviors in the educational environment show better academic performance, as corroborated by the Mann-Whitney $U$ statistical significance test. These findings answer research question RQ2, highlighting a relationship between the self-regulated profile identified by clustering algorithms and the academic performance of students.

## 4.3 Development Stage

Based on the theoretical findings and the results of exploratory analyses conducted in different educational contexts through the implementation of PoCs, it was possible to design a PA aimed at fostering SRL in VLEs. These PoCs, carried out sequentially using data from an international dataset and a national public institution, enabled the identification of student engagement patterns and self-regulatory profiles, which formed the foundation for defining the pedagogical and functional requirements of the architecture.

The PA was implemented in a real educational context. Its development and deployment involved the integration of various technological and pedagogical components within the VLE, including the creation of a custom plugin Time Tracker SRL designed to monitor students' time-on-task and provide feedback on time management strategies. This implementation aimed to foster students' autonomy and self-monitoring capabilities, aligning the environment with the principles of SRL.

The implementation and evaluation stage comprised the practical application of the proposed PA within the Introduction to Python online course, offered through the VLE. During this phase, the PA and its integrated tools were made available to students throughout the course modules, enabling them to engage with the resources designed to foster planning, monitoring, and reflection. The evaluation process focused on analyzing how students interacted with these resources and how such interactions supported


Figure 8 – Analysis of SRL Profiles in each cluster of the eight subjects.

self-regulated learning behaviors. Data collection combined system-generated logs and performance indicators from the VLE, along with responses to self-regulation questionnaires administered before and after the course. These data provided a comprehensive view of students' engagement patterns and learning progress throughout the implementation period. The detailed description of the proposed architecture, including its core components is presented in Chapter 5.

## 4.4 Quantitative Stage - Experiment

This stage aimed to empirically evaluate the effectiveness of the proposed PA in fostering SRL within a real educational context. Grounded on the theoretical and empirical insights obtained from the preceding PoCs and the architecture's development stage, the quantitative experiment was designed to analyze students' behavioral, motivational, and performance data during their interaction with the implemented pedagogical and technological resources in the VLE.

The experiment was carried out within the Introduction to Python online course, hosted on the Moodle platform and delivered between May and June 2024, with a total workload of 50 hours. Forty-nine students enrolled in the course, which was organized into four instructional weeks followed by three weeks dedicated to a final project. The course design combined synchronous and asynchronous sessions, allowing flexibility while maintaining regular opportunities for interaction and engagement.

The implementation of the PA in this course involved the active use of integrated tools such as the Time Tracker SRL plugin, Athena (OpenAI Chat), and Completion Progress that supported planning, monitoring, and reflection, aligned with the cyclical phases of SRL. Two adapted SRL questionnaires were administered at the beginning and at the end of the course to assess students' self-regulation strategies and to capture potential changes over time. The course and its associated data collection procedures were approved by the Human Research Ethics Committee[6], ensuring full compliance with ethical research standards and data protection principles.

Data collection during this experiment integrated multiple quantitative sources to provide a comprehensive representation of students' engagement and learning performance. The dataset comprised: (i) activity logs automatically generated by the VLE; (ii) records from the *Time Tracker SRL* plugin, which monitored students' time-on-task across learning activities; (iii) final course grades assigned to each participant; and (iv) the total time spent on the course. These data sources were systematically pre-processed and analyzed in order to identify behavioral patterns, engagement profiles, and correlations with academic achievement.

[6] Approval number: 78890524.5.0000.8158

Descriptive and inferential analyses were conducted to examine variations in engagement levels, time management, and the adoption of SRL strategies. Comparisons between pre- and post-course questionnaire results enabled the identification of potential developments in students' self-regulatory behaviors over the duration of the course. By integrating behavioral data, plugin metrics, and self-report measures, the analysis provided a multidimensional view of how the proposed PA influenced learners' autonomy, engagement, and academic outcomes.

The following subsections detail the instruments and procedures adopted for data collection, the administration of SRL questionnaires, the pre-processing of learning traces, and the clustering techniques employed to identify distinct SRL profiles within the experimental group.

### 4.4.1 Self-Regulated Learning Questionnaire

The application of self-regulation questionnaires is a crucial step in assessing students' self-regulation skills throughout the course. Below, we detail the procedures for administering the questionnaires and the instruments used. The self-regulation questionnaires were administered at two distinct points: at the beginning and at the end of the course. This schedule allowed for tracking the development of students' self-regulation skills over time. The initial administration of the questionnaire will serve as a baseline to measure the students' level of self-regulation before the intervention, while the final administration will allow us to assess the overall impact of the PA.

#### 4.4.1.1 Instruments Used

To assess students' motivation and learning strategies, an adapted version of the Motivated Strategies for Learning Questionnaire (MSLQ), based on Pintrich *et al.* (1991), was used. The complete MSLQ consists of 81 items distributed across 15 constructs and is organized into two main sections: Motivation and Learning Strategies. Further details are provided in Annex A.

An adapted questionnaire based on the works of Pintrich *et al.* (1991) and Maldonado-Mahauad, Pérez-Sanagustín and Beyle (2020) is used to assess students' motivation and learning strategies, as presented in Appendix A. The answers to this questionnaire use a six-point Likert scale: prefer not to answer, strongly disagree, disagree, not sure, agree, and strongly agree. This scale makes it possible to capture the intensity of students' perceptions and attitudes towards the proposed questions, providing detailed and robust data for analysis.

The data collected through the questionnaires will be analyzed using statistical methods to identify changes in students' self-regulation skills. The analysis will include comparisons between the initial and final scores for each construct of the questionnaire. Analysis

Of Variance (ANOVA) techniques will be employed to determine the significance of the differences observed over time.

Additionally, correlational analyses will be conducted to investigate the relationship between self-regulation skills and students' academic performance. The expected results indicate a significant positive correlation, showing that students with more developed self-regulation skills tend to achieve higher academic outcomes.

All stages of the questionnaire application strictly followed the ethical guidelines establishe by the Instituto Federal do Sul de Minas Gerais (IFSULDEMINAS) Ethics Committee. The participants were informed about the research objectives, and their consent was obtained prior to participation. The confidentiality and privacy of participants' data are ensured using anonymization techniques and secure data storage.

The systematic administration of self-regulation questionnaires will allow for a robust assessment of students' self-regulation skills and provide valuable insights into the impact of the proposed PA. Detailed analysis of the collected data will contribute to understanding how different components of the architecture influence the development of self-regulation skills, informing future improvements in course implementation and pedagogical practices.

### 4.4.2 Data collect

Data collection was conducted on Moodle, based on the participation of students enrolled in the Introduction to Python course at a public educational institution. Prior to the data collection phase, the project was submitted to and approved by the institution's Ethics Committee, thereby ensuring compliance with ethical research guidelines.

Moodle activity logs are used to monitor students' interactions with the platform, providing detailed information on the use of course resources. Four files are collected to build the dataset:

- **Grade Report:** which contains students' scores for each activity and the final course grade;
- **Log Report:** which records all events performed by users in the course;
- **Configurable Reports plugin report:** detailing the amount of time each user spent on each course;
- **Time Tracker SRL plugin report:** offering a detailed view of the time students spent on each course activity, such as forum participation.

To ensure the analytical consistency of the study, the data extracted from these sources were subsequently organized, cleaned, and transformed into a structured dataset suitable for quantitative analysis. This stage was essential for consolidating information from heterogeneous reports into a unified format, enabling the integration of behavioral, temporal, and performance indicators. The following subsection details the pre-processing procedures applied to the collected data, including the selection, aggregation, and refinement of attributes that served as the foundation for the subsequent analytical stages.

### 4.4.3 Pre-processing

The pre-processing and construction of the final dataset were conducted using the Pandas library in Python, which provides intuitive data structures and robust functionalities for the systematic analysis and manipulation of large datasets (MCKINNEY, 2010). This step was essential to ensure the consistency, reliability, and interpretability of the data before applying statistical and learning analytics techniques.

The course under analysis enrolled 49 students; however, 13 of them never accessed the platform, thereby producing no event logs. The raw log report generated a table comprising 12,113 records, distributed across 33 distinct event types. To enhance the analytical validity of the dataset, a sequence of pre-processing operations was applied. Initially, the event logs were aggregated by *user_id* to quantify the total number of interactions associated with each student over the course duration. From the 33 identified event types, three were excluded because they did not represent actions effectively performed by students and, therefore, were deemed irrelevant for the behavioral analysis: *_core_event_course_module_created*, *_core_event_user_profile_viewed*, and *_core_event_user_graded*.

Following this refinement, each log entry was examined and mapped to a corresponding SRL strategy, according to Zimmerman's framework (ZIMMERMAN; MARTINEZ-PONS, 1986), as presented in Table 18. This step ensured that platform interaction data could be reorganized into broader SRL categories, enabling the alignment of digital trace data with the theoretical constructs of the study.

As a result of this methodological process, the final dataset was constructed, comprising eight attributes. It is important to note that the *id* and *time* attributes do not correspond to actions directly performed by students within the platform; rather, they serve as identifiers and temporal reference points, respectively, as shown in Table 19. This dataset served as the foundation for the subsequent stages of the analysis, ensuring the methodological consistency between the data collection, pre-processing, and SRL focused evaluation phases.

### 4.4.4 Clustering

In the subsequent analytical stage, three clustering algorithms were applied to the pre-processed data extracted from the VLE in order to identify and characterize SRL profiles:

Table 18 – Association of SRL Strategies (ZIMMERMAN; MARTINEZ-PONS, 1986) and Moodle Logs.

| SRL Strategy | Log | Log Description |
|---|---|---|
| Seeking Social Assistance | *_mod_forum_discussion_subscription_created* | Represents subscription to a forum discussion, suggesting continued interest in the topic and social interactions. |
| | *_mod_forum_discussion_subscription_deleted* | Indicates the cancellation of a forum subscription, which may represent a change in learning strategy. |
| | *_assignsubmission_comments_comment_created* | It represents the creation of a comment in a task submission, interaction or request for feedback. |
| | *_mod_forum_discussion_viewed* | Records that a student viewed a forum discussion, suggesting involvement in academic debate. |
| | *_core_user_list_viewed* | Indicates that the student viewed the user list, possibly seeking contact with classmates or teachers. |
| | *_mod_forum_post_created* | It shows that the student created a new post on the forum, which may indicate that they are actively seeking academic support. |
| Goal-Setting and Planning | *_core_event_course_searched* | Indicates that the student used the search feature to look for a course, suggesting a proactive step in planning their studies. |
| | *_core_event_course_section_viewed* | Records when a student viewed a specific course section, possibly to plan which activities to prioritize. |
| | *_mod_assign_submission_form_viewed* | Records the viewing of the task submission form. |
| | *_mod_assign_assessable_submitted* | Indicates that the student has submitted an assignment for assessment. |
| | *_assignsubmission_onlinetext_submission_created* | Indicates the creation of a task submission. |
| | *_mod_quiz_attempt_submitted* | Demonstrates that the student has completed and submitted a questionnaire. |
| | *_mod_quiz_attempt_started* | Indicates that the student has initiated a quiz attempt. |
| | *_mod_assign_submission_submitted* | Indicates that the student has submitted an assignment, representing the execution of a planned academic task. |
| | *_mod_assign_submission_created* | Represents the creation of a task submission draft, reflecting planning and time management. |
| Self-Evaluation | *_core_event_course_module_completion_updated* | Indicates that the completion status of a course module was updated, suggesting self-monitoring. |
| | *_core_event_badges_list_viewed* | Represents that the student viewed the list of available badges, possibly as a form of evaluating achievements. |
| | *_mod_assign_submission_status_viewed* | Represents the query of the submission status of a task. |
| | *_mod_forum_post_updated* | Demonstrates that the student reviewed and edited a forum post, suggesting critical reflection on their contribution. |
| | *_mod_forum_post_deleted* | Indicates that the student deleted a forum post, possibly after reconsidering its content, an act of self-evaluation. |
| | *_gradereport_user_grade_report_viewed* | It demonstrates that the student has accessed details about their individual grades. |
| | *_gradereport_overview_grade_report_viewed* | Indicates that the student has viewed their overall grade report, a classic indicator of self-assessment. |
| | *_mod_quiz_attempt_reviewed* | Indicates that the student reviewed their quiz attempt—a reflective action for self-evaluation. |
| Keeping Records and Monitoring | *_core_course_viewed* | Represents access to the course, indicating that the student monitors their progress regularly. |
| | *_mod_assign_course_module_viewed* | Shows that a student viewed a task module, reinforcing the monitoring of academic progress. |
| | *_core_tag_added* | Indicates that a tag has been added, possibly to organize study content. |
| Environmental Structuring | *_tool_usertours_tour_started* | Demonstrates that a tour has been started, indicating that the student is exploring and configuring their virtual learning environment. |
| | *_tool_usertours_tour_ended* | Indicates that a tour has been completed, which may represent a successful adaptation to the digital environment. |
| Reviewing Records | *_mod_quiz_attempt_summary_viewed* | Indicates that the student reviewed the summary of a quiz attempt. |
| | *_mod_quiz_attempt_viewed* | Represents the viewing of a quiz attempt. |

Table 19 – Final Dataset Description.

| **Attribute** | **Description** |
|---|---|
| *iduser* | Unique identifier for the student within the system. |
| *seeking_social_assistance* | Set of data from logs indicating the active search for help from teachers and peers within the system. |
| *goal_setting_planning* | Set of data reflecting the student's actions in planning and setting goals within the platform. |
| *self_evaluation* | Aggregated data that reflects the student's self-assessment process, evaluating their own performance and progress in the virtual environment. |
| *keeping_records_monitoring* | Logs showing the monitoring and maintenance of the student's academic records, aimed at optimizing future learning. |
| *environmental_structuring* | Set of logs indicating how the student organizes and structures their educational environment within the platform to improve the learning process. |
| *reviewing_records* | Data indicating the process of reviewing study materials, tests, and activities performed by the student within the system. |
| *time* | Records the total time of the student's access and interaction within the learning environment. |

K-Means, HDBSCAN, and Agglomerative Clustering. The choice of these algorithms was aligned with the methodological framework of this research and grounded in their complementary analytical properties. K-Means was selected for its computational efficiency and ability to generate well-defined, non-overlapping clusters; Agglomerative Clustering was adopted for its capability to uncover hierarchical relationships and similarities among data points; and HDBSCAN was incorporated for its robustness in handling datasets with varying densities and the presence of outliers. This combined approach strengthens the reliability of the clustering process, enabling a more comprehensive examination of student behavioral patterns in the context of SRL, and providing relevant insights into engagement levels and academic performance.

After the execution of the clustering procedures, the K-Means configuration with $k = 2$ emerged as the most appropriate solution, based on the comparative evaluation of three established clustering validation metrics: Silhouette, Davies-Bouldin, and Calinski-Harabasz. The Silhouette coefficient, with a value of approximately 0.61, indicated satisfactory separation between the two clusters, with minimal overlap. The Davies-Bouldin index, around 0.51, reflected a balanced relationship between intra-cluster compactness and inter-cluster separation, suggesting structural consistency. The Calinski-Harabasz index, reaching approximately 96, denoted adequate discrimination between groups, particularly relevant in scenarios involving a reduced number of clusters.

From a methodological perspective, these results confirm that the configuration with $k = 2$ provides a coherent segmentation of the dataset, consistent with the analytical objectives established in this thesis. Such segmentation supports the identification of distinct SRL profiles, enabling subsequent correlation analyses between behavioral patterns, engagement metrics, and academic performance. The detailed presentation and discussion of these analytical results are provided in Chapter 6, where they are interpreted in light of the research questions and the theoretical framework that guided this study.

This clustering-based identification of distinct SRL profiles provided a solid foundation for the subsequent qualitative stage of the research. To complement the quantitative findings and gain deeper insights into the behaviors, perceptions, and challenges experienced by students within each identified profile, focus group interviews were conducted. This qualitative approach enabled the triangulation of results, enriching the interpretation of the clustering outcomes and offering a more comprehensive understanding of how SRL strategies manifest in the VLE.

## 4.5 Qualitative Stage - Experiment

The qualitative stage aimed to complement the quantitative findings by exploring students' perceptions of the proposed PA. To this end, a focus group interview was conducted with participants who completed the course, encouraging open discussion about their experiences with the learning environment and the integrated tools. The interview protocol addressed topics such as usability, usefulness of the resources, and perceived impact on learning organization, monitoring, and reflection. The collected transcripts were analyzed using qualitative content analysis, allowing the identification of recurring themes and insights regarding the effectiveness of the architecture in supporting SRL. This phase provided valuable contextual evidence that deepened the understanding of how students experienced and interpreted the pedagogical support offered by the PA.

## 4.6 Integration of Results

The integration of results was conducted through a triangulation process that brought together quantitative evidence—derived from behavioral, performance, and questionnaire data—with qualitative insights obtained from participants' narratives during the focus group interviews. This complementary approach ensured a comprehensive interpretation of the data, combining objective indicators of engagement and self-regulated learning behaviors with the subjective perspectives of students regarding their experiences with the proposed PA.

The quantitative dimension encompassed data extracted from Moodle logs, the Time Tracker SRL plugin, and pre- and post-course self-regulation questionnaires. These datasets provided measurable indicators of participation, time management, and academic performance. The qualitative analyses, in turn, focused on students' reflections about how the implemented tools supported their goal setting, monitoring, and self-assessment processes. By connecting both strands of evidence, the integration phase revealed how pedagogical and technological elements jointly influenced learners' autonomy and engagement in the VLE.

This triangulated analysis strengthened the validation of the research questions and offered a multidimensional understanding of the impact of the PA. Beyond numerical trends, it captured the nuanced ways in which students perceived and enacted SRL strategies during the course. The convergence of these findings confirmed the potential of AI-supported educational technologies to promote self-regulated learning behaviors in authentic academic contexts.

Finally, the knowledge generated through this integrative process was consolidated into scientific outputs, including publications in conferences and peer-reviewed journals, as well as the systematic writing of this doctoral thesis. Together, these outcomes demonstrate the theoretical, methodological, and empirical contributions of the study, ensuring coherence with the Sequential Explanatory Mixed-Methods Design and providing a foundation for the broader application of the proposed PA in diverse educational settings.

Chapter 5

# Pedagogical Architecture for Self-Regulated Learning

This chapter presents the proposed PA, which was informed by both the theoretical foundations established through the literature review and related works, and the empirical findings derived from the PoCs. These investigations, conducted using the OULAD dataset and log data from Moodle at IFCDM, provided essential insights into student engagement and self-regulation behaviors, which served as the empirical basis for shaping the overall structure and rationale of the PA.

These analyses allowed the identification of distinct SRL profiles and provided insights into the impact of student behavior and engagement in VLE on their academic performance. Additionally, the theoretical concepts explored, such as PA, SRL, LA, and EDM, as well as the strategies investigated in the related works, supported the development of the PA functionalities. This provided a more robust and theoretically grounded approach to personalizing pedagogical interventions and supporting the teaching-learning process.

Grounded in these theoretical and empirical foundations, the development of the proposed PA was systematically structured to integrate technological, analytical, and pedagogical dimensions in a cohesive manner. This stage aimed to translate the identified requirements into a functional and adaptable architecture capable of supporting the cyclical processes of self-regulated learning within the VLE. The following sections detail the methodological procedures adopted for the design and implementation of the architecture, including the selection of tools, definition of functionalities, and organization of components that operationalize its pedagogical principles. The outcomes derived from the implementation and evaluation of this architecture are subsequently presented and discussed in Chapter 6.

## 5.1 Development of Pedagogical Architecture

The results obtained from the PoCs, combined with the literature review and related works, were essential for the development of the PA. The identification of student profiles with different levels of self-regulation enabled a deeper understanding of the needs and challenges faced by each group, supporting the design of personalized and effective pedagogical interventions. Moreover, the observed correlation between student engagement and academic performance provided valuable insights for the implementation of continuous monitoring and adaptive feedback mechanisms within the PA. These mechanisms allow the system to proactively respond to students' difficulties and support the development of SRL strategies.

The findings from the PoCs also revealed which VLE resources students engage with the most, offering valuable insights into the types of activities that encourage greater participation and self-regulation. In this context, time management emerged as a key SRL strategy for academic success, highlighting the need to provide tools that assist students in optimizing their study time.

The literature review highlighted the importance of tools that assist in time management as a fundamental aspect for the development of SRL skills. Studies indicated that students who are able to organize and monitor the time dedicated to their learning activities are more likely to achieve academic success (SHARP; SHARP, 2016) and (HEMMLER; IFENTHALER, 2024). This finding reinforced the need to include a specific tool for this purpose within the PA. In response to this demand, the Time Tracker SRL was developed—a plugin that enables real-time monitoring of the time spent on learning activities. The tool was designed to provide students with a clear view of how they are managing their time, helping them optimize their study strategies and, consequently, improve their academic performance.

The PA is organized into four key aspects: organizational, content, methodological, and technological. It is the interaction and integration among these components that give the PA its coherence and pedagogical effectiveness. Each aspect plays a crucial role in shaping the learning environment, ensuring that instructional strategies, resources, and technologies are aligned to support student development. To provide a comprehensive and visual representation of this structure, Figure 9 presents a Unified Modeling Language (UML) Component Diagram that illustrates the essential elements of the proposed pedagogical architecture and their interconnections. The diagram highlights how each component contributes to the overall system, with the objective of fostering SRL and enhancing the educational experience in VLEs.

## 5.2 Overview of the Proposed Architecture

The UML Component Diagram (Figure 9) of the PA highlights the interconnections between four main modules: Organizational, Content, Methodological, and Technological Aspects. The Organizational Aspects define the course structure and planning, guiding the methodology used, as indicated by their connection with the Methodological Aspects. These, in turn, shape the content design, ensuring that SRL strategies and assessments are effectively integrated. The Content Aspects, focused on the organization and delivery of material, are influenced by the organizational structure and benefit from the available technologies to maximize student interaction. The Technological Aspects provide the necessary tools to implement the methodologies and distribute the content, completing the cycle of interdependence in which each component supports and strengthens the others to ensure an integrated and effective learning environment.

The proposed PA, implemented in VLE, was conceived to harmoniously integrate theory and practice, with emphasis on student autonomy and SRL. It aims to facilitate the acquisition of technical and theoretical knowledge while engaging students in active and reflective learning. Through interactive and adaptive methods, it promotes the development of autonomous learning skills, enabling students to direct their own educational process. This approach holds substantial implications for research and practice, fostering self-regulation skills essential for academic success and enabling personalized learning through individualized resources and support.

A component diagram is a type of UML diagram that represents the physical structure of a system, highlighting how different software components (modules, libraries, interfaces, etc.) interact (GUEDES, 2009). It describes the organization and dependencies among components, facilitating a clear visualization of the architecture. Each component is treated as an encapsulated unit that can provide or consume services through well-defined interfaces. The key modules depicted include:

- **Organizational Aspects:** Roles and responsibilities of teachers and students, course structure and schedule, and the overall learning model based on constructivist principles;

- **Content Aspects:** Modularized material, types of content and resources, and planning resources to facilitate SRL, including video lessons, readings, quizzes, discussion forums, and practical exercises;

- **Methodological Aspects:** Self-regulated teaching strategies, formative assessments, feedback mechanisms, and activities that promote self-assessment and reflection;

- **Technological Aspects:** Integration of tools and plugins within Moodle, such as Learning Analytics, Configurable Reports, Analytics Graphs, Completion Progress,

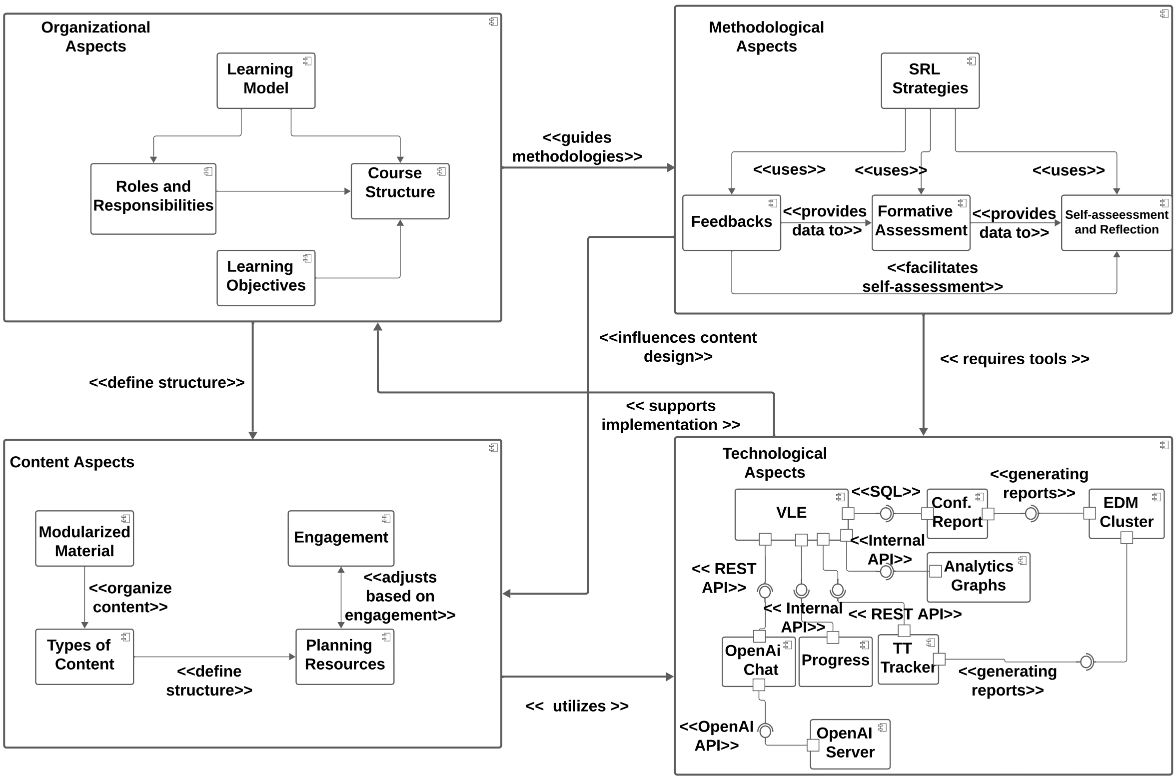



Figure 9 – Pedagogical Architecture.

Athena, and the Time Tracker SRL plugin, supporting the planning, execution, and reflection phases of SRL.

The following subsections describe each of these aspects in detail. The discussion begins with the Organizational Aspects, which define participant roles, course structure, and pedagogical orientation.

## 5.3 Organizational Aspects

The organizational aspects relate to the pedagogical approach designed for a specific context, covering the structuring of time and space, as well as the goals set for both teachers and students. In some cases, this planning may also include other participants, such as tutors or managers.

In this study, the organizational aspects for a 7-week online extension course are planned. The course was offered to students at the Federal Institute of Southern Minas Gerais (IFSULDEMINAS), Carmo de Minas campus, and included participants from the Integrated Technical High School (Computer Science and Food Technology) and the undergraduate Business Administration program. The main objective of the course is to provide students with essential programming skills in Python. Upon completion, students were expected to develop problem-solving programs using the Python language. Each weekly module consists of readings, explanatory videos, webinars, forums, quizzes and practical activities that align with the learning objectives.

The teacher was responsible for preparing and making teaching material available, conducting synchronous sessions, and interacting with students, clarifying doubts and providing feedback. Students were encouraged to actively participate, complete the proposed activities, engage in discussions and seek support when necessary. SRL was a central pillar, with students encouraged to set their own goals, monitor their progress and reflect on their achievements. A detailed description of the elements that comprise the Organizational Aspects is provided below, offering a comprehensive understanding of how each component contributes to the course's structure and execution.

### 5.3.1 Learning Model

The learning paradigm adopted is Constructivism, which posits that students develop their knowledge through practice and experimentation, supported by theoretical underpinning and concrete examples (JONASSEN, 1991). The course was structured with activities designed to foster interaction and collaboration, including discussion forums for sharing programming solutions, clarifying doubts, and group projects.

Constructivism is based on the idea that learning occurs as students actively construct their own knowledge and understanding rather than passively receiving information. In programming education, this model is particularly effective, as it emphasizes practical experimentation, problem-solving, and critical reflection (JONASSEN, 1991).

According to Bodner (1986), constructivism is an epistemological approach or a stream of thought that emerges from the advancements in contemporary sciences and Philosophy. It is a theory that enables us to understand and interpret the world we inhabit. The model exhibits an inherently subjective nature, offering students the flexibility to shape their learning journey. This process is grounded in the students' prior knowledge, their individual perceptions, and dynamic interactions with teachers, peers, and the environmental context in which they are immersed (SONEGO, 2019).

Furthermore, the organizational design of the course also considered the negotiation of the knowledge domain with the students, aligning curricular specificities with their interests and expectations. This flexible approach ensures that the pedagogical architecture remains open to adjustments during the learning process, thereby fostering active participation and ownership of the educational journey.

### 5.3.2 Learning Objectives

The learning objectives of the PA aim to provide a comprehensive educational experience, with a strong emphasis on SRL. The course was planned to integrate technical content with the development of autonomous learning skills. SRL is crucial, encouraging students to set their own goals, monitor and evaluate their progress, and adjust their strategies as needed. This learning model not only imparts technical knowledge but also cultivates essential skills for continuous and independent learning, preparing students to face future challenges with greater self-confidence and autonomy.

The learning objectives were conceived in an open and flexible manner, allowing them to be redefined throughout the course according to students' progress and group discussions. This flexibility is constitutive of the pedagogical architecture and supports the adaptation of strategies to different thematic focuses. In addition, students' prior knowledge was acknowledged as a fundamental element, serving as the starting point for new learning experiences. By recognizing and building upon what students already knew, the course promoted deeper engagement and more meaningful knowledge construction.

Autonomy in the learning process forms the fundamental basis for the construction of a SRL model. This stage is important as it fosters the development of students' ability to direct their own educational journey. This includes the competence to set personal learning goals, identify necessary resources, devise effective study strategies, and conduct periodic self-assessments of the progress made.

Various strategies to foster student autonomy are implemented in the PA. Students are encouraged to set goals at the beginning of the course, defining personalized learning objectives. The architecture assists in managing the students' time, promoting the development of skills through the creation of study schedules and the establishment of realistic deadlines for tasks.

### 5.3.3 Course Structure and Schedule

The Python extension course offered was designed to be completed over a seven-week period, with a total workload of 50 hours. The course was fully online, providing flexibility for students to engage with the materials, resources, and activities at their own pace, within the established deadlines. The course consisted of weekly modules with content and activities and three weeks to develop the final project, named according to the curricular components of the approved extension project at the campus. Each module consisted of up to four 15-minute video lessons, webinars, supplementary materials, quizzes applied directly within the VLE, forums, discussion groups, and programming exercises.

- ❏ **Week 1:** Introduction to Python and its fundamentals, including installation, configuring the development environment and understanding the basic syntax. Learning resources: Webinar, video lessons, reading material, forum, pre-test quizzes, post-test quizzes and self-regulation questionnaire;
- ❏ **Week 2:** Control structures, functions, and basic data manipulation in Python. Learning resources: Webinar, video lessons, reading material, forum, quizzes, and coding exercises;

- **Week 3:** Vectors and Matrices in Python. Learning resources: Webinar, video lessons, reading material, coding challenges, and coding exercises;

- **Week 4:** Methods and functions in Python. Deepen students' understanding of how to define and use functions, the difference between functions and methods, parameters, arguments, and return values. Learning resources: Webinar, video lessons, reading material, forum, and coding exercises;

- **Week 5, 6 and 7:** Practical application of the concepts learned through a final project, which will incorporate all the topics studied during the course. Learning resources: Webinar, video lessons, forum, quizzes, and self-regulation questionnaire.

The proposed PA includes a structured calendar to map out deadlines for assignments and assessments, as well as to include reminders and mark other important dates. This feature is crucial in fostering self-regulated learning, as it enables students to plan and manage their time effectively, promoting self-discipline and personal responsibility in meeting deadlines and educational objectives.

### 5.3.4 Roles and Responsibilities

The proposed PA involves two main actors: teachers and students. The teacher is responsible for the stages of teaching practice and for didactic planning. This planning is crucial to determine the course trajectory and to anticipate potential challenges and opportunities. Therefore, the teacher must know their target audience, establish clear objectives, define content, consider the workload, and be familiar with the teaching modality. The course was offered in a distance learning format, using the Moodle platform. The PA aims to assist the teacher in anticipating activities, content, technological resources, and potential unforeseen events, seeking to plan effective actions to overcome challenges and achieve desired outcomes.

The second key actor in our PA is the student, whose primary function is to develop skills that promote SRL. It is imperative that students adapt their time and devise their own strategies to facilitate the development and construction of knowledge. This process requires a proactive and reflective approach from the student, encouraging autonomy in learning management and content assimilation.

In summary, the Organizational Aspects module represents the essential interactions between the key components that support the structure of a course, highlighting how the Learning Model directly influences Roles and Responsibilities as well as the Course Structure and Schedule. The arrows indicate the direction of this influence, showing that the learning model defines the roles of teachers and students, which in turn affect the course's organization. Additionally, the Learning Objectives interact with the Course Structure, demonstrating that the way the course is structured impacts the achievement of educational goals. These relationships are visually represented in the component diagram shown in Figure 9.

## 5.4 Content Aspects

The second aspect of PA is content, which involves the instructional materials and resources designed to support and enrich students' learning experiences (BEHAR *et al.*, 2020). This includes a variety of learning objects, software, and tools. This dimension defines the scope of content to be covered and encompasses all materials intended to facilitate students' construction of subject-matter knowledge.

As previously mentioned, the extension course will be offered to Integrated Technical High School (Computer Science and Food Technology) and the undergraduate Business Administration program. In this context, the content aspects aim to meet the curricular requirements of the extension project approved at the campus. The primary goal is to ensure that the content not only transmits technical knowledge but also fosters SRL skills and the ability to solve practical problems. It is crucial that the content covered sparks interest, curiosity, and motivation in students. For this purpose, interactive content should be used, effectively and meaningfully engaging students in the learning process.

The content developed for the course is aligned with contemporary practices and emerging trends in the field of Python programming, ensuring that students acquire relevant and applicable skills in the current professional context. This approach guarantees an education that is not only theoretically grounded but also deeply rooted in the practical demands and realities of the job market in the information technology sector. In the following, each item of the content aspects of the proposed PA will be detailed.

### 5.4.1 Modularized Material

The course is offered on the platform in a modular format, facilitating sequential and in-depth learning of each topic. Each module is designed to address a specific topic or set of skills related to Python programming, enabling students to progress through the course in a structured and manageable manner. This modular structure gives students greater control over their learning process, allowing them to focus on topics that require additional attention or to advance more quickly through familiar content.

Each module of the course includes a range of learning resources: video lessons, readings, practical exercises, quizzes, and discussion forums. This multimedia approach caters to different learning styles and helps to keep students engaged and motivated. In the architecture, the modules are presented in a clear and accessible manner, with each section clearly demarcated and accompanied by descriptions and learning objectives. This aids in student usability and helps maintain focus on the learning objectives.

Incorporated into each module, assessment and feedback resources enable students to test their knowledge and receive guidance for their continuous development. This strategy facilitates immediate and necessary adjustments in the learning process. The modular structure of the material not only organizes the content efficiently but also promotes SRL, offering students the autonomy and responsibility to manage their own learning process. This approach ensures that each student can learn effectively, adapting to their own pace and individual needs.

### 5.4.2 Types of Content and Resources

The efficiency of a SRL experience in the VLE significantly influenced by the diversity and quality of the contents and resources provided. The PA we propose adopts a wide variety of content types and resources, each playing a distinct role in the educational process. The central goal of this structure is to cultivate and reinforce students' self-regulation skills, providing an environment that not only facilitates knowledge assimilation but also encourages autonomy and the ability to self-manage learning.

The PA has been designed to address the diverse learning needs of students by providing an interactive educational experience supported by a variety of learning resources. These resources were selected not only for their relevance and effectiveness, but also for their ability to engage students with different learning levels and styles. Each resource plays a crucial role in fostering a self-regulated and stimulating learning environment, encouraging students to explore, experiment, and reflect, thereby supporting a deep and lasting understanding of the content. The learning resources adopted in the PA are described below:

- ❑ **Video Lessons:** Video lessons are used as a key resource to introduce and explain theoretical concepts. Designed to be concise, engaging, and informative, these lessons incorporate practical examples that facilitate the application of knowledge. They aim to promote understanding and content assimilation, constituting a central element of the SRL strategy. This resource allows students to revisit the material as needed, supporting reflection and reinforcement of learning. The flexibility of access and the interactive nature of video lessons foster student autonomy and active engagement in the learning process;

- ❑ **Readings and Support Materials:** These materials are essential for providing a more in-depth perspective on the topics addressed in the video lessons. They facilitate the expansion of students' understanding by allowing deeper exploration of the themes discussed. Their purpose is to support a more elaborate comprehension of the concepts, fostering habits of critical reading and autonomous research. Moreover, these materials play a fundamental role in the SRL process by encouraging students to actively engage in their own learning and to integrate information beyond the core content.

- **Quizzes:** Quizzes are implemented as an effective strategy to reinforce learning and provide immediate feedback. This methodology is crucial for SRL, as it offers students an instant assessment of their performance, enabling them to identify and focus on areas needing improvement. Furthermore, these resources also assist educators in monitoring student progress and adapting teaching according to identified needs, contributing to a more personalized and efficient educational process;

- **Forums:** these virtual spaces are fundamental for fostering the exchange of ideas, clarification of doubts, and collaborative knowledge building. Such resources are crucial in forming an active learning community, promoting the exchange of perspectives, joint problem-solving, and enhancement of communication skills. Moreover, they play a vital role in SRL, as they encourage students to take an active stance in their educational process, fostering reflection, critical analysis, and the development of interpersonal competencies essential for autonomous and continuous learning;

- **Practical Exercises:** are designed to allow students to apply the concepts learned during the lessons, being fundamental for the development of practical skills. They facilitate the transition from theoretical knowledge to real-world application. The use of interactive tools within the platform enriches this learning experience, significantly contributing to SRL. These resources encourage students to take an active role in their learning process, autonomously applying the concepts acquired and reflecting on their applicability and effectiveness in practical scenarios.

### 5.4.3 Planning Resources

Planning resources play a crucial role in SRL, providing students with the necessary tools to efficiently organize and manage their learning process. These resources are fundamental in assisting students in setting goals, developing effective study schedules, and autonomously monitoring their progress. This approach not only facilitates more structured and guided learning, but also promotes the development of essential skills for self-management of knowledge. The following will detail the planning resources used:

- **Integrated Calendar and Schedule:** these resources are implemented within the architecture to facilitate the visualization of task deadlines, assessment dates, and significant course events. The tool is accessible and easy to use, enabling students to organize themselves effectively and precisely. Furthermore, they play a key role in SRL, as they assist students in efficiently planning their study time. This ensures that they are adequately prepared for all academic commitments while also promoting a balance between learning activities and other responsibilities;

- **Task List and Reminders:** these tools are indispensable for effective management of daily activities and academic deadlines. Students have the opportunity to create to-do lists specific to each module of the course and set reminders, ensuring that no essential task is overlooked. These functionalities are extremely valuable in maintaining organization and guiding students, contributing to the reduction of anxiety and optimization of study efficiency. Moreover, they are intrinsically aligned with the principles of SRL architecture, as they foster autonomous learning skills;

- **Monitoring Progress:** the proposed PA is designed to allow both students and teachers to monitor advancement throughout the program. This includes checking the completion of tasks, assessments performed, and participation in various activities. This feature is crucial for providing continuous feedback on student performance, aligned with the principles of SRL architecture. It enables students to have a clear view of the goals already achieved and those that still need to be met, assisting them in effectively adjusting their efforts and learning strategies adaptively. Thus, it significantly contributes to the self-management and development of the student's educational autonomy.

### 5.4.4 Engagement and Interactivity Resources

Engagement and Interactivity Resources are central elements for the effectiveness of a learning environment, especially in online contexts. These functionalities are employed with the goal of capturing students' interest, fostering active interaction, motivating, and creating a more engaging and dynamic educational experience. These resources are essential in the PA, as they stimulate active student participation in the learning process, encouraging autonomy, critical thinking, and continuous reflection. Some of the resources used are:

- **Interactive Tools:** the installed plugins play a fundamental role in enhancing student engagement and promoting interactivity. Tools such as Completion Progress provide real-time feedback on students' progress, participation, and completion of activities, allowing them to continuously monitor their performance. Athena (OpenAI Chat) promotes interactivity by enabling students to interact with an intelligent agent capable of answering questions, providing explanations, and offering real-time guidance on various topics. This conversational interface encourages students to actively seek information and clarification, fostering a more interactive and personalized learning experience. Additionally, the Time Tracker SRL plugin helps monitor the time spent on each activity, supporting effective time management and the development of self-regulation skills. By integrating these interactive tools, the PA promotes a more dynamic and engaging learning experience, encouraging active participation and continuous collaboration among students;

- **Forums:** essential for promoting collaboration and social interaction in an VLE. These spaces encourage students to share insights, debate ideas, and provide feedback to one another, fostering a community of learning. In SRL, peer interaction serves as a motivational tool, enabling students to reflect on their own learning through the perspectives of others, while also reinforcing the development of critical thinking and problem-solving skills;

- **Webinars and Live Sessions:** live learning sessions, such as webinars and videoconferences, are essential for facilitating real-time interaction between students and teachers. These sessions enhance the sense of connection and belonging to the educational community, in addition to providing valuable opportunities for clarification of doubts and immediate discussion of relevant topics. This teaching format is particularly beneficial in the context of Self-Regulated Learning, as it encourages active student participation, promotes collaborative knowledge building, and facilitates the adaptation of teaching to the individual needs of students, contributing to a more dynamic and responsive learning environment;

- **Collaborative Challenges and Competitions:** projects and tasks that require teamwork and collaboration among students are crucial elements in the course structure. These challenges foster critical thinking, creativity, and innovation. Competitions and group projects establish a dynamic and collaborative learning environment, encouraging students to engage more deeply with the course content and to apply their skills in a practical and meaningful way. Furthermore, such activities align with the principles of SRL, as they promote social interaction and problem-based learning, essential skills for the development of autonomy and self-efficacy in the educational process.

The Content Aspects module illustrates the interaction between essential components to promote SRL and student engagement. The key components include Modularized Material, which organizes the course into sequential modules; Content Types, which diversify the material into video lessons, readings, and quizzes; Planning Resources, which help students manage their time with tools like calendars and task lists; and Engagement, which encourages active interaction through interactive and collaborative activities.

## 5.5 Methodological Aspects

The Methodological Aspects module demonstrates the integration of pedagogical components aimed at promoting SRL and creating a dynamic educational environment. The key components are: SRL Strategies, Feedback, Formative Assessment, and Self-Evaluation and Reflection. SRL strategies guide the application of feedback and formative assessments, helping students develop self-regulation skills. Feedback provides insights into performance, while Formative Assessment monitors progress in real-time, allowing for adjustments in teaching strategies. Self-Evaluation and Reflection encourage students to review their learning based on feedback and assessment data, fostering continuous adjustments and improving the overall effectiveness of the educational process.

In an online educational environment that emphasizes SRL, the pedagogical methods and approaches adopted are fundamental in empowering students to take control of their own learning process (BEHAR *et al.*, 2020). This Methodological Aspects section addresses the teaching strategies that will be employed in the course, highlighting how each contributes to the development of essential skills for effective and autonomous learning. The chosen teaching methodologies play a crucial role in facilitating SRL. They are designed not only to transmit knowledge but also to spark curiosity, encourage critical reflection, and promote a deeper understanding of the topics covered. These methodological approaches assist students in becoming active learners, capable of evaluating, managing, and directing their own learning.

The focus will be on strategies that encourage active student participation, promote autonomy, stimulate collaboration, and provide continuous feedback (RIBEIRO, 2019). This includes a variety of practical activities, collaborative discussions, reflective tasks, and formative assessments. Throughout this section, we will explore in detail the various methodologies and practices that form the backbone of our course. The goal is to provide a clear overview of how each methodological strategy will contribute to a rich, interactive, and self-regulated learning environment.

### 5.5.1 Self-Regulated Teaching Strategies

The pedagogical strategies that promote self-regulation, such as project-based learning and problem-based learning, are employed. These approaches encourage students to actively engage in defining their own learning objectives, in seeking resources, conducting investigations, and in the practical application of acquired knowledge (SONEGO, 2019). Thus, they not only facilitate the assimilation of content but also foster critical skills in analysis, synthesis, and evaluation, which are fundamental elements for the development of autonomous and effective learning.

As an integral and culminating part of the course, students will be challenged to apply the knowledge they have acquired in a final project, which involves the development of an application in Python. The goal of this project-based learning goes beyond the mere practical use of programming concepts. It is designed to stimulate fundamental skills such as critical thinking, creativity, and collaboration. Furthermore, the project is perfectly aligned with the objectives of SRL, as it encourages students to take an active role in their educational process. They will be motivated to plan, execute, and evaluate their projects with a reflective and self-directed approach.

In this sense, by being challenged to critically investigate and produce both tangible and symbolic artifacts, students are instigated to define clear learning objectives (planning), to select and apply adequate resources to solve problems (organization and information-seeking), to continuously monitor the progress of their productions (self-assessment and monitoring), and to reflect on the outcomes obtained (self-reflection and self-consequences). Moreover, the cooperative character of these dynamics promotes strategies of help-seeking and social regulation of learning, which are fundamental to SRL in virtual environments. In this way, interactivist-problematizing dynamics not only enrich the formative process but also operationalize the development of self-regulatory competencies within the context of the PA.

### 5.5.2 Learning Feedback

There is a prioritization of regular and constructive feedback, which plays a fundamental role in guiding students to identify areas for improvement and in formulating action strategies for continuous development. The purpose is to foster a growth mindset in students, encouraging them to view challenges as valuable opportunities for learning and evolution. This type of feedback is essential in the context of SRL, as it stimulates self-assessment, reflection, and academic resilience (SONEGO, 2019). Moreover, the methodological design emphasizes distributed pedagogical mediation, where not only teachers but also students take on the role of mediators in the learning process. This occurs mainly in the discussion forums, where all students can read their peers' contributions and interact with them. This practice expands the mediation process, fostering a more collaborative and self-regulated learning environment.

The quizzes and tasks available on VLE will be enriched with an immediate feedback system. For each activity, students will receive automated and detailed comments based on a set of carefully pre-defined criteria. These feedbacks are designed to provide clear and specific guidance, helping students understand where they are excelling and in which areas they can improve. This not only accelerates the learning process but also ensures that students receive constant and constructive support in their educational journey.

As a crucial component of the feedback strategy oriented toward SRL, an interactive initiative focused on *seeking help* is introduced. Through these sessions, students have the opportunity to clarify doubts in real time and deepen their understanding of the covered concepts. This functionality is integrated into the VLE via ChatGPT, providing immediate and accurate responses that support timely and effective learning.

During these sessions, students can explore course themes in greater depth by asking specific questions and receiving detailed feedback without leaving the learning platform. The use of ChatGPT ensures that responses are prompt, clear, and adapted to individual learning needs of each student.

### 5.5.3 Formative Assessment

The continuous use of formative assessments is employed to provide immediate feedback on students' learning, in contrast to an exclusive focus on summative assessments at the end of the course. This approach is essential in assisting students to understand their progress in real time and, consequently, adjust their study strategies as needed (MORENO; PINEDA, 2020). Furthermore, formative assessment is a key component in SRL, as it promotes ongoing self-reflection and self-assessment, enabling students to identify areas for improvement and take proactive steps towards their continuous academic development.

In addition to formative assessment, the proposed PA adopts a processual and cooperative evaluation approach. Learning is assessed through the tangible and symbolic artifacts produced during the course, with the evaluation itself generating new learning products. For example, in the final Python project, students are encouraged to plan, implement, and present their applications in iterative cycles, receiving continuous feedback from both the instructor and their peers. The discussion forums also play a key role in this process, as students can review and comment on their colleagues' ideas, fostering reflection, constructive critique, and the reconstruction of solutions. These mechanisms exemplify how the evaluation process within the PA becomes cooperative and iterative, providing opportunities for knowledge co-construction and ensuring that evaluation is not only summative but also constitutive of the learning journey.

### 5.5.4 Self-Assessment and Reflection

The implementation of activities that promote critical reflection and self-questioning, such as learning diaries and review sessions, constitutes a fundamental aspect of the PA. This strategic approach encourages students to engage in a deeper analysis of their own learning processes, fostering the identification and adoption of effective strategies for continuous development (ZHANG; CHENG, 2019). These pedagogical practices support student autonomy and strengthen self-analysis, metacognition, and self-adjustment skills, which are essential for effective and self-directed learning.

## 5.6 Technological Aspects

Technology has been playing a key role in transforming education. In our SRL approach, Technological Aspects are carefully selected and integrated to enrich the teaching and learning process. This section explores how various tools and technological platforms are used to create a dynamic, interactive, and accessible educational environment, which not only facilitates the transmission of knowledge but also promotes autonomy, collaboration, and active engagement among students.

Our focus is on selecting and implementing technologies that not only complement but enhance learning. This includes the use of Moodle, a robust and versatile online learning platform, supported by AI technologies, and advanced tools like ChatGPT, which offers interactive and instantaneous support for student inquiries.

We will detail each technological aspect of the PA proposed, highlighting how they contribute to a rich and multifaceted SRL experience. Our aim is to equip students with the necessary tools and environment to become confident, capable, and adaptable learners, prepared for the challenges of the modern world.

### 5.6.1 Virtual Learning Environment

VLEs play a central role in supporting SRL by providing structured yet flexible spaces for content delivery, interaction, and feedback. These platforms enable the integration of diverse resources and tools that facilitate the planning, execution, and reflection phases essential to the SRL process. Within the proposed pedagogical architecture, the VLE serves not only as a repository of educational content but also as a space that fosters learner autonomy and engagement.

Among the various VLE platforms available, Moodle was selected to implement the proposed architecture due to its robust features and established presence in educational institutions. Students who participated in the intervention were already familiar with Moodle, as it had been previously adopted by the institution. Characterized by being open-source, Moodle provides an environment enriched with diverse educational resources and offers high flexibility, facilitating customization according to the specific needs of each course and learning model (MOODLE, 2024c).

At the core of our PA, Moodle plays an essential role in creating a learning environment that is both structured and flexible, supporting students' autonomy on their educational journey. The organization of course content into weekly modules facilitates intuitive navigation through the educational material, contributing to a consistent and accessible learning experience. Importantly, specific plugins have been integrated and developed with the purpose of enriching the platform, aligning with the three crucial phases of SRL: planning, execution, and reflection. This approach provides students with robust support to effectively manage their own learning process, promoting the continuous development of self-regulation skills.

The course content has been structured into weekly modules, organizing it into clear and distinct thematic sections. Within each module, there is a variety of pedagogical activities and resources selected and aligned with the specific learning objectives. This modular structure not only facilitates students' understanding of the course's scope and structure but also promotes a flexible approach to study, allowing them to progress at their own pace and according to their needs. Such organization is fundamental in optimizing the learning process, encouraging personal planning and self-management, essential elements for an effective and personalized educational experience.

### 5.6.2 Learning Analytics

LA can be implemented within the Moodle platform to derive in-depth analyses of student behavior and progress throughout their learning journey. Thus, the PA we propose aims to utilize these valuable insights to foster a more informed and personalized self-regulation approach. Through the use of LA, it is possible to identify learning patterns, anticipate challenges, and optimize pedagogical strategies, aligning them with the individual needs of each student (HAMDANE; MHOUTI; MASSAR, 2022). This process not only enriches the learning experience but also strengthens the students' ability to autonomously manage their own educational development, in accordance with the fundamental principles of SRL.

We conduct systematic collection and analysis of data from student interactions with the PA throughout the course. This detailed analysis enables the monitoring of student engagement, assessing the frequency and duration of study sessions on the platform. Furthermore, a qualitative analysis of interaction with the provided educational resources is carried out, identifying which materials are most used and eliciting interest or difficulties. We also investigate task submission patterns and examine student performance on assessments. This comprehensive analytical process enables the implementation of precise and targeted pedagogical interventions to optimize the educational process. As a result, it supports the development of a more effective and adaptive learning environment aligned with students' needs.

Through the identification of these patterns and behaviors, teachers can offer personalized support, adjusting content, pace, and teaching methodologies to better meet the needs of each student. Furthermore, students themselves are empowered with knowledge about their study habits and areas of strength and weakness, encouraging self-reflection and the self-regulation of their learning strategies. Therefore, adopting a data-driven approach to understand learning behaviors not only improves the support offered to students but also enriches the educational process, making it more inclusive, adaptive, and effective. The implementation of LA in identifying learning patterns and behaviors represents, thus, a crucial step in promoting a culture of SRL that prepares students for academic and personal success.

Based on the previously discussed concepts about LA and the characteristics of SRL, the significant potential of LA solutions to stimulate and promote SRL becomes evident (SILVA, 2018). LA solutions offer students opportunities to reflect on their learning process and develop metacognitive skills (DURALL; GROS, 2014). Silva (2018) notes that in contexts with a high volume of data, the process of self-observation can become challenging, and the overload of information may lead to superficial or disorganized metacognitive

Table 20 – Detailed description of collected data.

| **Parameter** | **Collected Data** |
|---|---|
| Logins | Weekly, mid-course, and total course logins. |
| Views | Daily number of views, weekly, mid-course; Total course views, number of unique resources accessed and types of the accessed resources. |
| Forums | Weekly, at mid-course and the overall total of number created posts. |
| Formative assessment | Grades of each formative assessment and participation in assessment regardless of the submission of the answers. |

*Source:* Saqr and Tedre (2017)

monitoring, compromising the essential stage of self-reflection in the self-regulation process. The author further emphasizes that LAsolutions provide crucial pedagogical feedback to assist the student, particularly in the processes of self-observation, self-assessment, and self-regulation.

### 5.6.3 Configurable Reports

We have installed the Configurable Reports plugin [1] in the PA proposed. It is a robust tool in Moodle that allows users to create custom reports without SQL knowledge, making it suitable for both administrators and teachers. It supports report export, scheduling of reports with the use of cronjobs, and sending notifications via email. An important feature is that, although the creation of advanced reports may require SQL knowledge, the plugin also provides flexibility by allowing the creation of basic reports without such knowledge. Further details on this functionality are available in the documentation (MOODLE, 2024b).

A wide range of data can be collected and stored in CSV or XLS formats. The reports generated in the architecture through the Configurable Reports plugin provide a spectrum of detailed information, as listed in Table 20, based on the work described by (SAQR; TEDRE, 2017). The availability of these data is crucial for promoting SRL, as it allows both students and teachers to gain precise insights into the learning process. With this information, it is possible not only to identify areas requiring more attention and development but also to recognize patterns of behavior and engagement that could be pivotal in enhancing individual learning strategies. Furthermore, the analysis of these data supports the creation of personalized and growth-oriented feedback, encouraging reflection, self-assessment, and the continuous adaptation of study techniques, which are essential elements for the effective self-regulation of learning.

[1] https://moodle.org/plugins/block_configurable_reports

### 5.6.4 Analytics Graphs

Analytics Graphs is an advanced visualization tool integrated into Moodle via a plugin [2], designed to facilitate the analysis of educational data through the generation of statistical charts (LLERENA; MORáN; ZAMORA-GALINDO, 2021). This tool is essential for monitoring and understanding student interactions with course materials. This plugin provides five graphs to facilitate the identification of students' profiles by teachers. The graphs are (MOODLE, 2024a):

1. **Chart Visualization of Grades:** Presentation of grades in graphic format, facilitating quick understanding of performance patterns and areas requiring additional attention, including students with problems;

2. **Content Accesses Chart:** Monitoring the frequency with which students access the resources available in the virtual environment, offering a clear view of student engagement over time;

3. **Number of Active Users Chart:** How many users are active in a certain time of day;

4. **Assignment Submissions Chart:** Which users have submited assignments on time or late (tasks and quizzes);

5. **Hits distribution Chart:** How each user is accessing the course and its resources in each cours week.

The Analytics Graphs plugin is an essential tool in the VLE, providing deep insights into students' interactions with course resources. This plugin not only reveals the dynamics of the virtual classroom but is also a key component in the pedagogical strategy, enabling educators to proactively monitor student engagement and progress. By accessing detailed analyses, teachers can identify learning patterns, recognize areas of difficulty, and pinpoint effective study behaviors, which is crucial for adapting teaching strategies to meet the individual needs of each student.

These insights empower educators to implement personalized interventions that promote self-regulation. For example, by providing specific feedback and fostering data-driven discussions, teachers encourage students to reflect on their learning strategies and adjust their approaches to enhance academic performance. In this way, the Analytics Graphs plugin directly aligns with the goals of self-regulated learning, serving as a vital resource for developing self-directed skills in students.

The Grade Distribution Chart (see Figure 10) provides a visual summary of students' performance across the different course assessments. Each boxplot represents the dispersion of grades within an activity, displaying the median, quartiles, and potential outliers.

[2] https://moodle.org/plugins/block_analytics_graphs

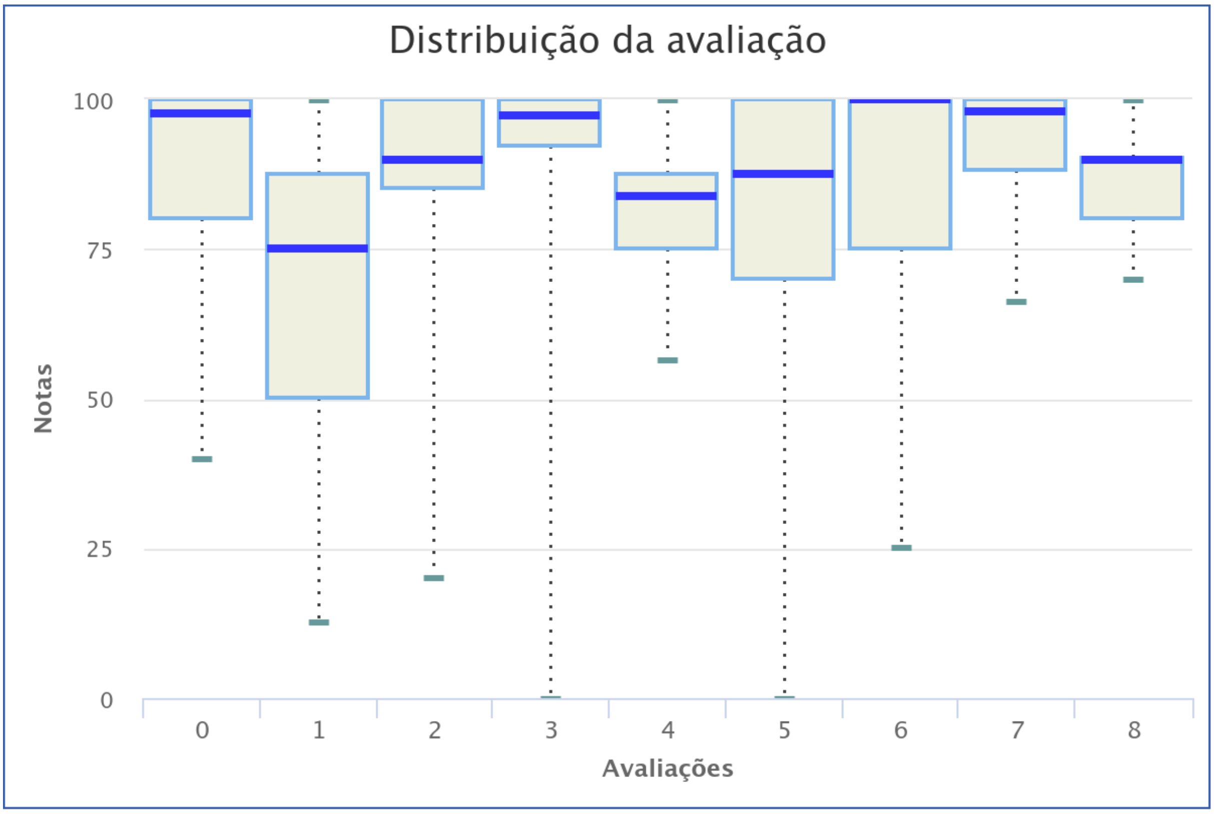



Figure 10 – Visualization of grade distribution across assessments in the course (In Portuguese).

This visualization allows instructors to promptly identify performance variability among learners, revealing both consistency in learning outcomes and areas where students encountered difficulties.

The chart indicates that, in general, student performance remained above the satisfactory threshold in most evaluations, with some variation observed in specific assessments—particularly those that required higher levels of abstraction or problem-solving. Such information is valuable for diagnosing pedagogical effectiveness and guiding instructional adjustments. From the perspective of SRL, this visualization supports reflective teaching practices by enabling educators to identify tasks that challenge students' metacognitive regulation and adapt feedback strategies accordingly.

The Content Access Distribution Chart illustrates the extent to which students interacted with the different resources and learning materials available in the VLE (see Figure 11). Each bar represents the number of participants who accessed or did not access a specific item, grouped by topic within the course structure. This visualization provides a clear overview of student engagement patterns, allowing instructors to identify which contents were most frequently accessed and which received limited attention.

The data reveal that access levels were generally high in the introductory and coding exercise sections, indicating active participation in fundamental learning components. However, a gradual decline in access frequency is observed in some later modules and dis-

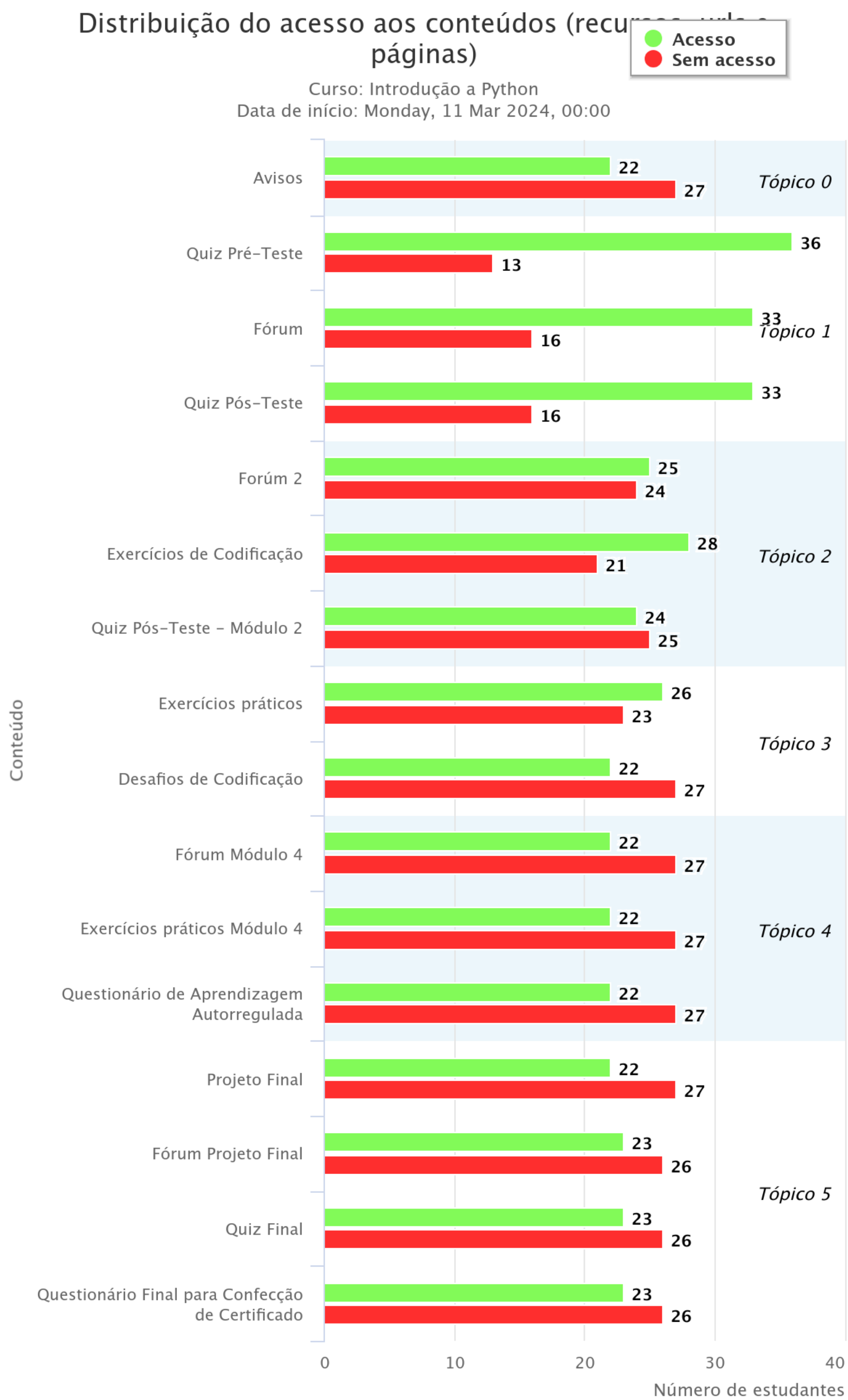



Figure 11 – Distribution of student access to course contents and resources (In Portuguese).

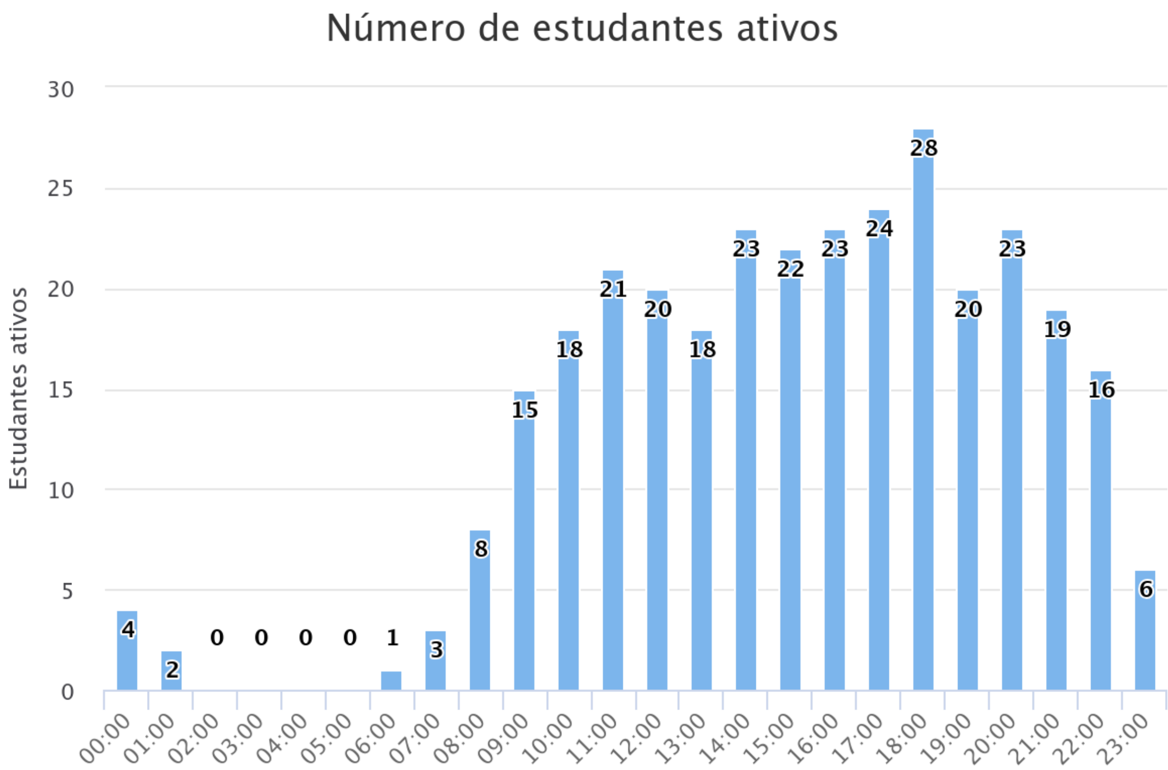



Figure 12 – Number of active students throughout the day (In Portuguese).

cussion forums. This suggests potential challenges in maintaining consistent engagement across the course, especially in collaborative activities that require peer interaction.

From a pedagogical standpoint, these insights are valuable for diagnosing motivational and behavioral aspects related to engagement. By identifying underutilized resources, instructors can implement targeted interventions to encourage participation and foster continuous interaction. Furthermore, in the context of SRL, such visualizations enable both educators and students to reflect on study behaviors, time allocation, and the degree of autonomy demonstrated throughout the course.

As shown in Figure 12, the Active Students Chart displays the distribution of learners' activity across different hours of the day, revealing the temporal dynamics of engagement within the VLE. This visualization provides instructors with valuable insights into students' preferred study periods and patterns of interaction, helping to identify when the virtual environment experiences higher levels of participation.

The data show a clear increase in activity from mid-morning to late afternoon, peaking between 17:00 and 18:00, with a gradual decline during the evening hours. These findings suggest that most students concentrated their learning activities during daytime hours, possibly aligning their study sessions with institutional schedules or personal availability. Minimal engagement was recorded during the early morning period, which aligns with common patterns in distance education contexts.

From an educational analytics perspective, this type of temporal mapping allows in-

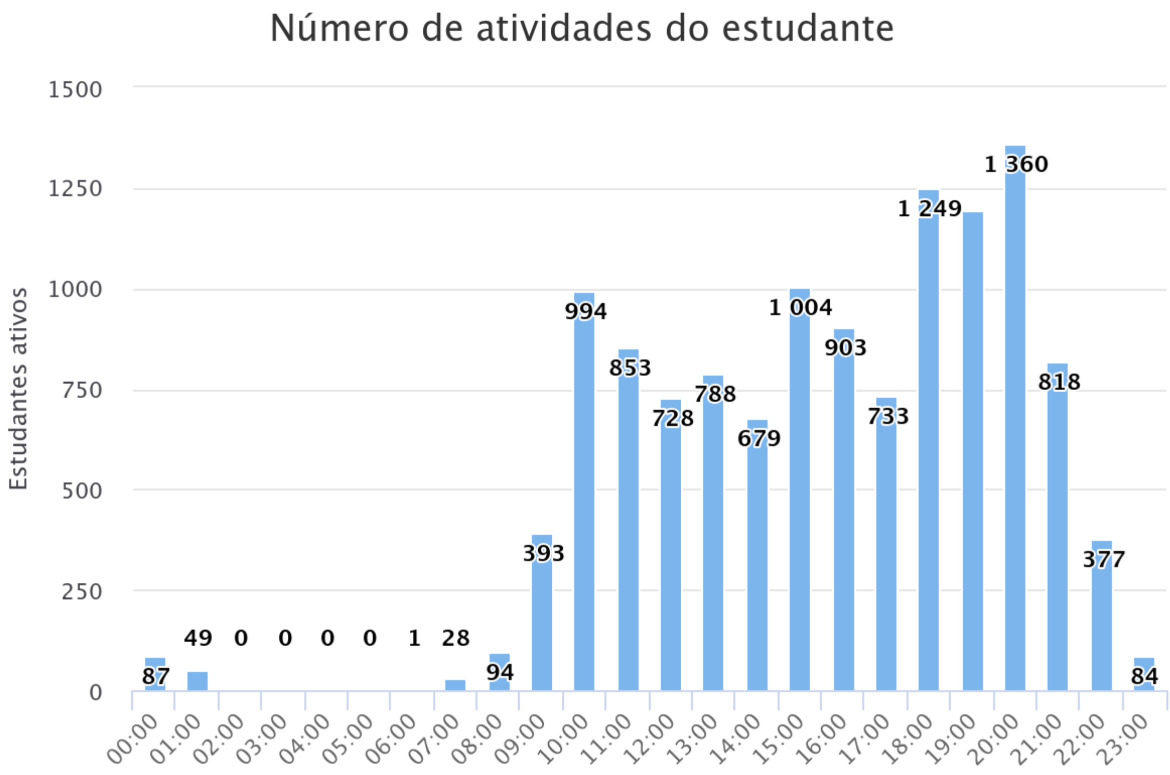



Figure 13 – Number of student activities recorded throughout the day (In Portuguese).

structors to optimize pedagogical interventions, such as scheduling synchronous sessions, publishing announcements, or providing feedback during peak engagement times. In the context of SRL, the visualization also offers evidence of students' time management and study routines, reflecting how learners distribute their cognitive effort and regulate their engagement throughout the day.

As depicted in Figure 13, the Student Activity Chart represents the volume of actions performed by learners within the VLE across different hours of the day. Unlike the chart of active students (Figure 12), which indicates presence, this visualization quantifies the total number of interactions, including resource views, assignment submissions, and quiz participation. This distinction offers a clearer interpretation of engagement patterns.

The results reveal that the intensity of student activity follows a pattern similar to that of active participation, with a gradual increase beginning in the morning and a pronounced peak between 18:00 and 20:00. This indicates that learners not only logged into the environment during these hours but also engaged more deeply in learning tasks. The reduction in activity during the late evening and early morning hours reinforces the predominance of daytime study routines.

From an analytical standpoint, this visualization allows instructors to distinguish between mere access and actual engagement, helping to identify when students are most productive. In relation to SRL, the concentration of activities within specific periods reflects students' time management strategies and self-organization, highlighting their

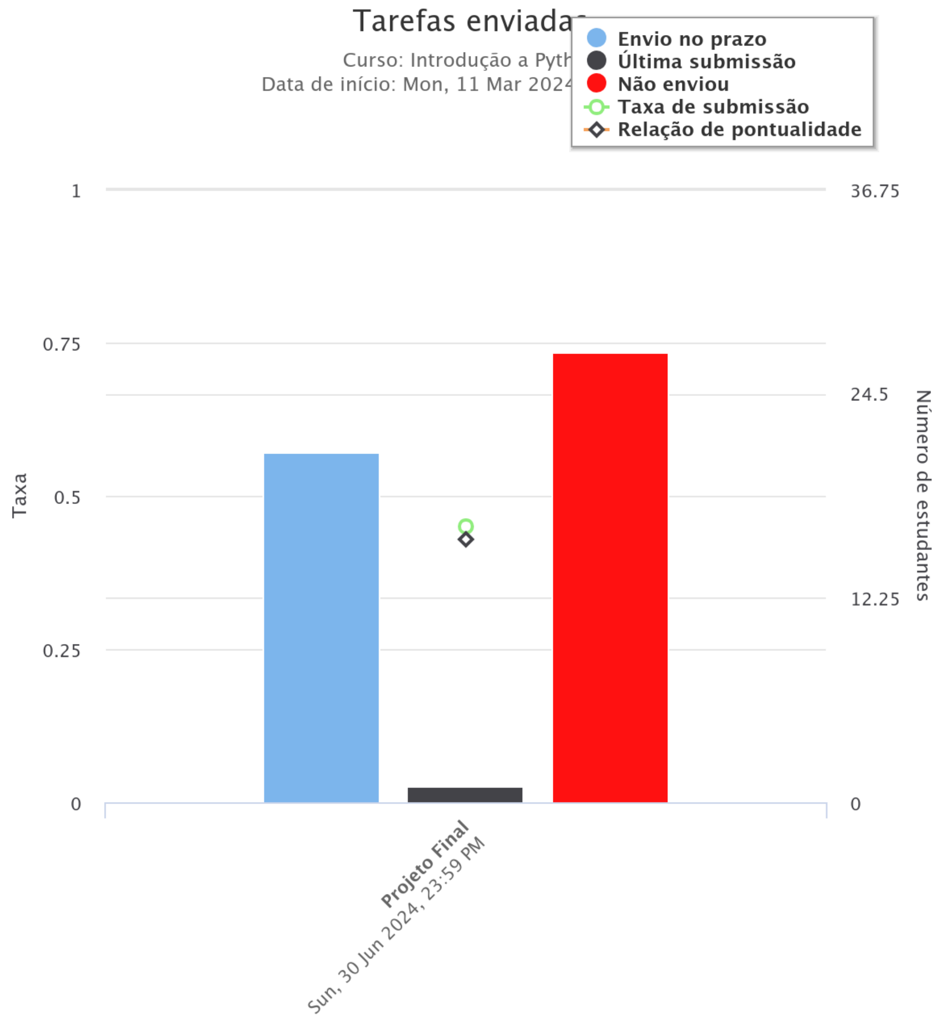



Figure 14 – Submission status and punctuality of student assignments (In Portuguese).

ability to plan study schedules according to personal availability and cognitive rhythm.

As illustrated in Figure 14, the Assignment Submissions Chart presents an overview of students' task submission behavior in the course, specifically regarding the Final Project. The chart differentiates submissions made on time, late submissions, and cases in which the task was not submitted, offering a clear view of punctuality and engagement with evaluative components.

The results show that slightly more than half of the students submitted their work within the deadline, while a significant portion failed to complete the final project. This distribution indicates heterogeneous commitment levels among participants, which may reflect differences in time management, motivation, or perceived task difficulty. The punctuality rate, represented in the chart, serves as an important indicator of learners' organizational and self-regulatory capacity within the VLE.

From a pedagogical perspective, these findings emphasize the importance of scaffolding strategies that promote consistent engagement and support time management skills. In terms of SRL, punctual task submission is strongly associated with the forethought and

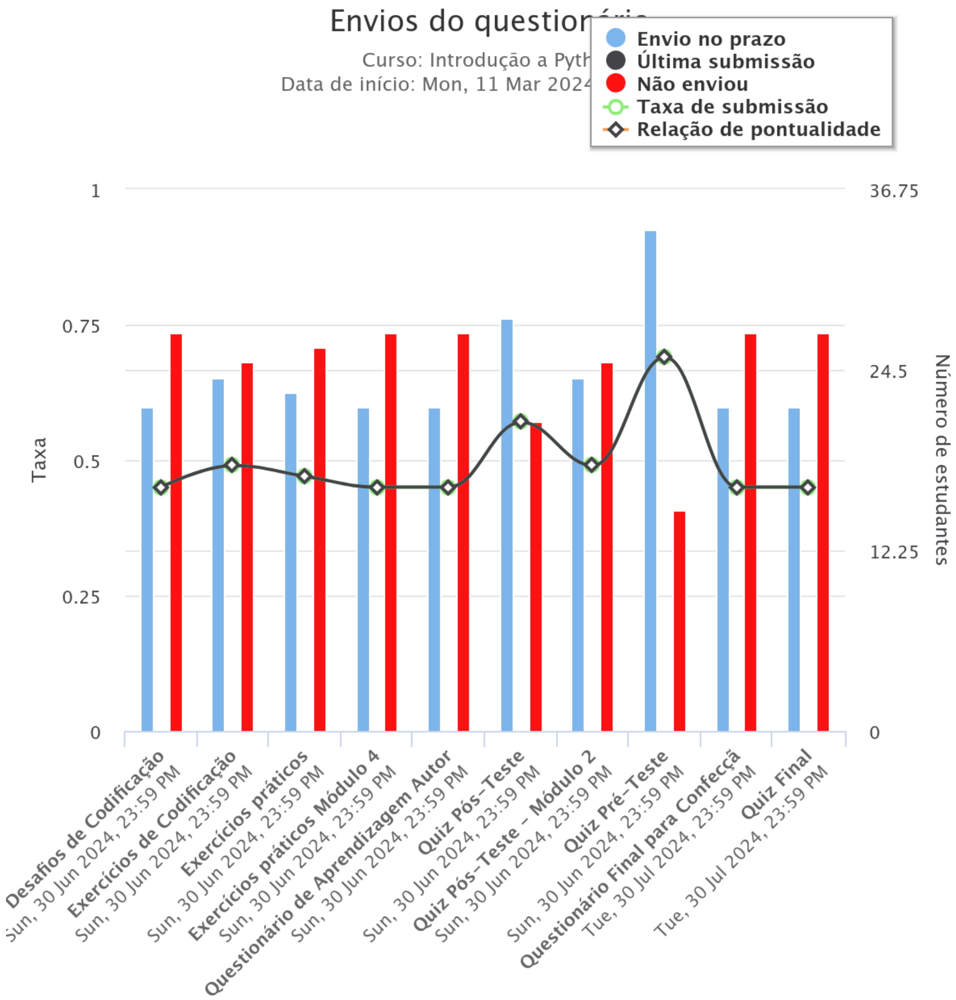



Figure 15 – Submission rate and punctuality across course questionnaires and assessments (In Portuguese).

performance phases described by Zimmerman and Martinez-Pons (1986), as it involves goal setting, planning, and monitoring of progress. Therefore, this visualization not only provides instructors with actionable data for feedback but also highlights behavioral evidence of students' self-regulation practices.

As shown in Figure 15, the Questionnaire Submission Chart summarizes students' submission patterns across the different evaluative components of the course, including self-regulation questionnaires, quizzes, and practical exercises. The chart distinguishes on-time submissions, late submissions, and non-submissions, while also displaying the overall submission rate and punctuality ratio. This visualization enables a comparative analysis of students' commitment to multiple formative and summative tasks.

The data indicate a moderate submission rate overall, with noticeable variation across the different course activities. Higher punctuality levels were observed in quizzes and structured exercises, suggesting stronger engagement with mandatory and graded activities. These findings are pedagogically significant, as they help identify areas where students may require additional motivational or organizational support. In the context of SRL, the submission behavior illustrated in Figure 15 reflects key dimensions of the forethought and performance phases, particularly in relation to goal setting, time management, and self-monitoring. The visualization thus provides valuable insights for both instructors and researchers, linking behavioral indicators to the development of autonomous and self-reflective learning practices.

In addition to the class-level analyses previously discussed, the Analytics Graphs plugin also enables the generation of individualized visualizations for each student, providing a detailed overview of their engagement patterns, activity levels, and academic progress within the VLE. These personalized charts allow instructors to identify students who may require additional support or targeted feedback, facilitating timely pedagogical interventions. By combining aggregate and individual perspectives, this tool enhances the instructor's ability to monitor learning processes, promote adaptive teaching strategies, and foster the development of SRL behaviors among students.

### 5.6.5 Completion Progress

Another relevant plugin that facilitates visualization in Moodle is the Completion Progress[3], which serves both teachers and students. The plugin allows for detailed configuration of the activities that students must complete to meet the course completion requirements (MôNEGO, 2019).

This plugin provides a progress bar that displays the percentage of activities completed by the student, using an intuitive color coding to facilitate interpretation: green indicates completed activities, yellow signals activities in progress, red activities pending or not completed and blue represents activities not started. This visual tool not only provides greater clarity for students about their progress in the course, allowing for effective and autonomous monitoring of their educational progress, but also stands out as an essential time management tool.

The Figure 16 illustrates a screenshot of the Completion Progress plugin for a student in the course. This tool offers a clear and comprehensive view of the course activities, allowing the student to intuitively track and reflect on their overall progress. In the context of SRL, Completion Progress plays a key role by providing continuous visual feedback, enabling the student to monitor their activities, better plan their time, and adjust their study strategies according to the established goals. In the proposed PA, this tool is essential for fostering self-regulation, as it encourages students to be more proactive in managing their activities.

[3] https://moodle.org/plugins/block_completion_progress

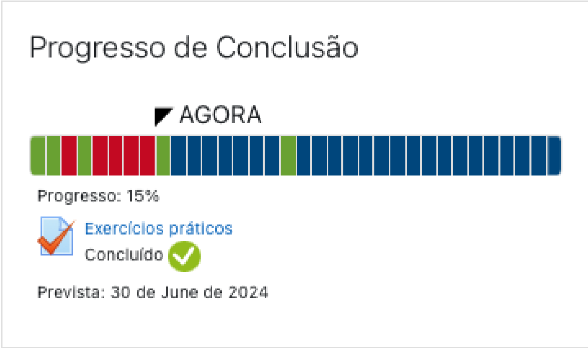



Figure 16 – Screenshot of the Completion Progress for a student in the course (In Portuguese).

In addition to benefiting students, the Completion Progress plugin offers a special feature for teachers through the 'overview of students'. This functionality allows educators to view the progress of all students in the course at once, quickly identifying those who may be at risk due to delays in activities. This panoramic view is crucial for early intervention, ensuring that adequate support is provided to help students stay on track.

Upon accessing the plugin, the teacher can view a screen displaying the student's full name, the last access to the course, the student's visual progress, and the corresponding percentage. Additionally, reports can be generated with this data in various formats, such as .csv and .xls, either for all students or a specific student. The teacher can also send messages directly to a group of students or an individual student and make notes. Figure 17 shows a screenshot of the teacher's view, with the students' names removed to ensure data privacy.

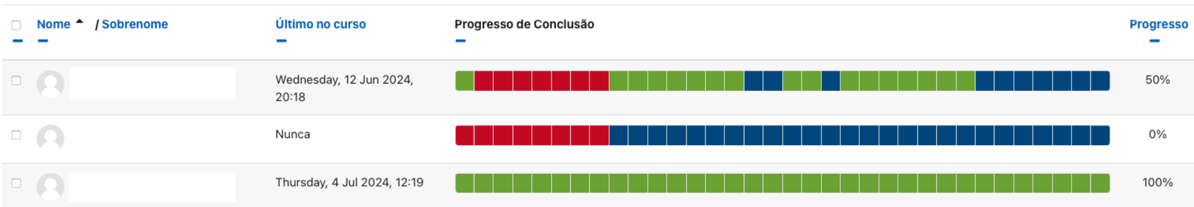
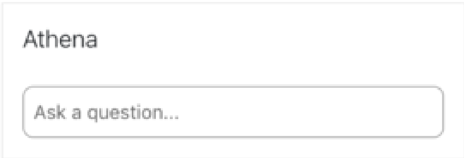
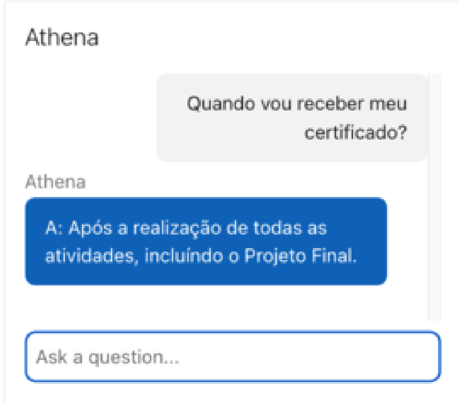



Figure 17 – Screenshot of the Completion Progress for a Teacher (In Portuguese).

The Completion Progress is essential for promoting SRL, as it allows students to visually track the status of their activities through color-coded bars. This visualization aids in identifying whether activities are completed, in progress, or not started, encouraging self-assessment and proactive management of their own learning. This feature strengthens student autonomy, a key component of SRL, by empowering them to recognize where they need to focus their efforts to successfully meet academic objectives. The teacher, by monitoring each student's progress, can intervene in a personalized manner, providing additional support when necessary, making the plugin a crucial tool for SRL in the educational environment.




(a) Home screen




(b) Automated prompt

Figure 18 – Athena screenshot (In Portuguese).

### 5.6.6 Athena

Within the proposed architecture, another installed block plugin was OpenAI Chat[4], renamed as *Athena*, in honor of the Greek goddess of wisdom. This plugin integrates ChatGPT into VLE, providing a graphical chat interface within the course environment, facilitating communication and access to information by students (FIRAT, 2023).

As discussed in Lin (2023), the integration of ChatGPT into VLE offers valuable functionalities, including reflective and diagnostic feedback. This feature helps clarify students' mistakes and suggests improvements, being particularly useful for learners facing specific difficulties. By supporting the identification and correction of underlying issues, this mechanism promotes SRL and encourages students to monitor and adjust their learning strategies.

After installing the plugin from the official Moodle plugin repository, the administrator must add the block to the Moodle interface. Activating the plugin requires an API key from OpenAI, which can be obtained by registering on the OpenAI [5]. Users must then create their secret keys and enter them into the plugin settings in Moodle. This procedure ensures the effective installation and configuration of the plugin, enabling the use of ChatGPT's capabilities in the educational support provided by Moodle.

Figure 18 illustrates two distinct screens of the Athena plugin, integrated into Moodle. Figure 18(a) displays the initial screen of Athena, where students can begin their interactions and seek information directly through a dynamic chat with ChatGPT. Figure 18(b) shows an interface with pre-inserted prompts, demonstrating how students can effectively use the system to seek help or clarification on specific topics, thus facilitating an autonomous and interactive learning experience.

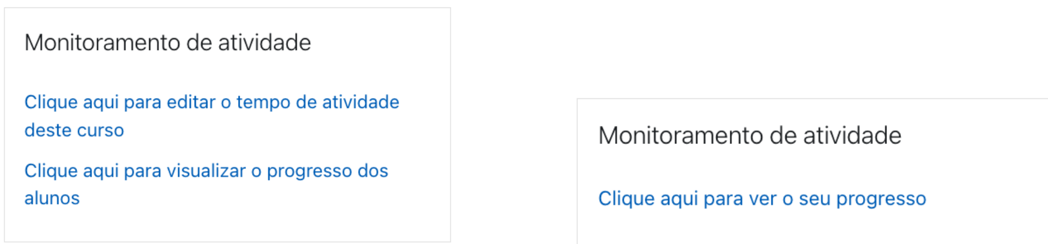
[4] https://moodle.org/plugins/block_openai_chat
[5] https://platform.openai.com/

### 5.6.7 Time Tracker SRL

In a prior study conducted as part of the current research Lima *et al.* (2024), the significance of integrating technological mechanisms within learning environments to foster SRL was emphasized. However, it was noted that the learning environment should not only offer support but also proactively promote self-regulation. This observation underlines the necessity for developing tools like the Time Tracker SRL plugin to enhance the effectiveness of self-regulated learning processes.

Moreover, a significant challenge identified is the scarcity of tools to support SRL in VLEs, as highlighted in a systematic literature review (CERÓN *et al.*, 2021). Therefore, the creation of such tools is essential, as they have been shown to benefit students in terms of SRL, specifically supporting strategies such as goal setting, strategic planning, organization, note-taking, and time management.

The **Time Tracker SRL** plugin, designed to integrate with VLE, is an essential component of the implemented PA. This plugin focuses on facilitating SRL, an approach that encourages students to monitor, regulate, and reflect on their own learning process. The Figure 19 is the screenshot of the Time Tracker SRL. Figure 19(a) illustrates the main screen of the plugin for the teacher, while Figure 19(b) shows the main screen for the student.




(a) Teacher (b) Student

Figure 19 – Screenshot of the Time Tracker SRL (In Portuguese).

Through tracking the time spent on evaluative activities, the tool provides valuable data that allows students to assess the efficiency of their study strategies and adjust them as necessary.

When adding an activity to the platform, the teacher defines an estimated amount of time that students should spend to complete it. This supports effective pedagogical time management, which is essential for the success of SRL processes. In addition, teachers can generate comprehensive reports detailing the time each student spends on individual activities as well as the total time devoted to the course, with outputs available in .csv format. These reports enable educators to monitor and analyze time management patterns across their courses.

The main objective in developing this plugin is to enhance SRL by allowing users to view and manage time invested in various activities such as quizzes, assignments and forums. For teachers, it offers the ability to establish expected duration for each activity. On the other hand, students can monitor the actual time spent, compare it to these predefined expectations, and adjust their study methods accordingly. This dual functionality not only promotes an effective learning environment, but also facilitates the development of critical self-assessment and time management skills, essential for academic success.

- **Time Tracker SRL Modeling**

The modeling of the Time Tracker SRL is a fundamental component in the development of the plugin, aimed at refining the functionalities that support the self-regulated learning process. This section details the methodological approach adopted and the integration strategies with the Moodle platform, aiming to maximize the pedagogical effectiveness and usability of the system. Through this modeling, the goal is to create an environment that not only facilitates time management for users but also promotes a reflective analysis of their own learning.

**Identification of User Needs**

The requirements gathering process began with the clear identification of the main stakeholders' needs: students and teachers in virtual learning environments. The objective was to understand how study time is managed and perceived by these users and how a tool could help them improve time management and self-regulated learning.

The Time Tracker SRL plugin was developed to meet the needs of users, including both teachers and students, by providing an effective tool for time management in assessment activities within the Pedagogical Architecture.

In the "Preparatory" phase of self-regulation, teachers can set clear expectations regarding the time required for each activity, helping students plan their study strategies. During the "Execution" phase, the plugin allows students to monitor in real-time the time spent on each task, aiding them in maintaining focus and adjusting their behavior as needed. In the "Assessment' phase, the Time Tracker SRL provides detailed feedback on time usage, enabling students to reflect on their effectiveness and make adjustments to improve future performance. Thus, the plugin not only facilitates time management but also promotes continuous and reflective analysis of the self-regulated learning process.

**Functional Requirements**

The development of the Time Tracker SRL plugin is driven by the need to enhance SRL through effective time management. This topic presents the functional requirements of the plugin, which were defined to address the specific needs of each stakeholder. Through the integration of these features, the plugin provides comprehensive support for both teachers and students in VLEs.

The key functionalities of the plugin for each stakeholder are delineated as follows:

1. **Time Allocation by Activity:** teachers can set an estimated time for completing each type of activity, facilitating structured and anticipated guidance;

2. **Time Monitoring:** students are able to view the time expended on each activity and the total time devoted to the course, enhancing their ability to manage and allocate their time effectively;

3. **Report Generation:** Teachers should be able to generate detailed reports, either individually or by class, documenting the time each student spends on each assessment activity and the total time dedicated to the course;

4. **Automated Feedback:** the plugin generates feedback based on time spent on tasks, helping students better understand their individual learning needs.

#### Non-Functional Requirements

In addition to the functional requirements, several non-functional requirements have been established to ensure the overall effectiveness and reliability of the Time Tracker SRL plugin. These requirements address critical aspects such as usability, performance, security and scalability, which are essential to providing a robust and seamless user experience. By adhering to these standards, the plugin aims to meet the diverse needs of students and teachers within the proposed PA.

The non-functional requirements are as follows:

1. **Usability:** the interface should be intuitive and easy to use for both students and teachers;

2. **Performance:** the plugin must operate with quick response times, avoiding any noticeable delays in the learning environment;

3. **Security:** ensure the protection of user time data, with access restricted according to the roles defined in Moodle;

4. **Scalability:** capable of operating efficiently in courses with a large number of students.

### ❏ Development of Time Tracker SRL

The development of the Time Tracker SRL plugin involved a comprehensive approach to ensure it met the needs of students and teachers in PA. This section provides an

```
block_course_activity_time/
├── block_course_activity_time.php
├── edit_course_activities.php
├── README.md
├── student_metrics.php
├── students_progress.php
├── styles.css
├── version.php
├── amd/
|   ├── build/
|   └── src/
├── classes/
|   ├── external/
|   ├── local/
|   ├── observer/
|   ├── output/
|   └── tasks/
├── db/
|   ├── events.php
|   ├── install.xml
|   ├── services.php
|   └── tasks.php
├── lang/
|   ├── en/
|   └── pt_br/
└── templates/
    ├── components/
    └── pages/
```

Figure 20 – General Structure of Time Tracker SRL.

overview of the plugin's organizational structure, the main elements incorporated in its design, and the technologies employed in its development. By detailing these aspects, we intend to highlight the systematic and technical foundations that support the functionality and effectiveness of the plugin.

### Organizational Structure of the Time Tracker SRL

After defining the pedagogical, functional, and non-functional requirements, and understanding the Moodle data model, we selected the necessary data and information for the development of our Activity Monitoring plugin. This plugin will be available in the Moodle environment as an additional block, following the best practices recommended in the tutorials available on the Moodle website [6].

The Figure 20 shows the general structure of the Time Tracker SRL. The root file is named **block_course_activity_time**, and the structure includes various files and directories. Some of these files are mandatory and must exist within a Moodle component, while others are optional, as described in the developer tutorials[7]. Some important files are detailed below:

1. ***block_course_activity_time.php***: this is the main file that defines the block class, including essential methods such as *init()*, *get_content()*, and

[6] https://docs.moodle.org/dev/Tutorial
[7] https://moodledev.io/docs/4.4/apis/commonfiles

*applicable_formats()*. It is responsible for initializing the block and defining its behavior in the Moodle environment;

2. ***edit_course_activities.php, student_metrics.php, and students_progress.php:*** these PHP scripts handle specific functionalities of the plugin. They process and display information about course activities, student metrics, and student progress, respectively;

3. ***styles.css:*** CSS stylesheet that defines the visual appearance of the block and its associated pages, ensuring a consistent and pleasant user interface;

4. ***version.php:*** contains plugin metadata, such as the current version, Moodle dependencies, and the component name. This file is essential for version management and the plugin's compatibility with different Moodle versions;

5. ***db/:*** directory containing files related to the plugin's database;

6. ***lang/:*** directory containing language files.

The organization of the **block_course_activity_time** plugin is well-structured and follows best practices for Moodle plugin development. Each file and folder has a specific role, ensuring that the plugin is modular, easy to maintain, and expandable. This structure facilitates understanding and collaboration while ensuring that all functionalities are well integrated and manageable.

### Technologies Used in Development

The development of the Time Tracker SRL plugin utilized a combination of technologies and programming languages to ensure effective integration with Moodle and provide a robust and intuitive user experience. This section details the main technologies employed during the development process:

1. **PHP:** The core language used for developing the plugin, PHP was essential due to its compatibility with Moodle, which is built on this language. PHP scripts handle the main logic, database interactions, and server-side operations of the plugin;

2. **JavaScript:** Used to enhance the interactivity and responsiveness of the plugin's user interface. JavaScript enables dynamic content updates and user interactions without requiring full page reloads, thereby improving the overall user experience;

3. **HTML and CSS:** These standard web technologies were used to structure and style the plugin's user interface. HTML provides the framework for the content, while CSS ensures a visually appealing and consistent design across different pages and devices;

4. **MySQL:** As Moodle's default database management system, MySQL was used to store and retrieve data efficiently. The plugin interacts with MySQL to save user data, track time spent on activities, and generate reports;

5. **XML:** Used to define the database structure and configure plugin events and services;

6. **Moodle API:** Leveraged to integrate seamlessly with the Moodle platform, the Moodle API provides the necessary functions and protocols to interact with Moodle's core features. This ensures that the plugin adheres to Moodle's standards and maintains compatibility with various Moodle versions.

By utilizing these technologies, the development of the Time Tracker SRL plugin ensured not only effective integration with the Moodle platform but also the creation of a user-friendly and functional tool for end users. Each technology played a vital role in building a robust, scalable, and intuitive system focused on improving time management and self-regulated learning in virtual teaching environments.

The Technological Aspects module indicates that the VLE provides data through REST API and internal APIs, accessed by plugins such as Configurable Reports, which generates reports later used by the EDM Clustering component for data analysis. Analytics Graphs is responsible for creating interactive visualizations from Moodle's internal data, while OpenAI Chat communicates with Moodle via REST API and uses the OpenAI API to offer real-time support to students. Completion Progress utilizes internal APIs to monitor student progress, and Time Tracker SRL tracks the time spent on activities, also using REST APIs.

The definition of the technological aspects included the installation of Moodle on a server using Google Cloud infrastructure. To ensure the stability and accessibility of the virtual environment, a custom domain and a fixed IP were configured. This initial setup is crucial to guarantee that the learning environment is robust and reliable, providing a solid foundation for subsequent activities.

Then, Moodle was configured, including the development, installation, and adjustment of essential plugins to support SRL. A new plugin, called Time Tracker SRL, was specifically developed to monitor the time students dedicate to activities, allowing the teacher to set an estimated time that students should spend on each task.

The development process involved the use of PHP, the core language for plugin development. JavaScript was employed to enhance interactivity and responsiveness of the user interface, allowing for dynamic updates without the need for a full page reload. HTML and CSS were used to structure and style the interface. Efficient data management was enabled by MySQL, Moodle's default database management system, while XML was utilized to define the database structure and configure plugin events and services. Additionally, Moodle's API was essential for ensuring seamless integration with the platform's core functionalities, maintaining compatibility with various Moodle versions.

In addition to the developed plugin, other plugins were installed, such as Analytics Graphs[8], which provides detailed visualizations of learning data, Configurable Reports[9], which allows for the creation of custom reports on student performance, and Completion Progress[10], which enables students to graphically track their progress in the course. Lastly, the OpenAI Chat plugin[11] was integrated, incorporating ChatGPT into Moodle. This tool provides a graphical interface for chat interaction within the course environment, facilitating communication and access to information for students.

- **Overview of the Time Tracker SRL**

The Time Tracker SRL plugin offers a range of features that support SRL for both teachers and students. This topic details these features, accompanied by screenshots that illustrate how they are presented on the Moodle platform.

For teachers, the plugin provides a tool that allows them to estimate the time for each activity provided, monitor the actual time spent by students, and generate reports. The main features available for teachers are:

1. **Setting time estimates:** Teachers can set estimated durations for each activity, helping students understand the expected time commitment. Figure 19(a) shows the main screen of the plugin for teachers, where they can choose to edit the time of course activities or view students' progress. Additionally, Figure 21 provides a detailed view of the interface where teachers can enter the estimated time for each activity;

2. **Monitoring Student Time:** Teachers can view the time spent by each student on individual activities, as well as the total time spent on the course. This functionality is shown on the teacher's dashboard in Figure 22. The teacher can filter by student by searching for a specific name or email, or by filtering for a specific date. In the Figure 22, the students' names have been changed to Student A, B, C, and D to maintain privacy.

   The teacher has access to the total number of students enrolled in the course, the total time spent on the planned activities, and the progress on these activities. By clicking on the name of a specific student, a detailed report on each activity is presented to the teacher, as shown in Figure 23, which displays the detailed report for Student B. This same information is made available individually to each student;

[8] https://moodle.org/plugins/block_analytics_graphs
[9] https://moodle.org/plugins/block_configurable_reports
[10] https://moodle.org/plugins/block_completion_progress
[11] https://moodle.org/plugins/block_openai_chat

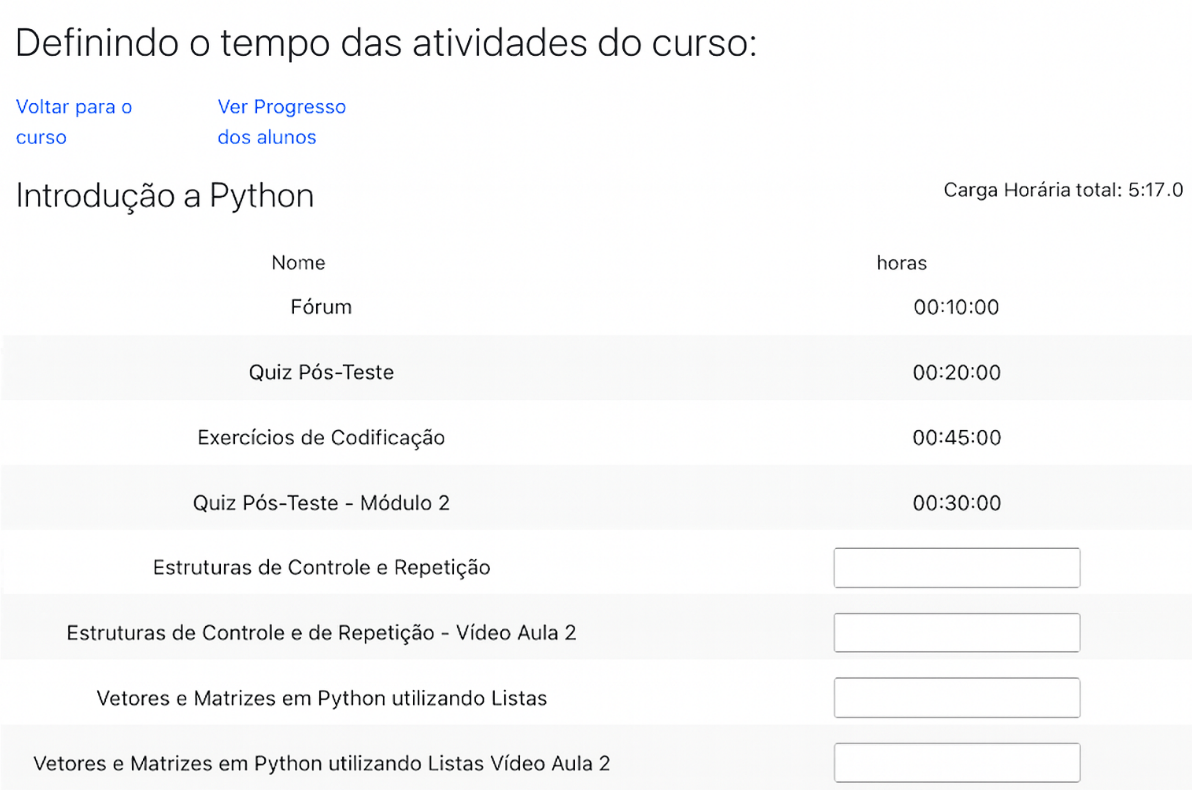



Figure 21 – Screenshot of Time Estimation Insertion (In Portuguese).

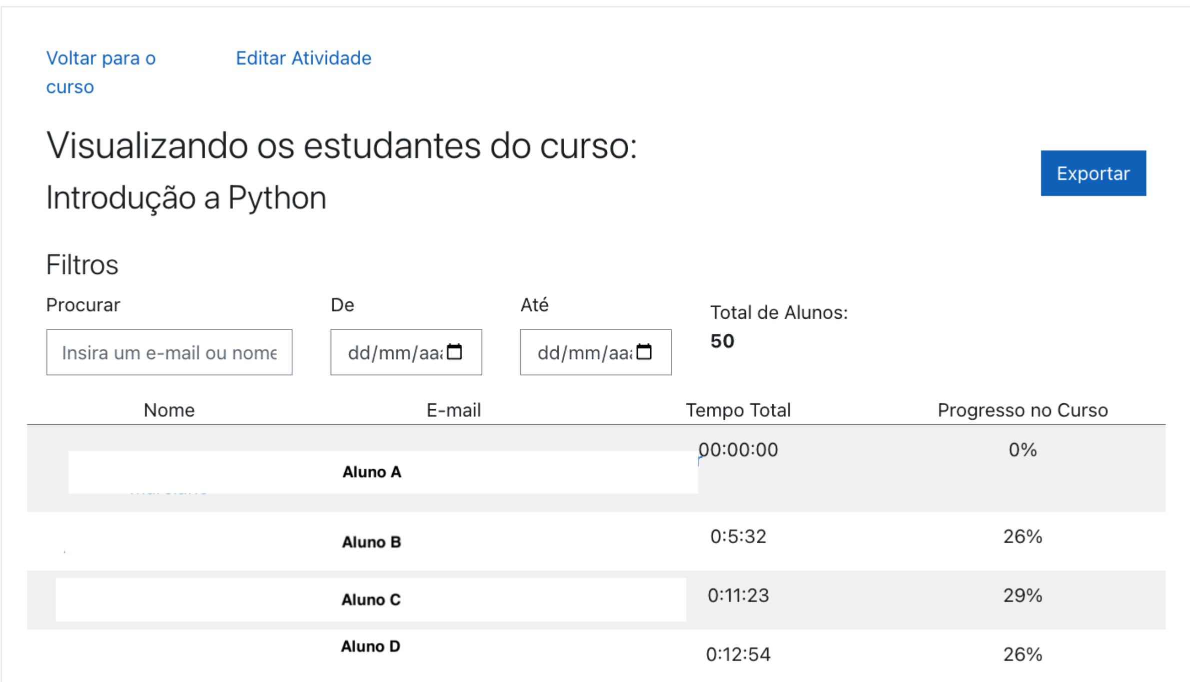



Figure 22 – Screenshot of monitoring time spent by students (In Portuguese).

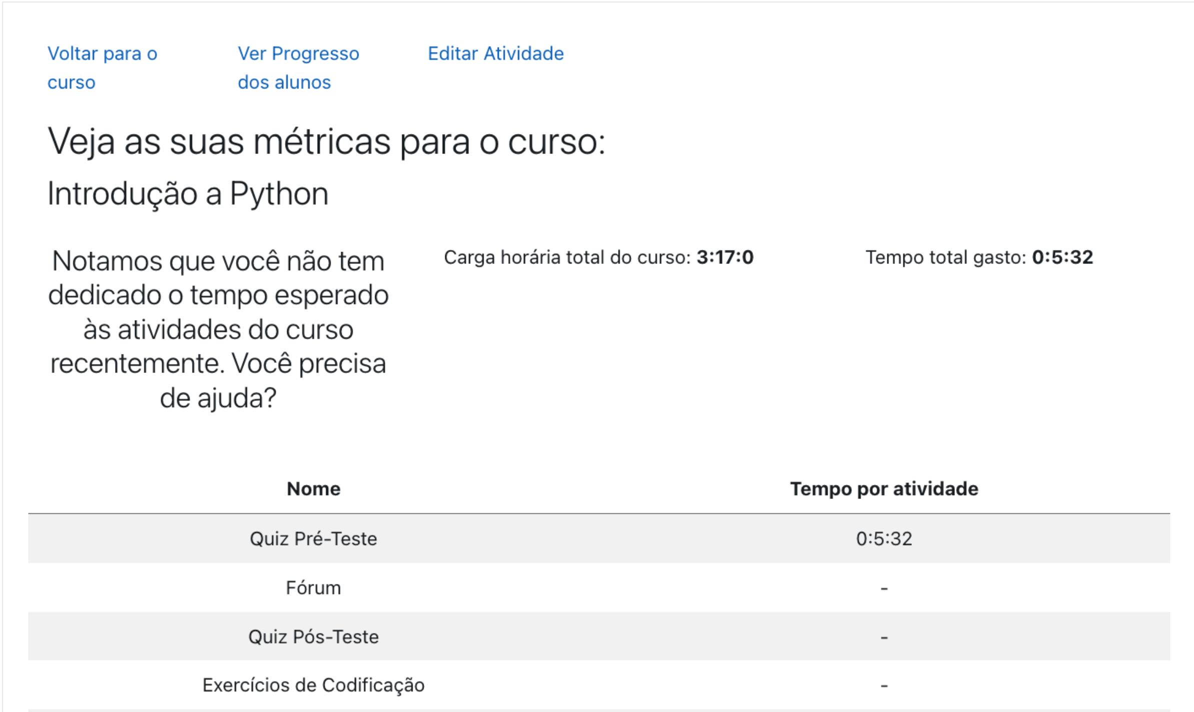



Figure 23 – Screenshot of monitoring time spent by Student B (In Portuguese).

3. **Generating Reports:** Teachers can generate and export reports in .csv format, detailing the time spent by students on different activities. These reports provide insights into students' time management and help identify areas where students might need additional support.

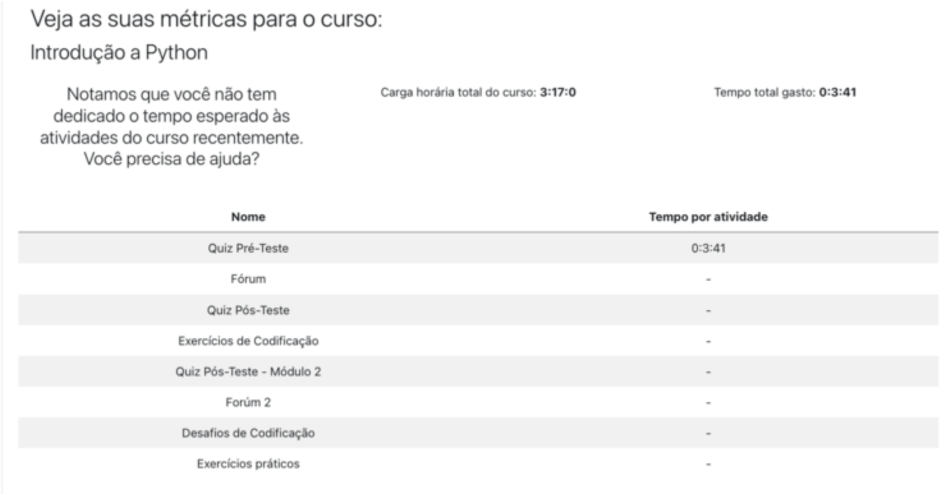
For students, the plugin offers functionalities that allow them to monitor the time spent on activities, compare it with the estimates provided by teachers, and receive real-time feedback. Figure 19(b) illustrates the student's main screen showing the time tracking interface. The main functionalities available to students are:

1. **Real-Time Monitoring:** Students can track the time they spend on each activity and compare it to the estimated time set by the teacher. This helps students manage their time more effectively and adjust their study habits as needed;

2. **Feedback on Time Usage:** The plugin provides feedback based on the time spent on tasks, helping students to reflect on their study strategies and make necessary adjustments to improve their learning efficiency;

3. **Time Management Insights:** Students can view summaries of their time usage across all activities, helping them to better understand their study patterns and identify areas for improvement.

Figure 24 shows three different types of feedback automatically presented to the student. These screenshots are from three different students. Figure 24(a) represents a student whose time spent on activities is significantly below the time planned by the teacher.




(a) Feedback 1


(b) Feedback 2


(c) Feedback 3

Figure 24 – Types of Automated Feedback Provided by the Time Tracker SRL (In Portuguese).

The following feedback is displayed: "We noticed that you have not been dedicating the expected amount of time to the course activities recently. Do you need help"?

Figure 24(b) shows a student who is dedicating the expected amount of time to the course activities, as planned by the teacher. The student sees the following feedback: "You are on schedule with the planned course activities".

When the student spends more time than planned by the teacher, they receive the following feedback: "Your commitment to the course activities is impressive". Figure 24(c) illustrates this scenario, showing the feedback received by the student.

The functionalities of the Time Tracker SRL have been developed to provide a robust and intuitive user experience, promoting SRL through effective time management. By offering detailed insights and real-time monitoring, the plugin empowers both teachers and students to optimize their teaching and learning processes.

Through integration with Moodle, Time Tracker SRL ensures that the learning environment is not only supportive, but also proactive in encouraging self-regulation. Teachers are equipped with tools to set clear expectations and monitor progress, while students receive the feedback they need to reflect on and improve their study strategies.

Overall, the Time Tracker SRL stands out as an important tool in the educational landscape, addressing the need for effective time management and SRL in VLEs. By leveraging technology to enhance pedagogical practices, this plugin significantly contributes to the advancement of personalized and effective education.

## 5.7 Interfaces of the Pedagogical Architecture in the VLE

To complement the description of the proposed PA, this subsection presents the main interfaces of its implementation in the VLE, both from the instructor's and the student's perspectives. The visual organization of the course, the arrangement of resources, and the integration of technological tools are designed to operationalize the pedagogical strategies outlined in the architecture. By aligning content delivery with monitoring and feedback mechanisms, these interfaces demonstrate how the virtual environment was adapted to foster SRL through planning, monitoring, and reflection.

Figure 25 presents the instructor's view of the PA within the VLE, as used in the delivery of the *Introduction to Python* course. The content is organized into modules, with *Module 1: Introduction to Python and Basic Concepts* displayed. This module comprises a variety of educational resources, including tutorial videos, supporting materials, a diagnostic quiz (Pre-Test), and a discussion forum, all designed to foster active learning and student engagement.

On the right side of the interface, features associated with promoting SRL are integrated, such as the (1) Activity Monitoring Panel, (2) Completion Progress, and (3)

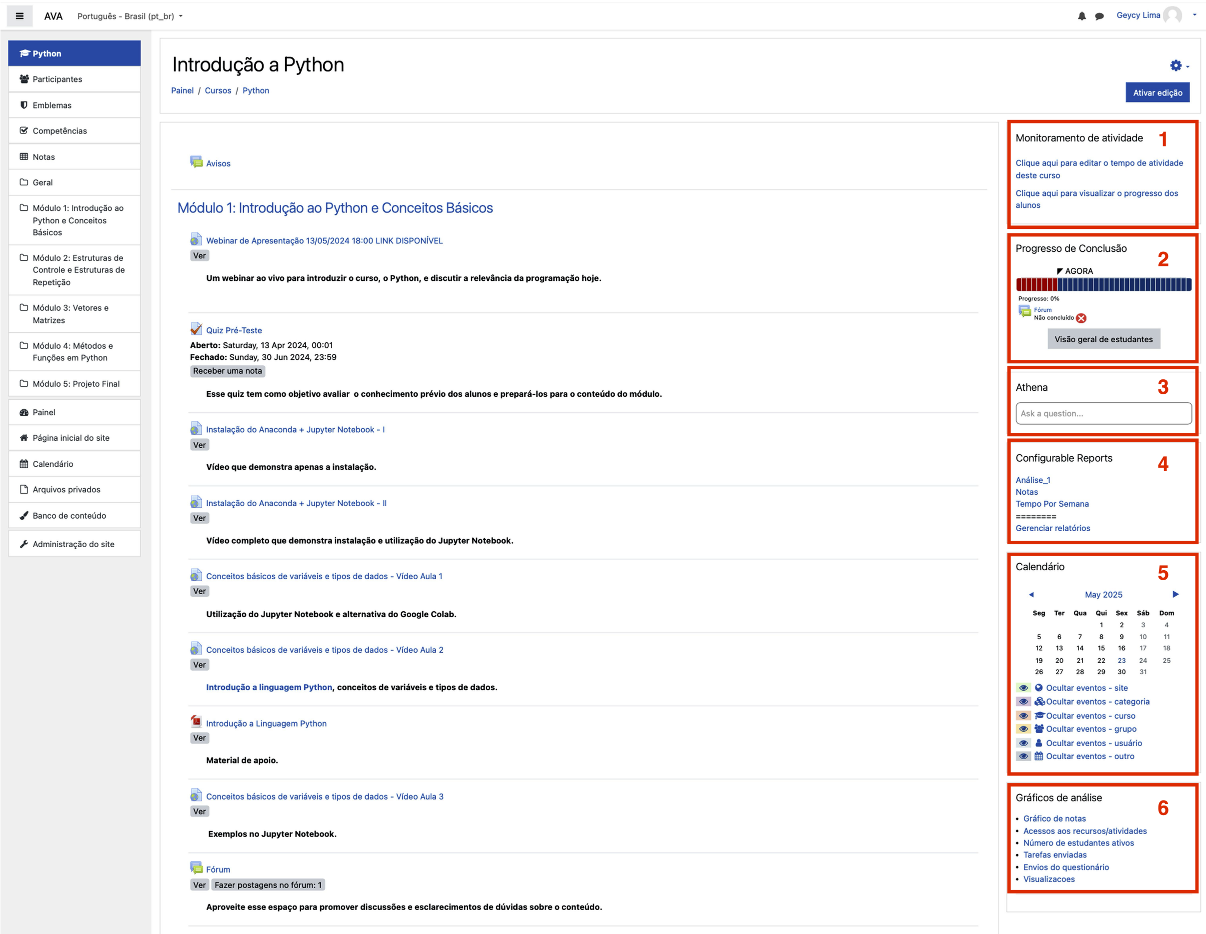



Figure 25 – Screenshot of the VLE (teacher view) showing course organization and SRL support tools (In Portuguese).

Athena—an AI-based interaction plugin. These tools support instructors in tracking overall class performance, monitoring individual student progress, and responding to questions through intelligent assistants, enabling continuous and timely feedback. Additionally, (4) Configurable Reports and (6) Analytics Graphs provide data-driven insights to inform pedagogical interventions, while a (5) visual calendar facilitates instructional planning and time management. This configuration illustrates how the environment combines technological resources with pedagogical strategies to strengthen planning, monitoring, and reflection throughout the learning process.

Figure 26 shows the student's view of the PA in the VLE, also in the context of the *Introduction to Python* course. The screenshot highlights the (1) modular organization of the course and, on the right side of the interface, key features aimed at fostering SRL.

Among the main tools is the (2) Activity Monitoring plugin (Time Tracker SRL), which enables students to track the time spent on each activity, promoting greater awareness of time management. The (3) Completion Progress Panel displays visual indicators of completed activities and a chronological task timeline, supporting continuous monitoring and progress tracking. The (4) Athena panel, based on artificial intelligence, facilitates autonomous information seeking and problem-solving through interaction with intelligent

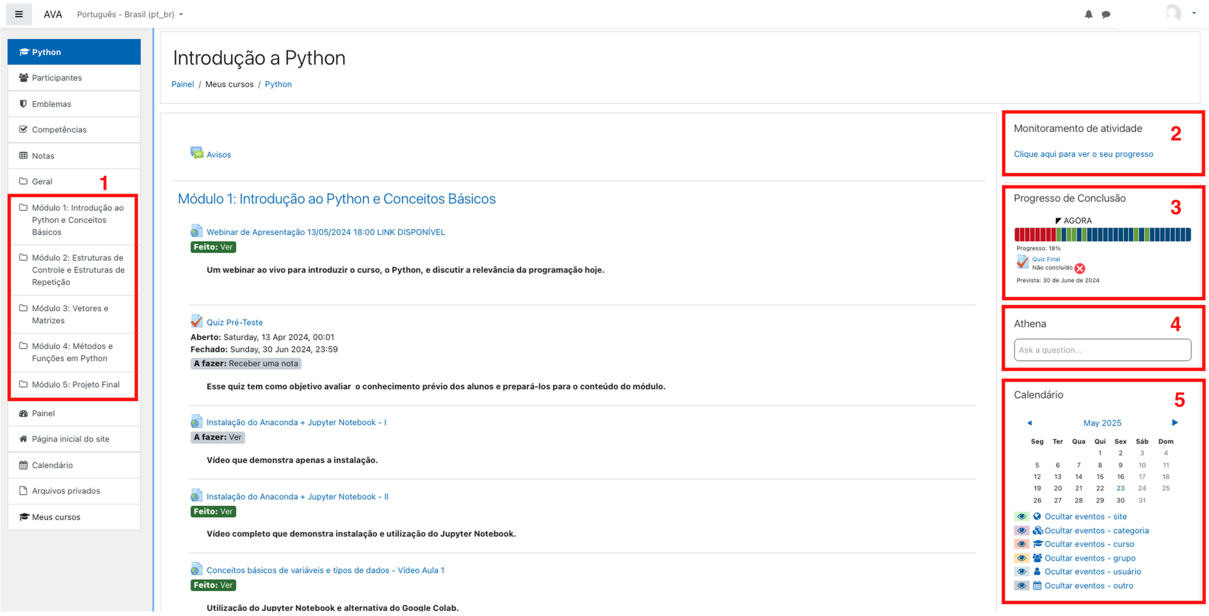


Figure 26 – Screenshot of the VLE (student view) showing course organization and SRL support tools (In Portuguese).

agents, encouraging self-responsibility and metacognitive strategy use. The integrated (5) calendar further aids in planning by clearly indicating deadlines and significant academic events. Overall, this interface configuration reinforces the PA's role in promoting SRL by providing tools that stimulate planning, monitoring, and self-reflection, thereby enhancing students' autonomy and metacognitive awareness within the virtual environment.

In summary, the instructor and student interfaces of the proposed PA illustrate how technological resources, pedagogical strategies, and course organization converge to support the development of SRL within the virtual environment. By integrating tools for planning, monitoring, and reflection directly into the learning interface, the architecture transforms the VLE into an active space for fostering autonomy and metacognitive skills. These configurations not only operationalize the conceptual design of the architecture but also establish the foundation for its empirical validation, the outcomes of which are presented and discussed in the following chapter.

# Chapter 6

# Results

This chapter presents the results and analyses derived from the experimental phase of the research, in which the proposed PA was implemented and empirically evaluated within an authentic VLE. Building upon the design principles and technological components described in Chapter 5, this phase aimed to validate the effectiveness of the PA in fostering students' SRL behaviors, promoting engagement, and supporting academic achievement.

To achieve this goal, the study adopted a Sequential Explanatory Mixed-Methods Design (IVANKOVA; CRESWELL; STICK, 2006), integrating quantitative and qualitative approaches to provide a comprehensive understanding of the pedagogical and technological impact of the proposed architecture. The quantitative phase focused on behavioral and performance data extracted from the VLE including system logs, time-on-task indicators, and responses to pre- and post-course SRL questionnaires—to identify patterns of engagement and self-regulation. The qualitative phase complemented these analyses by examining students' perceptions and experiences through focus group interviews. This approach provided contextual insights and enabled triangulation between behavioral and perceptual evidence.

The combination of these two analytical perspectives enabled a comprehensive evaluation of both the pedagogical and technological dimensions of the proposed architecture. The following sections present the results obtained from the implementation of the proposed approach in an authentic educational context. These sections include analyses derived from system logs, questionnaires, and focus group interviews, highlighting how the integration of quantitative and qualitative evidence contributed to validating the effectiveness of the PA in fostering students' engagement and self-regulated learning behaviors. Together, these results provide a solid empirical basis for discussing the impact and implications of the proposed architecture within VLEs.

This section presents and analyzes the results obtained from the implementation of the proposed PA in VLE. The evaluation aimed to assess the architecture's effectiveness in fostering SRL, improving student engagement, and supporting academic achievement. To achieve this, multiple analyses were conducted, including: (i) examination of VLE log data to identify behavioral patterns and interaction frequency; (ii) evaluation of time-on-task data collected through the Time Tracker SRL plugin; (iii) analysis of completion and progress indicators; (iv) assessment of clustering results to profile students according to SRL-related behaviors; (v) statistical correlation between engagement metrics, self-regulation strategies, and academic performance; (vi) evaluation of questionnaire responses on self-regulation skills; and (vii) thematic analysis of focus group interviews to gather students' perceptions and experiences regarding the use of the architecture. This multi-faceted approach integrates quantitative and qualitative evidence to provide a comprehensive assessment of the proposed architecture's impact.

## 6.1 Clustering

Figure 27 presents the distribution of attributes related to SRL strategies for the two groups identified using the K-Means algorithm. **Cluster 0** consists of 23 students, while **Cluster 1** includes 13 students. Overall, **Cluster 0** exhibits a higher frequency of actions across most of the analyzed strategies, suggesting a more self-regulated learner profile.

With respect to the *seeking social assistance* attribute, **Cluster 0** demonstrated higher levels of engagement in collaborative interactions, including participation in discussion forums and posting of comments, whereas **Cluster 1** exhibited a markedly lower frequency of such actions. The distinction between the two groups becomes even more evident in relation to the *goal setting and planning* attribute. Students in **Cluster 0** recorded a substantially greater number of planning-related events, reflecting higher levels of organization and proactive learning behavior. In contrast, **Cluster 1** displayed limited engagement in these activities, suggesting a less structured and potentially more reactive approach to the learning process. These attributes were derived from the VLE log data and mapped to their corresponding SRL strategies during the data pre-processing stage.

The *self-evaluation* and *progress monitoring* attributes further accentuated the behavioral divergence between the two clusters. Students in **Cluster 0** engaged more frequently in actions such as consulting grades, revisiting previous assessment attempts, and monitoring the completion of activities, evidencing a deliberate and systematic approach to learning regulation. Conversely, **Cluster 1** demonstrated very limited participation in these practices, with distributions concentrated near zero, indicating minimal adoption of metacognitive monitoring behaviors within the VLE.

The *environmental structuring* attribute presented similar distributions for both clusters, indicating that this variable did not play a decisive role in differentiating student profiles. In contrast, the *reviewing records* attribute followed the general trend observed in other behavioral indicators: students in **Cluster 0** demonstrated greater activity, registering a higher number of events related to revisiting attempts and reviewing completed activities, whereas students in **Cluster 1** displayed comparatively lower engagement in

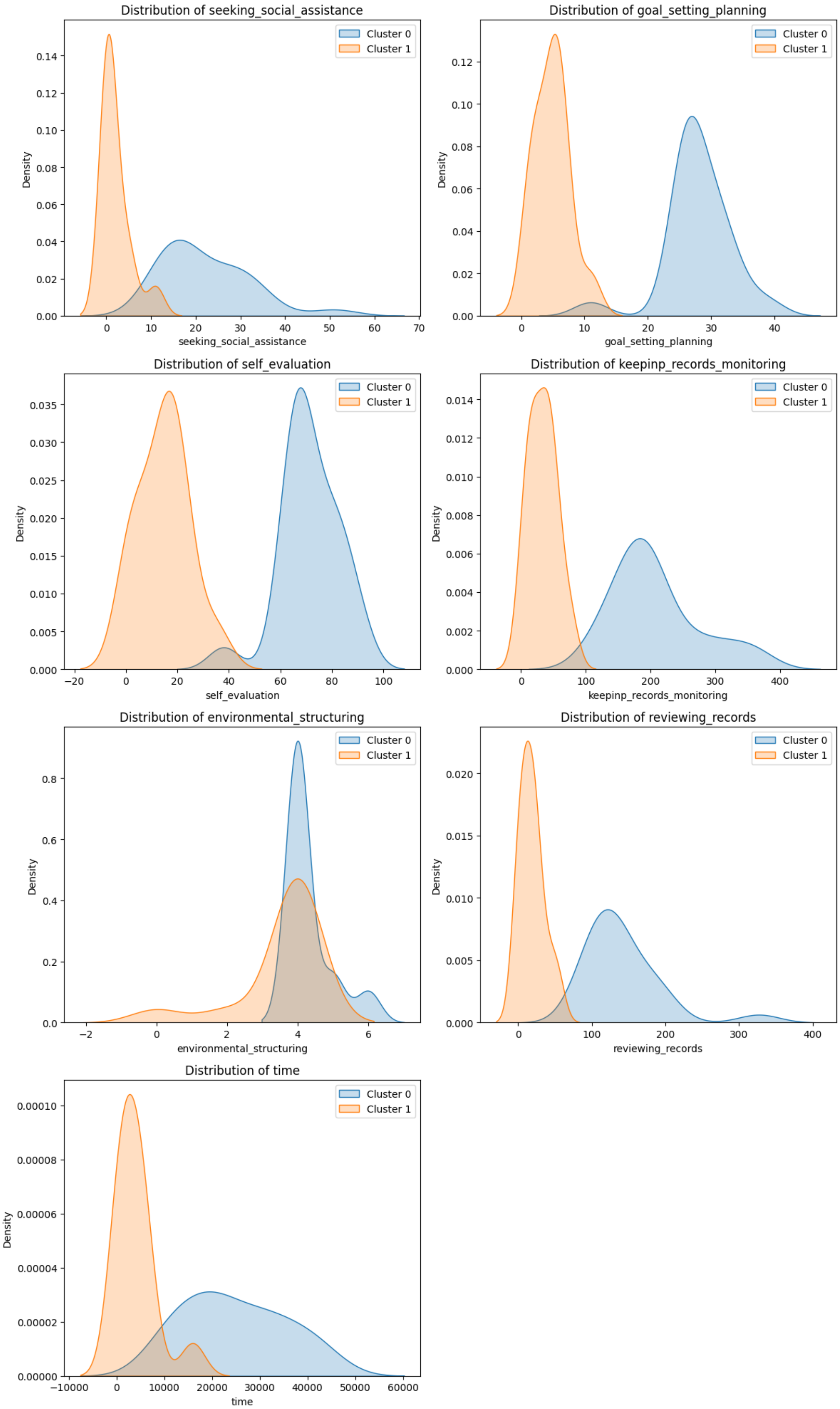



Figure 27 – Density graphs for each attribute.

such reflective practices.

Finally, the *time* attribute, representing the total time spent on the platform, also revealed substantial differences between the two clusters. Students in **Cluster 0** remained active in the virtual environment for longer periods, reinforcing their pattern of higher engagement and consistent application of SRL strategies. In contrast, **Cluster 1** comprised students with markedly less time invested in the system, which aligns with their lower levels of interaction in other behavioral attributes.

In summary, the results indicate that students in **Cluster 0** exhibit behavioral patterns consistent with a self-regulated learner profile, characterized by frequent and diversified engagement in multiple SRL related strategies. The consistent application of these strategies, combined with a higher overall level of activity and time invested in the VLE, reflects proactive learning management and metacognitive awareness. In contrast, students in **Cluster 1** demonstrate lower engagement in these strategies, suggesting a more passive and dependent learning approach. The segmentation obtained through the clustering process reinforces the methodological relevance of incorporating multiple behavioral attributes to identify distinct SRL profiles in online education, providing valuable evidence to guide targeted pedagogical interventions.

Figure 28 presents the distribution of final grades among students grouped by cluster, together with the average time spent in discussion forums. **Cluster 1**, predominantly composed of students with **Grade C** (92.3%) and a small proportion with **Grade B** (7.7%), exhibits a weaker SRL profile. This group also recorded a substantially lower average time in forums, approximately 42.4 seconds—indicating limited engagement in academic and collaborative interactions within the VLE.

In contrast, **Cluster 0** demonstrates characteristics consistent with a stronger SRL profile. The majority of its members achieved **Grade A** (78.3%), followed by **Grade B**(17.4%) and a small proportion with **Grade C** (4.3%). This cluster also presented a markedly higher average forum participation time — approximately 786.3 seconds — suggesting greater involvement in discursive and collaborative activities, which are known to contribute to the development of self-regulatory behaviors.

To assess whether the differences in grades between clusters were statistically significant, the Mann–Whitney U test was applied (URDAN, 2010), with results indicating statistical significance $valor_p < 0.05$. This non-parametric test was selected based on the Shapiro–Wilk test, which confirmed that the grade distributions in each cluster did not meet the assumption of normality.

These findings suggest a strong association between the adoption of self-regulation strategies, engagement in forums, and academic performance. Students who actively participate in interactive activities—particularly those involving peer collaboration and knowledge exchange—tend to exhibit more consistent SRL behaviors, which are reflected in superior academic outcomes. Therefore, the joint examination of cluster composition

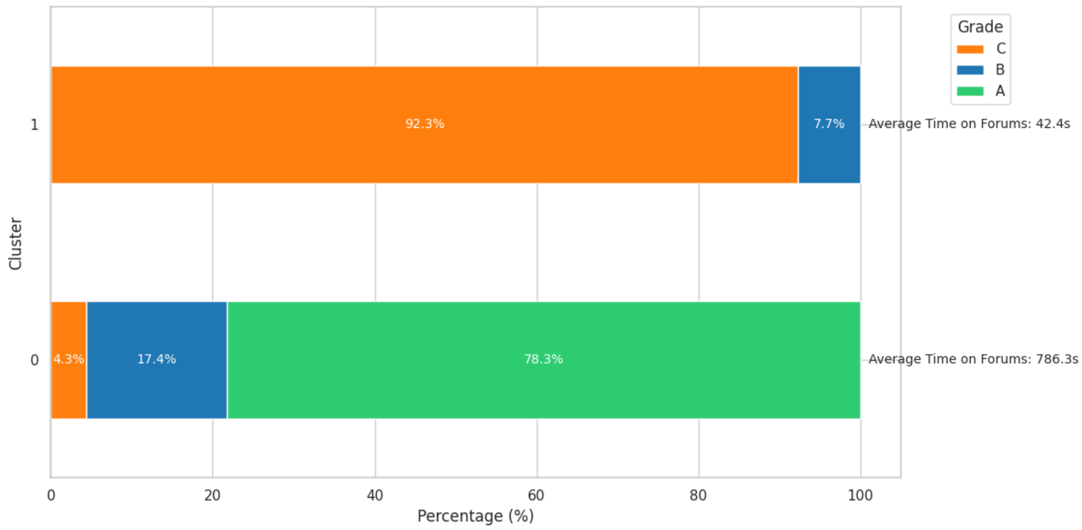



Figure 28 – Distribution of Final Grades and Average Forum Time by Cluster.

and average forum participation time provides a relevant analytical perspective for identifying and promoting pedagogical practices that foster SRL in online education contexts.

The results obtained from the clustering analysis in the PA are consistent with the findings from the PoCs analyses, further validating the effectiveness of the proposed architecture. In both analyses, distinct profiles of student engagement and SRL were identified, emphasizing the importance of time management, participation, and interaction with course resources as key factors in academic success. In addition to these behavioral and performance indicators, the next stage of the analysis examines the results of the self-regulation questionnaire. This complementary perspective allows for the comparison between the SRL profiles identified through clustering and students' self-reported regulation skills, providing further evidence to validate and refine the characterization of learner profiles within the VLE.

## 6.2 Self-Regulated Learning Questionnaire

Self-regulation questionnaires were administered at two distinct points in time at the beginning and at the end of the course with 22 students completing both applications. The instrument was adapted from the validated questionnaires proposed by Pintrich *et al.* (1991) and Maldonado-Mahauad, Pérez-Sanagustín and Beyle (2020), comprising a total of 18 items. The questions addressed core dimensions of SRL, encompassing motivational factors, cognitive and metacognitive learning strategies (such as planning and goal setting), and resource management strategies, including help-seeking, time management, and study environment organization.

To analyze the impact of the course on the development of self-regulation strategies, a comparison of students' average responses before and after the course was conducted according to the SRL dimensions. The questionnaire applied was structured based on

Table 21 – Average scores of SRL dimensions, with differences and *p_values*.

| **Dimension SRL** | **Pre-course** | **Post-course** | **Difference** | *p_value* |
|---|---|---|---|---|
| **Motivation** | 4.03 | 4.25 | +0.23 | 0.017 |
| **Cognitive Strategies** | 3.38 | 3.23 | -0.15 | 0.400 |
| **Metacognitive Strategies** | 3.31 | 3.72 | +0.41 | 0.004 |
| **Resource Management** | 3.52 | 3.58 | +0.21 | 0.018 |

theoretical frameworks of SRL, with its items organized into four dimensions. The Motivation dimension (items 1 to 5) sought to capture students' interest in the content, the value attributed to the activity, perceived self-efficacy, and the desire for achievement. The Cognitive Strategies dimension (items 6 to 8) encompassed the use of basic study techniques, such as rehearsal, organization, and elaboration of information. The Metacognitive Strategies dimension (items 9 to 13) comprised processes of planning, monitoring, and self-evaluation of learning, which are fundamental for the conscious regulation of one's own performance. Finally, the Resource Management dimension (items 14 to 18) included practices related to time and study environment management, as well as seeking support from peers, teachers, and other external resources.

Based on this grouping, mean scores were calculated for each dimension in both the pre-course and post-course stages, allowing the evaluation of students' progress in specific aspects of SRL. Subsequently, a paired t-test was applied to verify whether the observed differences were statistically significant (URDAN, 2010). The results of this analysis are summarized in Table 21.

To assess students' progress in the different dimensions of SRL, a paired t-test was applied between pre-course and post-course scores. The results revealed statistically significant differences ($p\_value < 0.05$) in the dimensions of Motivation , Metacognitive Strategies , and Resource Management , whereas the dimension of Cognitive Strategies did not present statistical significance. The results showed a statistically significant increase in three dimensions of self-regulation: Motivation (+0.23), Metacognitive Strategies (+0.41), and Resource Management (+0.21). In contrast, the Cognitive Strategies dimension ($-0.15$) did not present a significant variation between the pre- and post-test. These findings indicate that the course had a positive influence particularly on aspects related to motivational engagement, monitoring of one's own learning, and the use of resources to support the learning process, although it did not lead to substantial changes in the cognitive strategies employed by the students.

The reliability of the questionnaire was verified using Cronbach's $\alpha$, a metric widely employed to assess internal consistency among items. The instrument presented ($\alpha = 0.85$) in the pre-test and ($\alpha = 0.79$) in the post-test, indicating good overall reliability. Cronbach's $\alpha$ is one of the most commonly used measures for evaluating internal consistency, as it estimates an instrument's reliability based on the correlations among its items, with values above 0.70 generally considered acceptable (CRONBACH, 1951).

Building upon this general analysis, the complete question-level results are presented in Table 22. In accordance with the methodological framework of this study, this complementary analysis focused on identifying which specific items of the questionnaire exhibited the greatest variation between pre- and post-course applications. The results revealed notable improvements in key areas of SRL, particularly in metacognitive strategies, where question 10 ("I follow through with the goals I set for myself in my learning activities") recorded the largest mean increase (+1.045), indicating that students developed greater ability to set clear goals and monitor their learning progress. Likewise, question 11 ("I communicate with my classmates and instructors through forums to reflect on the content I am learning in this course") showed a meaningful gain (+0.909), suggesting an enhanced capacity for critical reflection and self-evaluation.

In contrast, question 6 ("When studying for this course, I read my lecture notes and study materials multiple times") exhibited a mean decrease (−0.318), which may reflect a reduced tendency to rely on repetitive reading strategies. This could indicate a shift toward more autonomous and self-directed learning practices; however, it may also suggest a lower perceived need for peer or instructor support after the course. Table 22 provides a detailed overview of the students' responses, complementing the dimension-level results. In addition, it highlights the potential of the proposed PA to foster specific SRL aspects.

To establish a clearer relationship between students' behavioral data and their self-reported strategies, an integrative comparison was conducted between the clustering results derived from the VLE logs and the dimensions assessed in the self-regulation questionnaire. Table 23 summarizes these findings according to the phases of Zimmerman's model of self-regulated learning *forethought*, *performance control*, and *self-reflection*. This synthesis provides a comprehensive perspective on how the behavioral evidence aligns with students' perceptions and self-assessments, enabling a more holistic understanding of the development of SRL throughout the course.

As shown in Table 23, there is a strong convergence between the behavioral and self-reported data. The *forethought* phase revealed significant improvements in planning and goal setting, evidenced by the higher frequency of planning-related actions in **Cluster 0** and the increase in motivation and metacognitive scores in the questionnaires. The *performance control* phase demonstrated that students in **Cluster 0** exhibited more effective time management and monitoring behaviors, corroborated by significant gains in the Resource Management dimension. Finally, in the *self-reflection* phase, both log data and questionnaire responses indicated greater engagement in self-evaluation and reflective practices. These convergent results reinforce that the proposed PA successfully supported students in completing the self-regulatory cycle—planning, executing, and reflecting on their learning.

Table 22 – Average Scores of Self-Regulation Questionnaire Before and After the Course.

| **Questions** | **Pre-course** | **Post-course** | **Difference** |
|---|---|---|---|
| 1) In a course like this, I prefer study materials that spark my curiosity, even if they are difficult to understand. | 3.86 | 4.04 | 0.181 |
| 2) If possible, I want to achieve better grades in this course than most other students. | 3.40 | 3.63 | 0.227 |
| 3) I believe I will be able to apply what I learned in this course to other courses. | 4.59 | 4.63 | 0.045 |
| 4) I am very interested in the content of this course. | 4.45 | 4.68 | 0.227 |
| 5) I am confident that I can do a good job on the assignments and tests in this course. | 3.81 | 4.27 | 0.454 |
| 6) When studying for this course, I read my lecture notes and study materials multiple times. | 3.27 | 2.95 | -0.318 |
| 7) When I study for this course, I gather information from different sources such as lectures, readings, and discussions. | 4.27 | 4.27 | 0.000 |
| 8) When I study the readings for this course, I create outlines, take notes, or make diagrams and charts to help organize my ideas. | 2.59 | 2.45 | -0.136 |
| 9) I set goals to help manage the time I spend learning. | 3.40 | 3.04 | -0.363 |
| 10) I follow through with the goals I set for myself in my learning activities. | 3.04 | 4.09 | 1.045 |
| 11) I communicate with my classmates and instructors through forums to reflect on the content I am learning in this course. | 2.40 | 3.31 | 0.909 |
| 12) I try to understand how what I've learned impacts my work, study habits, or academic practice. | 3.90 | 4.04 | 0.136 |
| 13) When I face a problem in this course, I can usually come up with several possible solutions. | 3.77 | 4.09 | 0.318 |
| 14) I reach out to someone knowledgeable about the course content when I need help. | 3.45 | 3.77 | 0.318 |
| 15) I share my challenges with online classmates so that we understand what we are working on and how to solve problems together. | 2.72 | 2.86 | 0.136 |
| 16) Even though we don't have daily classes, I still try to spread my study time evenly across the days. | 2.90 | 2.86 | -0.045 |
| 17) I work hard to do everything well in this course, even when I don't enjoy the tasks. | 3.81 | 4.27 | 0.454 |
| 18) I find a comfortable place to study and choose a study schedule that helps me avoid distractions. | 3.95 | 4.13 | 0.181 |

## 6.3 Focus Group Interviews

As part of the qualitative phase, focus group interviews were conducted with 19 students, distributed into five groups according to their educational level, distinguishing between high school and higher education participants. The purpose of this stage was to complement the quantitative findings obtained in the previous analyses, enabling a more comprehensive understanding of students' behaviors and perceptions regarding the use of the proposed PA within the VLE. The sessions took place between August 14 and 15, 2024, at the Instituto Federal do Sul de Minas Gerais – Campus Carmo de Minas.

The interviews followed a semi-structured format, guided by scripts designed to explore students' perspectives on their interaction with the VLE and their experiences with the implemented PA. This methodological choice allowed flexibility for the interviewer to probe emerging themes beyond the predefined questions, thus capturing a wider range of perceptions and reflections (LAZAR; FENG; HOCHHEISER, 2017). The discussion scripts (see Appendix B) focused on key aspects related to engagement, learning autonomy, and the perceived usefulness of the tools integrated into the architecture.

Particular attention was given to resources explicitly designed to promote SRL, such

Table 23 – Integrated analysis of behavioral and self-reported SRL data based on Zimmerman's model.

| SRL Phase | Behavioral Evidence (VLE Log Clusters) | Self-Reported Evidence (SRL Questionnaire) | Integrated Interpretation |
|---|---|---|---|
| **Forethought** | **Cluster 0** showed higher frequency of *goal-setting* and *planning* actions, such as defining objectives and accessing planning tools. | Significant increases were observed in **Motivation** (+0.23; $p = 0.017$) and **Metacognitive Strategies** (+0.41; $p = 0.004$), indicating improved commitment and planning skills. | Convergent evidence suggests stronger proactive regulation and goal orientation among highly engaged students. |
| **Performance Control** | **Cluster 0** demonstrated greater engagement in *monitoring*, *record reviewing*, and *environmental structuring*, reflecting active control of learning processes. | The **Resource Management** dimension increased significantly (+0.21; $p = 0.018$), revealing improvements in time and study environment management. | Findings indicate enhanced self-monitoring and use of regulatory strategies during task execution. |
| **Self-Reflection** | **Cluster 0** exhibited more frequent *self-evaluation* events, revisiting completed tasks and grades. | Higher post-test means in **Metacognitive Strategies** confirm greater emphasis on evaluation and reflective learning. | Behavioral and self-reported data jointly indicate consolidation of reflective and adaptive learning behaviors. |
| **Cognitive Strategies** | Not directly captured in log data; clusters mainly reflect behavioral and metacognitive aspects. | No significant change observed in **Cognitive Strategies** ($-0.15$; $p = 0.400$). | The difference likely stems from the nature of the instruments, reflecting complementarity rather than contradiction between behavioral and self-reported approaches. |

as self-assessment quizzes, collaborative activities, progress tracking, time-on-task monitoring, and the Athena assistant. Students were encouraged to reflect on how these tools supported the core phases of self-regulated learning planning, monitoring, and self-reflection. Across all groups, participants described the learning environment as intuitive and well-structured. This perception was largely attributed to the integration of these tools within the course organization.

The Completion Progress bar was consistently mentioned as a valuable feature for planning and organization. Students reported that the visual tracking of completed and pending activities helped them manage deadlines, allocate study time, and maintain a clear sense of progression throughout the course—aligning with the forethought and performance phases described in Zimmerman's model of SRL (ZIMMERMAN; MARTINEZ-PONS, 1986). Similarly, the Time Tracker SRL plugin was praised for its role in encouraging reflection on time management, as it allowed learners to monitor and adjust their study habits based on real usage data.

Weekly quizzes were also identified as one of the most engaging and beneficial components. Students highlighted that these formative assessments facilitated knowledge consolidation and self-evaluation, enabling them to identify topics requiring further study. This metacognitive engagement contributed to strengthening monitoring and reflection skills, reinforcing a feedback loop essential to the development of SRL.

On the other hand, some challenges were reported regarding the use of discussion forums. Although these activities were designed to foster collaboration and help-seeking, participation remained limited. Several students attributed this to interface complexity or to the rule that required posting before viewing peers' messages, which some perceived as demotivating. These observations suggest the need for interface and instructional design improvements to better support the social dimension of SRL.

Another prominent feature mentioned by participants, particularly those from higher education, was the use of the Athena assistant (OpenAI Chat). Students emphasized its importance in enabling immediate and individualized feedback, especially for questions they might hesitate to ask publicly. This functionality was regarded as an effective mechanism for supporting help-seeking and reflection, reducing cognitive barriers to participation and increasing learners' confidence and autonomy.

Furthermore, the final project-based activity was highlighted as a motivating and integrative element of the course. Students reported that it provided opportunities to apply knowledge, demonstrate practical skills, and reflect on their progress, contributing to the self-reflection phase of SRL. Several participants expressed that completing the project increased their sense of accomplishment and self-efficacy, confirming the link between well-structured tasks and the development of autonomous learning behaviors.

The qualitative evidence obtained through these focus group sessions corroborates the quantitative findings presented earlier, reinforcing the effectiveness of the proposed PA in fostering SRL. Features such as progress visualization, modular content sequencing, formative assessment, and AI-assisted feedback were consistently recognized as key contributors to students' engagement and self-regulatory behavior. Nonetheless, the underutilization of collaborative tools underscores the need for future refinements aimed at enhancing peer interaction and social regulation within virtual learning environments.

In summary, the focus group interviews provided rich insights that complemented the behavioral patterns identified in the quantitative phase. This triangulation of evidence strengthened the methodological rigor of the study, offering a multidimensional understanding of how the proposed PA supported autonomy, engagement, and reflective learning practices.

## 6.4 Integration of Results

Based on the data analysis, the results address **RQ1**, demonstrating how the design of the PA facilitates the development of SRL skills. The integration of tools such as the Time Tracker SRL, Athena, Completion Progress, quizzes, forums, and self-assessment resources played a fundamental role in supporting students, empowering them to plan, monitor, and adjust their learning strategies effectively.

The data indicate that students from **Cluster 0**, who more frequently used these PA resources, exhibited an SRL profile. The Time Tracker SRL, for instance, proved to be a valuable resource for the development of self-regulation by enabling students to monitor the time spent on activities, identifying the need to adjust their study routines. This is particularly important, as one of the main SRL skills is the ability to manage time autonomously and strategically.

Furthermore, the frequent use of quizzes and forums (Figure 28) by the students contributed to an improvement in self-assessment skills and goal-setting, as confirmed by the increase in averages in the self-regulation questionnaires. The results of the attribute density analysis (Figure 27) showed that students more engaged with these resources demonstrated a more developed ability to reflect on their progress and adjust their learning goals. Based on the presented results, **H1** was confirmed. The implementation of the AI-supported PA resulted in a statistically significant improvement in some of the students' self-regulation skills.

**RQ2** investigates how students' engagement with PA resources correlates with their academic performance. The results reveal a significant correlation between the level of student engagement and their academic outcomes, demonstrating that intensive use of PA resources is associated with better performance.

Specifically, the students from **Cluster 0**, who exhibited greater interaction with the available resources, achieved higher grades and completed more evaluative activities compared to the students in **Cluster 1**, who participated less. The density analysis of the attributes (see Figure 27) revealed clear differences between the two groups, with the students in **Cluster 0** demonstrating significantly higher dedication in terms of course time and active participation in the resources. These more frequent and in-depth interactions are strongly correlated with superior academic performance.

Furthermore, the analysis of students' grades (Figure 28) reinforces this correlation. **Cluster 0**, composed mainly of students with Grade A (78.3%), Grade B (17.4%), and Grade C (4.3%), reflects a group that engaged more intensively with the resources of the PA. On the other hand, **Cluster 1**, consisting primarily of students with Grade C (92.3%) and Grade B (7.7%), shows a clear association between lower engagement and lower academic performance. These data indicate that involvement with educational resources not only increases participation but also has a direct impact on students' academic outcomes. Therefore, these results confirm **H2**, providing evidence of a positive correlation between students' engagement with the PA resources and their academic performance.

The results obtained also provide an answer to **RQ3**, which investigates how the interpretations of the data obtained from the analyses validate the proposed PA. Through the combination of EDM techniques, self-regulation questionnaires, and interviews, it was possible to gain a comprehensive view of the effectiveness of the PA in supporting the development of SRL skills. The results of the quantitative analyses provide clear evidence that the PA promotes self-regulation by integrating resources that encourage reflection and learning monitoring. The interview analysis also reinforces these findings, revealing that students perceived improvements in the organization, planning, and monitoring of their academic activities, as well as an increase in autonomy and self-assessment skills. These qualitative reports corroborate the patterns identified in the quantitative analyses, validating the PA as an effective pedagogical tool for fostering SRL practices in VLEs.

The self-regulation questionnaires, administered at the beginning and end of the course, provided a direct measure of the evolution of SRL skills over time. Students who engaged more deeply with the resources of the PA reported a significant improvement in their ability to plan, monitor, and adjust their study practices, which corroborates the data collected in the EDM analyses.

**H3** was confirmed, as the quantitative analyses conducted through EDM and the self-regulation questionnaires indicate that the PA is effective in fostering students' SRL skills and in identifying distinct self-regulatory profiles. The clustering analysis clearly distinguished students with behavioral patterns aligned to a self-regulated learner profile—characterized by higher engagement, diversified use of learning strategies, and superior academic performance—from those exhibiting lower adoption of such strategies. The questionnaire results revealed significant improvements in key dimensions, particularly metacognitive strategies and self-evaluation, reinforcing the contribution of the architecture to enhancing students' abilities in planning, monitoring, and reflection.

Moreover, the qualitative findings from the focus group interviews corroborated these results, indicating that the resources integrated into the PA were perceived as facilitators of the self-regulated learning process. Taken together, these findings confirm that the PA not only supports the development of SRL but also enables the identification of distinct learner profiles, thereby providing valuable insights to guide personalized pedagogical interventions in VLE contexts. The following chapter presents the main conclusions of the study, highlighting its contributions to the field, its limitations, and directions for future research.

Chapter 7

# Conclusion

This study proposed, developed, and evaluated an AI-supported PA designed to foster SRL in a VLE. The research began with a systematic literature review aimed at identifying technologies and pedagogical approaches that effectively support SRL in digital environments. This review served as the theoretical and methodological foundation for the design of the proposed PA, guiding the integration of intelligent tools and learning analytics resources to promote goal setting, planning, monitoring, and reflection. Building on these findings, educational data mining (EDM) techniques were subsequently employed to identify distinct SRL profiles. These results informed the implementation of pedagogical mechanisms aimed at enhancing students' engagement, autonomy, and regulation of their learning processes.

Building upon these foundations, the research also pursued specific objectives related to the implementation and evaluation of the proposed PA. These included analyzing the correlations between students' engagement with the resources integrated into the architecture, their SRL skills, and their academic performance. The study further involved the development of the Time Tracker SRL plugin, a time management tool designed to monitor the time students dedicated to learning activities within the VLE. Another objective focused on assessing the effectiveness of the PA in fostering SRL skills through empirical investigations that combined self-regulation questionnaires, performance data analyses, and focus group interviews. Altogether, these objectives were successfully accomplished, and the quantitative and qualitative results consistently demonstrated the positive impact of the PA on promoting autonomy, engagement, and SRL behaviors among students.

As part of this research process, an initial PoC was conducted to identify key EDM techniques used to reveal SRL profiles in VLEs. Three clustering algorithms from different categories—partitional, hierarchical, and model-based—were applied to data collected from the OULAD dataset. Based on internal validation measures, K-Means ($K = 2$) yielded the best performance and identified two SRL profiles. These profiles correspond to **Cluster 0 (No self-regulation)** and **Cluster 1 (Evidence of self-regulation)**.

Further analysis of student performance showed that those in **Cluster 1**, who demonstrated greater use of SRL strategies, achieved higher academic outcomes, with an 83% pass rate compared to 46% in **Cluster 0**. These findings highlight the potential of clustering to group students based on their SRL profiles, allowing educators to offer more personalized support tailored to the specific needs of each group. This approach provides valuable insights into student behavior and helps guide targeted interventions in online learning environments.

Building upon these findings, the second PoC focused on examining the SRL profiles of students enrolled in a subsequent technical course at a public educational institution. For this purpose, interaction data extracted from Moodle were analyzed using EDM techniques, following a rigorous pre-processing stage to ensure data quality and reliability. The analysis employed clustering algorithms—namely K-Means, HDBSCAN, and Agglomerative Clustering—to identify distinct behavioral patterns. This stage complemented the initial analysis, aiming to inform and refine the design of the PA by capturing variations in student behavior across different educational contexts.

The main findings of this second analysis are as follows: analysis of events recorded in Moodle logs and their pre-processing resulted in refined datasets that enabled an in-depth analysis of student interaction patterns and performance. The K-Means, HDBSCAN, and Agglomerative Clustering algorithms were compared, with HDBSCAN and K-Means showing tied results, as 4 datasets presented better outcomes for each. It was observed that students with evidence of SRL profiles generally achieve higher academic performance in the courses, as the clusters with greater engagement in learning resources showed a higher proportion of students with high grades.

The results of the PoCs were crucial in guiding the design of the PA. The identification of distinct SRL profiles, such as students who exhibited higher engagement and better academic outcomes, emphasized the importance of creating an environment that actively promotes the development of these skills. The findings from the initial analyses, which revealed a correlation between self-regulation and academic success, directed the development and use of specific tools in the PA, such as the Time Tracker SRL, Completion Progress, Analytics Graphs, and Athena (OpenAI Chat), which were employed to help students monitor, plan, and adjust their learning strategies. Therefore, the data obtained were essential to the PA design, ensuring that its features were grounded in the real self-regulation needs of the students identified in the analyses.

Based on the results of the preliminary analysis, an Pedagogical AI-based Architecture was implemented with the aim of encourage SRL through technological tools integrated into VLE, within the context of an online Python extension course offered to students. This course was structured with interactive activities designed to encourage self-regulation, involving the use of various plugins, such as Athena (OpenAI Chat) for interactive support, Completion Progress for progress tracking, and the Time Tracker SRL plugin, specifically developed to monitor the time dedicated to learning activities.

The data from VLE logs, reports generated by the Configurable Reports plugin, and the Time Tracker SRL were subjected to a rigorous pre-processing process to build a structured dataset. Initially, a detailed data pre-processing step was carried out on Moodle interaction logs, which were subsequently classified according to the self-regulation strategies described in Zimmerman's model (ZIMMERMAN; MARTINEZ-PONS, 1986). This categorization clearly identified student actions related to strategies such as Help Seeking, Goal Setting and Planning, Self-Evaluation, Monitoring, Reviewing Activities, and Environmental Structuring, providing a solid foundation for subsequent analysis.

Based on this structured dataset, clustering techniques — including K-Means, HDBSCAN, and Agglomerative Clustering — were applied to identify behavioral patterns and group students according to their learning profiles. The results revealed that students who made more active use of the available resources, such as forums and interactive activities, achieved higher academic performance. Among the algorithms tested, K-Means achieved the best validation measures, demonstrating greater consistency in identifying self-regulated learner profiles. The course implementation, along with the installed plugins, played a key role in promoting student engagement and developing their autonomy, fostering more effective SRL and improving academic outcomes.

The findings indicate that the PA, through its various elements and resources, effectively supported the development of SRL skills in the VLE **(RQ1)** . Tools such as the Time Tracker SRL plugin, Completion Progress, and Athena (OpenAI Chat) facilitated goal setting, activity monitoring, and self-assessment, enabling students to plan their learning, track their progress, and reflect on outcomes. Students' perceptions reinforced this effect, as they consistently highlighted these tools as essential for organizing study routines, consolidating knowledge, and seeking help when needed.

A significant relationship was also observed between students' engagement with PA resources and their academic performance **(RQ2)**. The clustering results revealed two distinct profiles: students who actively engaged with monitoring and collaborative resources not only demonstrated more frequent use of SRL strategies but also achieved higher grades and greater participation in forums, while those with limited engagement tended to achieve lower outcomes. This association reinforces the central role of SRL behaviors in promoting academic success in online contexts.

Finally, the integration of quantitative and qualitative results validates the proposed PA as an effective framework for fostering SRL **(RQ3)**. The clustering analysis provided measurable evidence of different SRL profiles, while the interviews contextualized these profiles by capturing students' perceptions of how the architecture supported their learning processes. The convergence between statistical evidence and students' experiences confirms the PAs effectiveness and points to its potential scalability for broader educational contexts.

The research hypotheses were validated throughout the development of the study. The implemented PA proved effective in developing students' SRL skills and in identifying different SRL profiles within VLEs. The EDM analyses identified clear behavioral patterns, distinguishing profiles of students with higher and lower levels of self-regulation. The self-regulation questionnaires administered during the course confirmed the improvement of SRL skills among students who were more engaged with the PA resources. Additionally, the focus group interviews corroborated these results, revealing positive perceptions regarding the role of the PA in promoting awareness and strengthening self-regulation skills. Thus, the PA was validated as effective both in supporting the development of SRL skills and in identifying SRL profiles.

This research highlights the crucial role of integrating an AI-based Pedagogical Architecture in promoting SRL in VLEs. The clustering analyses carried out over several phases of the study consistently revealed a strong correlation between students' active use of technological resources and their academic success. Students who engaged more frequently with the resources available in the PA demonstrated significantly superior academic performance, emphasizing the impact of these technologies on enhancing autonomy, engagement, and learning outcomes. These findings not only confirm the effectiveness of SRL strategies in improving academic performance but also underscore the need for advanced technological tools to create more personalized and immersive learning experiences. This research contributes to the growing field of SRL in VLEs and provides a solid foundation for future educational technology innovations, particularly those that leverage data-driven approaches to support and improve student learning.

Unlike previous studies that typically focus on isolated technological tools, this research integrates several key resources, such as the Time Tracker SRL, Completion Progress, Analytics Graphs, and Athena, into a cohesive AI-based PA. This holistic integration provides continuous feedback, real-time data analysis, and personalized interventions tailored to the specific SRL profiles of students, identified through advanced EDM techniques. The development of tools like the Time Tracker SRL, specifically designed to support time management—a critical skill for self-regulation—further emphasizes the innovative nature of this work.

In addition, the qualitative analyses of the interviews conducted with students enrolled in the course played a crucial role in validating the resources of the PA. These interviews offered in-depth insights into students' perceptions of the effectiveness of the implemented tools, enabling the identification of aspects that enhance or limit the development of their self-regulation skills. By combining these qualitative findings with the quantitative evidence from interaction data and self-regulation questionnaires, the study presents a comprehensive validation of the PA, reinforcing its potential to foster SRL and improve academic outcomes.

Moreover, the ability of the PA to both promote the development of SRL skills and identify distinct self-regulation profiles across different educational contexts represents a significant advancement over the state of the art. The application of clustering algorithms to analyze student behavior in terms of their SRL abilities enables a more precise and informed approach to the personalization of pedagogical interventions. This data-driven behavioral analysis offers valuable insights that can guide strategic pedagogical decisions, aimed at enhancing academic performance and providing a more effective and adaptive learning experience.

Thus, this research goes beyond traditional approaches, offering a novel, scalable, and adaptable solution that enhances engagement and academic performance through the targeted development of SRL skills. By combining robust data analyses with practical pedagogical tools, this approach sets a new standard for how VLEs can promote autonomy, engagement, and academic success. Future innovations in educational technology can build upon this foundation to further refine and expand the capabilities of AI-driven pedagogical systems, providing even more effective, adaptive, and personalized learning experiences that more precisely meet the individual needs of students.

## 7.1 Limitations

We would like to highlight some limitations of this work. First, the relationship we made between the student's SRL profiles and their final result was based on a correlation analysis, the results may therefore not have exposed all the factors that could contribute to their approved or failed. So, for future work, it is interesting to consider other variables besides the number of clicks on the resources of the VLE.

The study was conducted over a relatively short period, which may limit the full understanding of the development of SRL skills, as these tend to evolve over longer periods. As a result, the available data may not fully capture long-term changes in SRL profiles. Future research could benefit from continuous monitoring of students over several semesters, providing a more robust analysis of the development of these skills over time.

Additionally, the EDM analyses relied heavily on log data from the VLE and self-regulation questionnaires completed by students. While these data are valuable, they may not capture all aspects of students' behavior and self-regulation strategies, particularly those occurring outside the learning management system. This could limit the comprehensive understanding of self-regulatory interactions and behaviors.

Finally, it is important to recognize the technical limitations of the tools adopted in this study. The monitoring performed through the Time Tracker SRL plugin and the Completion Progress relies on variables recorded in the Moodle logs, which may not fully capture the actual time devoted to activities or the quality of student engagement. These metrics, while essential for implementing large-scale analyses in VLEs, are inherently constrained by the platform's tracking capabilities, as they cannot account for passive time with an activity open, interruptions, or study conducted outside the system. Nevertheless, the use of these Moodle-generated variables is fundamental for enabling consistent, automated, and replicable monitoring of student behavior, serving as a practical and widely applicable basis for EDM in authentic educational contexts.

## 7.2 Publications

The following scientific publications are directly related to this doctoral project:

**Published manuscripts**

- Lima, Geycy; Araújo, Rafael D.; Dorça, Fabiano. Uma Análise dos Recursos Tecnológicos Utilizados na Estimulação da Aprendizagem Autorregulada em Ambientes Educacionais na Última Década. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 2020 . p. 732-741. Qualis A3
- Lima, Geycy and Costa, Juliete and Dorça, Fabiano and Araújo, Rafael. Exploring the Relationship between Students Engagement and Self-Regulated Learning: A Case Study using OULAD Dataset and Machine Learning Techniques. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 2023. Qualis A3
- Lima, Geycy and Costa, Juliete and Dorça, Fabiano and Araújo, Rafael. An Analysis of Technological Resources to Encourage Self-Regulated Learning Behavior in Virtual Learning Environments in the Last Decade. Int. J. Learn. Technol. 19, 1 (2024), 85–108. https://doi.org/10.1504/ijlt.2024.137897. Qualis B2
- Lima, Geycy and Costa, Juliete and Dorça, Fabiano and Araújo, Rafael. Arquitetura Pedagógica Apoiada por IA para o Desenvolvimento da Aprendizagem Autorregulada em Estudantes. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 2024. Qualis A3
- Costa, Juliete and Lima, Geycy and Dorça, Fabiano and Araújo, Rafael. Exploring Self-Regulated Learning in Virtual Environments: An Experimental Clustering-Based Approach. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 2025. Qualis A3
- Lima, Geycy and Costa, Juliete and Dorça, Fabiano and Araújo, Rafael. An AI-Supported Pedagogical Architecture to Foster Self-Regulated Learning in Virtual Environments. In Computer Applications in Engineering Education, 2025. Qualis A3

**Manuscripts under review**

- Costa, Juliete and Lima, Geycy and Dorça, Fabiano and Araújo, Rafael. Self-Regulated Learning Traits in Students' Behavior Interactions in a Ubiquitous Learning Environment. In Revista Brasileira de Informática na Educação. Qualis A4

## 7.3 Future Work

Although this research has successfully validated the proposed Artificial Intelligence (AI)-based Pedagogical Architecture for fostering SRL in VLEs, several opportunities remain open for future investigations and enhancements. First, it is recommended to extend the duration and scope of the studies. The implementation was carried out in a single course over a relatively short period, which may limit the understanding of the progressive development of SRL skills, as these tend to consolidate over longer time horizons. Longitudinal studies, encompassing multiple semesters and different subjects, may provide a more robust and in-depth view of the evolution of SRL in diverse educational contexts.

Second, the PA can be further strengthened through the incorporation of more advanced learning analytics techniques, such as Deep Learning and Natural Language Processing (NLP). These methods could enable more refined identification of SRL strategies and more precise personalization of pedagogical interventions. For instance, analyzing forum posts or chat interactions may reveal cognitive and metacognitive strategies that are not fully captured by traditional log data, thereby broadening the understanding of students' self-regulatory processes.

Additionally, it is recommended to explore the impact of the PA from the instructors' perspective. Providing teachers with detailed information on students' SRL profiles and strategies may support pedagogical decision-making, assisting in the planning of targeted interventions and the design of more effective teaching strategies. This approach contributes to the development of more holistic and collaborative learning ecosystems. In these environments, both students and instructors share responsibilities in the teaching–learning process.

In summary, these research directions expand the theoretical, methodological, and practical contributions of this study, while advancing the integration of artificial intelligence, learning analytics, and the promotion of SRL in digital education contexts.

# Bibliography

ACKERMANN, M. R. *et al.* Analysis of agglomerative clustering. **Algorithmica**, v. 69, n. 1, p. 184–215, 2014. Available at: <https://doi.org/10.1007/s00453-012-9717-4>.

AGGARWAL, C. C. **Data Mining: The Textbook**. Cham: Springer, 2015. Available at: <https://doi.org/10.1007/978-3-319-14142-8>.

AINSCOUGH, L. *et al.* Characterizing university students' self-regulated learning behavior using dispositional learning analytics. In: **Proceedings of the 5th International Conference on Higher Education Advances (HEAd'19)**. Valencia, Spain: Universitat Politècnica de València, 2019. p. 819–827. Available at: <https://doi.org/10.4995/HEAd19.2019.9153>.

AL-HATEM, A.; MASOOD, M.; AL-SAMARRAIE, H. Fostering student nurses' self-regulated learning with the second life environment: An empirical study. **Journal of Information Technology Education: Research**, Santa Rosa, CA, USA, v. 17, n. 1, p. 285–307, 2018. Available at: <https://doi.org/10.28945/4110>.

ALDOWAH, H.; AL-SAMARRAIE, H.; FAUZY, W. M. Educational data mining and learning analytics for 21st century higher education: A review and synthesis. **Telematics and Informatics**, Amsterdam, Netherlands, v. 37, n. 1, p. 13–49, 2019. Available at: <https://doi.org/10.1016/j.tele.2019.01.007>.

ALHAZBI, S. *et al.* Using learning analytics to measure self-regulated learning: A systematic review of empirical studies in higher education. **Journal of Computer Assisted Learning**, v. 40, n. 4, p. 1658–1674, 2024. Available at: <https://doi.org/10.1111/jcal.12982>.

ALMAAZMI, A. *et al.* Learning analytics and machine learning. In: **Proceedings of the 5th International Conference on Signal Processing and Information Security (ICSPIS)**. Dubai, United Arab Emirates: IEEE, 2022. p. 164–168. Available at: <https://doi.org/10.1109/ICSPIS57063.2022.10002701>.

ANDERSON, T. W.; DARLING, D. A. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. **The Annals of Mathematical Statistics**, Beachwood, OH, USA, v. 23, n. 2, p. 193–212, 1952. Available at: <https://doi.org/10.1214/aoms/1177729437>.

ARAKA, E. *et al.* Using educational data mining techniques to identify profiles in self-regulated learning: An empirical evaluation. **The International Review of Research in Open and Distributed Learning**, v. 23, n. 1, p. 131–162, 2022. Available at: <https://doi.org/10.19173/irrodl.v22i4.5401>.

BAHRI, A. *et al.* Blended learning integrated with innovative learning strategy to improve self-regulated learning. **International Journal of Instruction**, Ankara, Turkey, v. 14, n. 1, p. 779–794, 2021. Available at: <https://doi.org/10.29333/iji.2021.14147a>.

BAKER, R.; ISOTANI, S.; CARVALHO, A. Mineração de dados educacionais: Oportunidades para o brasil. **Revista Brasileira de Informática na Educação**, Porto Alegre, RS, Brazil, v. 19, n. 2, p. 3–13, 2011. Available at: <https://doi.org/10.5753/RBIE.2011.19.02.03>.

BARNARD-BRAK, L.; PATON, V. O.; LAN, W. Y. Profiles in self-regulated learning in the online learning environment. **The International Review of Research in Open and Distributed Learning**, Athabasca, AB, Canada, v. 11, n. 1, p. 61–80, 2010. Available at: <https://doi.org/10.19173/irrodl.v11i1.769>.

BARRIA-PINEDA, J.; GUERRA-HOLLSTEIN, J.; BRUSILOVSKY, P. A fine-grained open learner model for an introductory programming course. In: **Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization**. New York, NY, USA: Association for Computing Machinery, 2018. (UMAP '18), p. 53–61. ISBN 9781450355896. Available at: <https://doi.org/10.1145/3209219.3209242>.

BEHAR, P. *et al.* A pedagogical architecture from a hybrid course in a u.s. media course: Focused on digital competences. **Educação Unisinos**, São Leopoldo, RS, Brazil, v. 24, n. 1, p. 1–27, 2020. Available at: <https://doi.org/10.4013/edu.2020.241.42>.

BIANCARDI, C. *et al.* Apa2i - uma arquitetura pedagógica aberta, adaptativa e inteligente para construção cooperativa de conhecimento. **Revista Novas Tecnologias na Educação**, v. 18, n. 2, p. 131–140, jan. 2021. Available at: <https://seer.ufrgs.br/index.php/renote/article/view/110211>.

BIANCARDI, C.; MENEZES, C.; VILHAGRA, L. Uma arquitetura pedagógica para construção cooperativa de resenhas reflexivas no contexto de ensino a distância. In: **Anais do XXXI Simpósio Brasileiro de Informática na Educação**. Porto Alegre, RS, Brasil: SBC, 2020. p. 262–271. ISSN 0000-0000. Available at: <https://sol.sbc.org.br/index.php/sbie/article/view/12782>.

BODNER, G. M. Constructivism: A theory of knowledge. **Journal of Chemical Education**, v. 63, n. 10, p. 873–878, 1986. Available at: <https://doi.org/10.1021/ed063p873>.

BOEKAERTS, M. Motivated learning: bias in appraisals. **IJER**, v. 12, n. 3, p. 267–280, 1988. ISSN 0883-0355. Available at: <http://www.sciencedirect.com/science/article/pii/0883035588900055>.

BORREGO, M.; DOUGLAS, E. P.; AMELINK, C. T. Quantitative, qualitative, and mixed research methods in engineering education. **Journal of Engineering Education**, Hoboken, NJ, USA, v. 98, n. 1, p. 53–66, 2009. Available at: <https://doi.org/10.1002/j.2168-9830.2009.tb01005.x>.

BOTELHO, J. S. e Edson Pimentel e Silvia Dotta e W. Estudo comparativo de plugins moodle para análise e acompanhamento da aprendizagem. **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação - SBIE)**, v. 30, n. 1, p. 189, 2019. ISSN 2316-6533. Available at: <http://milanesa.ime.usp.br/rbie/index.php/sbie/article/view/8723>.

BROADBENT, J.; FULLER-TYSZKIEWICZ, M. Profiles in self-regulated learning and their correlates for online and blended learning students. **Educational Technology Research and Development**, New York, NY, USA, v. 66, n. 6, p. 1435–1455, 2018. Available at: <https://doi.org/10.1007/s11423-018-9595-9>.

BROADBENT, J.; PANADERO, E.; FULLER-TYSZKIEWICZ, M. Effects of mobile-app learning diaries vs online training on specific self-regulated learning components. **Educational Technology Research and Development**, v. 68, n. 5, p. 2351–2372, 2020. Available at: <https://doi.org/10.1007/s11423-020-09781-6>.

CAMPELLO, R. J. G. B.; MOULAVI, D.; SANDER, J. Density-based clustering based on hierarchical density estimates. In: **Advances in Knowledge Discovery and Data Mining: 17th Pacific-Asia Conference (PAKDD 2013)**. Gold Coast, Australia: Springer, 2013. p. 160–172. Available at: <https://doi.org/10.1007/978-3-642-37456-2_14>.

CAO, T. *et al.* Utilizing clickstream data to reveal the time management of self-regulated learning in a higher education online learning environment. **Interactive Learning Environments**, Routledge, v. 31, n. 10, p. 6555–6572, 2023. Available at: <https://doi.org/10.1080/10494820.2022.2042031>.

ÇEBI, A.; GÜYER, T. Students' interaction patterns in different online learning activities and their relationship with motivation, self-regulated learning strategy and learning performance. **Education and Information Technologies**, Dordrecht, Netherlands, v. 25, n. 5, p. 4035–4053, 2020. Available at: <https://doi.org/10.1007/s10639-020-10151-1>.

CERÓN, J. *et al.* Self-regulated learning in massive online open courses: A state-of-the-art review. **IEEE Access**, v. 9, p. 511–528, 2021. Available at: <https://doi.org/10.1109/ACCESS.2020.3045913>.

CHEN, C.-M.; HUANG, S.-H. Web-based reading annotation system with an attention-based self-regulated learning mechanism for promoting reading performance. **British Journal of Educational Technology**, v. 45, n. 5, p. 959–980, 2014. Available at: <https://doi.org/10.1111/bjet.12119>.

CHEN, L. *et al.* Design of learning analytics dashboard supporting metacognition. In: **Proceedings of the International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2019)**. Cagliari, Italy: IADIS, 2019. p. 211–218. Available at: <https://doi.org/10.33965/celda2019_201911L022>.

CHIU, T. K. *et al.* Teacher support and student motivation to learn with artificial intelligence (ai) based chatbot. **Interactive Learning Environments**, Routledge, v. 0, n. 0, p. 1–17, 2023. Available at: <https://doi.org/10.1080/10494820.2023.2172044>.

CHIU, T. K. F. A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: A case of chatgpt. **Educational Technology Research and Development**, 2024. Available at: <https://doi.org/10.1007/s11423-024-10366-w>.

CICCHINELLI, A. *et al.* Finding traces of self-regulated learning in activity streams. In: **Proceedings of the 8th International Conference on Learning Analytics and Knowledge**. New York, NY, USA: Association for Computing Machinery, 2018. (LAK '18), p. 191–200. ISBN 9781450364003. Available at: <https://doi.org/10.1145/3170358.3170381>.

COBOS, R.; WILDE, A.; ZALUSKA, E. Predicting attrition from massive open online courses in futurelearn and edx. In: **Proceedings of the LAK17 FutureLearn Workshop on "What does it Mean to Learn?" in the Age of AI**. Vancouver, Canada: CEUR Workshop Proceedings, 2017. Available at: <https://ceur-ws.org/Vol-1967/FLMOOCS_Paper5.pdf>.

COMAN, C. *et al.* Online teaching and learning in higher education during the coronavirus pandemic: Students' perspective. **Sustainability**, v. 12, n. 24, p. 10367, 2020. Available at: <https://doi.org/10.3390/su122410367>.

COSTA, J.; DORÇA, F.; ARAÚJO, R. Avaliação do comportamento de estudantes em um ambiente educacional ubíquo. In: **Anais do XXXI Simpósio Brasileiro de Informática na Educação**. Porto Alegre, RS, Brasil: SBC, 2020. p. 182–191. Available at: <https://sol.sbc.org.br/index.php/sbie/article/view/12774>.

CRONBACH, L. J. Coefficient alpha and the internal structure of tests. **Psychometrika**, v. 16, n. 3, p. 297–334, 1951. Available at: <https://doi.org/10.1007/BF02310555>.

DELEN, E.; LIEW, J.; WILLSON, V. Effects of interactivity and instructional scaffolding on learning: Self-regulation in online video-based environments. **Computers & Education**, v. 78, p. 312–320, 2014. Available at: <https://doi.org/10.1016/j.compedu.2014.06.018>.

DILLENBOURG, P.; SCHNEIDER, D.; SYNTETA, P. Virtual learning environments. In: DIMITRACOPOULOU, A. (Ed.). **Proceedings of the 3rd Hellenic Conference on Information and Communication Technologies in Education**. Rhodes, Greece: [s.n.], 2002. p. 3–18. Available at: <https://telearn.archives-ouvertes.fr/hal-00190701>.

DO, C. B.; BATZOGLOU, S. What is the expectation maximization algorithm? **Nature Biotechnology**, v. 26, n. 8, p. 897–899, 2008. Available at: <https://doi.org/10.1038/nbt1406>.

DURALL, E.; GROS, B. Learning analytics as a metacognitive tool. In: **Proceedings of the 6th International Conference on Computer Supported Education (CSEDU 2014)**. Barcelona, Spain: SCITEPRESS, 2014. p. 380–384. Available at: <https://doi.org/10.5220/0004933203800384>.

EFKLIDES, A. Interactions of metacognition with motivation and affect in self-regulated learning: The masrl model. **Educational Psychologist**, v. 46, n. 1, p. 6–25, 2011. Available at: <https://doi.org/10.1080/00461520.2011.538645>.

ELSAYED, A. A. *et al.* Research in learning analytics and educational data mining to measure self-regulated learning: A systematic review. In: **World Conference on Mobile and Contextual Learning**. Delft, Netherlands: LearnTechLib, 2019. p. 46–53. Available at: <https://www.learntechlib.org/d/210600/>.

FARAHMAND, A.; DEWAN, M. A. A.; LIN, F. Student-facing educational dashboard design for online learners. In: **Proceedings of the IEEE International Conference on Dependable, Autonomic and Secure Computing (DASC) / Pervasive Intelligence and Computing (PiCom) / Cloud and Big Data Computing (CBDCom) / Cyber Science and Technology Congress (CyberSciTech)**. Calgary, Canada: IEEE, 2020. p. 345–349. Available at: <https://doi.org/10.1109/DASC-PICom-CBDCom-CyberSciTech49142.2020.00067>.

FELDMAN-MAGGOR, Y.; BLONDER, R.; TUVI-ARAD, I. Let them choose: Optional assignments and online learning patterns as predictors of success in online general chemistry courses. **The Internet and Higher Education**, v. 55, p. 100867, 2022. ISSN 1096-7516. Available at: <https://www.sciencedirect.com/science/article/pii/S1096751622000239>.

FIRAT, M. Integrating ai applications into learning management systems to enhance e-learning. **Instructional Technology and Lifelong Learning**, v. 4, n. 1, p. 1–14, 2023. Available at: <https://doi.org/10.52911/itall.1244453>.

FISCHER, C. *et al.* Mining big data in education: Affordances and challenges. **Review of Research in Education**, v. 44, n. 1, p. 130–160, 2020. Available at: <https://doi.org/10.3102/0091732X20903304>.

FRANÇA, R. S. d.; TEDESCO, P. C. d. A. R. Um modelo colaborativo para a aprendizagem do pensamento computacional aliado à autorregulação. In: **Anais do Simpósio Brasileiro de Informática na Educação (SBIE)**. Dourados, MS, Brasil: Sociedade Brasileira de Computação, 2014. p. 1133–1142. Available at: <https://doi.org/10.5753/cbie.sbie.2014.1133>.

FUNG, C. Y.; ABDULLAH, M.; HASHIM, S. Improving self-regulated learning through personalized weekly e-learning journals: A time series quasi-experimental study. **The E-Journal of Business Education & Scholarship of Teaching**, v. 13, n. 1, p. 30–45, 2019. Available at: <https://files.eric.ed.gov/fulltext/EJ1239150.pdf>.

FURLANETTO, G. *et al.* Algoritmos de agrupamento aplicados à detecção de fraudes. In: **Anais da XIII Escola Regional de Alto Desempenho de São Paulo**. Porto Alegre, RS, Brasil: SBC, 2022. p. 29–32. ISSN 0000-0000. Available at: <https://sol.sbc.org.br/index.php/eradsp/article/view/21924>.

GADELLA, L. U. *et al.* Application of learning analytics to study the accuracy of self-reported working patterns in self-regulated learning questionnaires. In: **Proceedings of the IEEE Global Engineering Education Conference (EDUCON)**. Porto, Portugal: IEEE, 2020. p. 1201–1205. Available at: <https://doi.org/10.1109/EDUCON45650.2020.9125405>.

GAETA, M. *et al.* Metacognitive learning environment: A semantic perspective. **Journal of e-Learning and Knowledge Society**, v. 7, n. 2, p. 69–80, 2011. Available at: <https://doi.org/10.20368/1971-8829/522>.

GAMIE, E. A.; EL-SEOUD, M. S. A.; SALAMA, M. A. A layered-analysis of the features in higher education data set. In: **Proceedings of the 2019 8th International Conference on Software and Information Engineering**. New York, NY, USA: ACM, 2019. (ICSIE 2019), p. 237–242. Available at: <https://doi.org/10.1145/3328833.3328850>.

GARCIA, R.; FALKNER, K.; VIVIAN, R. Systematic literature review: Self-regulated learning strategies using e-learning tools for computer science. **Computers & Education**, v. 123, p. 150 – 163, 2018. ISSN 0360-1315. Available at: <http://www.sciencedirect.com/science/article/pii/S0360131518301064>.

GUEDES, G. T. A. **UML 2: Uma Abordagem Prática**. São Paulo: Novatec, 2009.

HADWIN, A. F.; JäRVELä, S.; MILLER, M. Self-regulated, co-regulated, and socially shared regulation of learning. In: ZIMMERMAN, B. J.; SCHUNK, D. H. (Ed.). **Handbook of Self-Regulation of Learning and Performance**. New York: Routledge, 2011. p. 65–84. Available at: <https://doi.org/10.4324/9780203839010>.

HAMDANE, K.; MHOUTI, A. E.; MASSAR, M. How can learning analytics techniques improve the learning process? an overview. In: **Proceedings of the International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)**. Meknes, Morocco: IEEE, 2022. p. 1–5. Available at: <https://doi.org/10.1109/IRASET52964.2022.9738003>.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. H. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. ed. New York: Springer, 2009. Available at: <https://doi.org/10.1007/978-0-387-84858-7>.

HAYNES, C. C. The role of self-regulated learning in the design, implementation, and evaluation of learning analytics dashboards. In: **Proceedings of the Seventh ACM Conference on Learning @ Scale**. New York, NY, USA: Association for Computing Machinery, 2020. (L@S '20), p. 297–300. ISBN 9781450379519. Available at: <https://doi.org/10.1145/3386527.3406732>.

HE, H. *et al.* Visual analysis of the time management of learning multiple courses in online learning environment. In: **Proceedings of the IEEE Visualization Conference (VIS)**. Vancouver, Canada: IEEE, 2019. p. 56–60. Available at: <https://doi.org/10.1109/VISUAL.2019.8933778>.

HEMMLER, Y. M.; IFENTHALER, D. Self-regulated learning strategies in continuing education: A systematic review and meta-analysis. **Educational Research Review**, p. 100629, 2024. Available at: <https://doi.org/10.1016/j.edurev.2024.100629>.

HLOSTA, M.; ZDRAHAL, Z.; ZENDULKA, J. Ouroboros: Early identification of at-risk students without models based on legacy data. In: **Proceedings of the Seventh International Learning Analytics & Knowledge Conference**. New York, NY, USA: Association for Computing Machinery, 2017. (LAK '17), p. 6–15. ISBN 9781450348706. Available at: <https://doi.org/10.1145/3027385.3027449>.

HSU, P.-Y. Self-regulated learning strategies in technology-enhanced learning: Teaching self-monitoring and english oral presentation skills in graduate program. In: **Proceedings of the 2020 11th International Conference on E-Education,**

**E-Business, E-Management, and E-Learning**. New York, NY, USA: Association for Computing Machinery, 2020. (IC4E 2020), p. 25–29. ISBN 9781450372947. Available at: <https://doi.org/10.1145/3377571.3377579>.

HUANG, H. *et al.* Use of behavior dynamics to improve early detection of at-risk students in online courses. **Mobile Networks and Applications**, v. 27, n. 1, p. 441–452, 2022. Available at: <https://doi.org/10.1007/s11036-021-01844-z>.

HUANG, T.-C. *et al.* Developing a self-regulated oriented online programming teaching and learning system. In: **Proceedings of the IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)**. Wellington, New Zealand: IEEE, 2014. p. 115–120. Available at: <https://doi.org/10.1109/TALE.2014.7062599>.

ILVES, K.; LEINONEN, J.; HELLAS, A. Supporting self-regulated learning with visualizations in online learning environments. In: **Proceedings of the 49th ACM Technical Symposium on Computer Science Education**. Baltimore, MD, USA: Association for Computing Machinery, 2018. p. 257–262. Available at: <https://doi.org/10.1145/3159450.3159509>.

IVANKOVA, N. V.; CRESWELL, J. W.; STICK, S. L. Using mixed-methods sequential explanatory design: From theory to practice. **Field Methods**, v. 18, n. 1, p. 3–20, 2006. Available at: <https://doi.org/10.1177/1525822X05282260>.

JACAúNA, R. D. *et al.* Construção de uma arquitetura pedagógica apoiada em jogos digitais: uma experiência de aprendizagem. In: **Anais do XXVIII Workshop de Informática na Escola**. Porto Alegre, RS, Brasil: SBC, 2022. p. 100–109. Available at: <https://sol.sbc.org.br/index.php/wie/article/view/22341>.

JANSEN, R. S. *et al.* Supporting learners' self-regulated learning in massive open online courses. **Computers & Education**, v. 146, p. 103771, 2020. ISSN 0360-1315. Available at: <https://www.sciencedirect.com/science/article/pii/S0360131519303240>.

JHA, N. I.; GHERGULESCU, I.; MOLDOVAN, A.-N. Oulad mooc dropout and result prediction using ensemble, deep learning and regression techniques. In: **Proceedings of the 11th International Conference on Computer Supported Education**. Heraklion, Crete, Greece: SciTePress, 2019. p. 154–164. Available at: <https://doi.org/10.5220/0007767901540164>.

JONASSEN, D. H. **Objectivism versus Constructivism: Do We Need a New Philosophical Paradigm?** [s.n.], 1991. v. 39. 5–14 p. Available at: <https://doi.org/10.1007/BF02296434>.

JR., R. A. C. *et al.* Self-regulated learning in online learning environments: Strategies for remote learning. **Information and Learning Sciences**, Emerald Publishing Limited, v. 121, n. 5/6, p. 321–329, 2020. Available at: <https://doi.org/10.1108/ILS-04-2020-0114>.

JUNUS, K.; SANIOSO, H. B.; SADITA, L. The use of self-monitoring tools for linear algebra course in student-centered e-learning environment. In: **2014 IEEE Frontiers in Education Conference (FIE) Proceedings**. Madrid, Spain: IEEE, 2014. p. 1–4. Available at: <https://ieeexplore.ieee.org/document/7044296>.

KAPP, F.; BRAUN, I.; HARA, T. Evaluating lectures through the use of mobile devices: Auditorium mobile classroom service (amcs) as a means to bring evaluation to the next level. In: **Proceedings of the 8th International Conference on Computer Supported Education**. Rome, Italy: SciTePress, 2016. p. 251–257. Available at: <https://doi.org/10.5220/0005811202510257>.

KARAMI, S. *et al.* The impact of application of electronic portfolio on undergraduate english majors' writing proficiency and their self-regulated learning. **International Journal of Instruction**, v. 12, n. 1, p. 1319–1334, 2019. Available at: <https://doi.org/10.29333/iji.2019.12184a>.

KARIMI, H. *et al.* Online academic course performance prediction using relational graph convolutional neural network. In: **Proceedings of the 13th International Conference on Educational Data Mining (EDM 2020)**. International Educational Data Mining Society, 2020. Available at: <https://educationaldatamining.org/files/conferences/EDM2020/papers/paper_45.pdf>.

KAUFFMAN, D. F.; ZHAO, R.; YANG, Y.-S. Effects of online note taking formats and self-monitoring prompts on learning from online text: Using technology to enhance self-regulated learning. **Contemporary Educational Psychology**, v. 36, n. 4, p. 313–322, 2011. Available at: <https://doi.org/10.1016/j.cedpsych.2011.04.001>.

KAUFMAN, L.; ROUSSEEUW, P. J. **Finding Groups in Data: An Introduction to Cluster Analysis**. Hoboken, NJ: Wiley, 2009. Available at: <https://doi.org/10.1002/9780470316801>.

KHIAT, H. Using automated time management enablers to improve self-regulated learning. **Active Learning in Higher Education**, v. 23, n. 1, p. 3–15, 2019. Available at: <https://doi.org/10.1177/1469787419866304>.

KIA, F. S. *et al.* How patterns of students dashboard use are related to their achievement and self-regulatory engagement. In: **Proceedings of the Tenth International Conference on Learning Analytics & Knowledge**. New York, NY, USA: Association for Computing Machinery, 2020. (LAK '20), p. 340–349. ISBN 9781450377126. Available at: <https://doi.org/10.1145/3375462.3375472>.

KIM, D. *et al.* Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in south korea. **Computers & Education**, v. 127, p. 233–251, 2018. ISSN 0360-1315. Available at: <https://www.sciencedirect.com/science/article/pii/S0360131518302331>.

KINNARI-KORPELA, H.; SUHONEN, S. Active learning: How to promote it? In: **Engaging Engineering Education: SEFI 48th Annual Conference Proceedings**. Enschede, Netherlands: University of Twente, 2020. Available at: <https://www.theseus.fi/handle/10024/356404>.

KITSANTAS, A. Fostering college students' self-regulated learning with learning technologies. **Hellenic Journal of Psychology**, Psychological Society of Northern Greece, v. 10, n. 3, p. 235–252, 2013. Available at: <https://pseve.gr/wp-content/uploads/2018/03/Volume10_Issue3_Kitsantas.pdf>.

KUZILEK, J.; HLOSTA, M.; ZDRÁHAL, Z. Open university learning analytics dataset. **Scientific Data**, v. 4, p. 170171, 2017. Available at: <https://doi.org/10.1038/sdata.2017.171>.

LACHHEB, A. *et al.* Ai in higher education: A bibliometric analysis, synthesis, and a critique of research. **The Internet and Higher Education**, v. 67, p. 101021, 2025. ISSN 1096-7516. Available at: <https://www.sciencedirect.com/science/article/pii/S1096751625000302>.

LAW, C.-Y. *et al.* User perceptions of using an open learner model visualisation tool for facilitating self-regulated learning. In: **Proceedings of the Nineteenth Australasian Computing Education Conference**. New York, NY, USA: Association for Computing Machinery, 2017. (ACE '17), p. 55–64. ISBN 9781450348232. Available at: <https://doi.org/10.1145/3013499.3013502>.

LAWRIE, G. A. *et al.* Development of scaffolded online modules to support self-regulated learning in chemistry concepts. In: **Technology and Assessment Strategies for Improving Student Learning in Chemistry**. Washington, DC: American Chemical Society, 2016, (ACS Symposium Series, v. 1235). p. 1–21. Available at: <https://doi.org/10.1021/bk-2016-1235.ch001>.

LAZAR, J.; FENG, J.; HOCHHEISER, H. **Research Methods in Human-Computer Interaction**. Morgan Kaufmann, 2017. ISBN 9780128093436. Available at: <https://books.google.com.br/books?id=hbkxDQAAQBAJ>.

LEE, S.; BARKER, T.; KUMAR, V. Models of elearning: The development of a learner-directed adaptive elearning system. In: GREENER, S.; ROSPIGLIOSI, A. (Ed.). **Proceedings of the 10th European Conference on e-Learning**. Brighton, UK: Academic Conferences, 2011. p. 390–398. ISBN 9781908272225.

LEONARDOU, A.; RIGOU, M.; GAROFALAKIS, J. Opening user model data for motivation and learning: The case of an adaptive multiplication game. In: **Proceedings of the 11th International Conference on Computer Supported Education (CSEDU 2019)**. Heraklion, Crete, Greece: SciTePress, 2019. p. 383–390. Available at: <https://doi.org/10.5220/0007735603830390>.

LI, H. *et al.* Measuring behaviors and identifying indicators of self-regulation in computer-assisted language learning courses. **Research and Practice in Technology Enhanced Learning**, v. 13, n. 1, p. 19, 2018. Available at: <https://doi.org/10.1186/s41039-018-0087-7>.

LI, Q.; BAKER, R.; WARSCHAUER, M. Using clickstream data to measure, understand, and support self-regulated learning in online courses. **The Internet and Higher Education**, v. 45, p. 100727, 2020. ISSN 1096-7516. Available at: <https://www.sciencedirect.com/science/article/pii/S1096751620300038>.

LI, Q. *et al.* Can student-facing analytics improve online students' effort and success by affecting how they explain the cause of past performance? **Computers & Education**, v. 185, p. 104517, 2022. ISSN 0360-1315. Available at: <https://www.sciencedirect.com/science/article/pii/S0360131522000884>.

LIMA, A. F. O.; PIMENTEL, E. P. Mecanismos para suporte à auto-regulação da aprendizagem do estudante. In: **Anais do XXIV Simpósio Brasileiro de Informática na Educação**. Sociedade Brasileira de Computação, 2013. p. 296. ISSN 2316-6533. Available at: <http://milanesa.ime.usp.br/rbie/index.php/sbie/article/view/2508>.

LIMA, G.; ARAÚJO, R.; DORÇA, F. Uma análise dos recursos tecnológicos utilizados na estimulação da aprendizagem autorregulada em ambientes educacionais na Última década. In: **Anais do XXXI Simpósio Brasileiro de Informática na Educação**. Porto Alegre, RS, Brasil: SBC, 2020. p. 732–741. ISSN 0000-0000. Available at: <https://sol.sbc.org.br/index.php/sbie/article/view/12829>.

LIMA, G. D. *et al.* An analysis of technological resources to encourage self-regulated learning behaviour in virtual learning environments in the last decade. **Int. J. Learn. Technol.**, Inderscience Publishers, Geneva 15, CHE, v. 19, n. 1, p. 85–108, apr 2024. ISSN 1477-8386. Available at: <https://doi.org/10.1504/ijlt.2024.137897>.

LIMA, J. C. e Rafael Araújo e Fabiano Dorça e G. Exploring the relationship between students engagement and self-regulated learning: A case study using site dataset and machine learning techniques. **Anais do Simpósio Brasileiro de Informática na Educação (SBIE)**, p. 1154–1165, 2023. ISSN 0000-0000. Available at: <https://sol.sbc.org.br/index.php/sbie/article/view/26743>.

LIMA, M. *et al.* Uso de atributos de código para classificação da facilidade de questões de codificação. In: **Anais do Simpósio Brasileiro de Educação em Computação**. Porto Alegre, RS, Brasil: SBC, 2021. p. 113–122. ISSN 0000-0000. Available at: <https://sol.sbc.org.br/index.php/educomp/article/view/14477>.

LIN, J. Chatgpt and moodle walk into a bar: A demonstration of ai's mind-blowing impact on e-learning. **SSRN Electronic Journal**, 2023. Available at: <https://doi.org/10.2139/ssrn.4393445>.

LIU, J.; ZHENG, Y.; JIANG, X. Design of augmented interactive video based on script engine. In: **Proceedings of the International Workshop on Artificial Intelligence and Education**. New York, NY, USA: Association for Computing Machinery, 2019. (WAIE 2019), p. 1–5. ISBN 9781450376235. Available at: <https://doi.org/10.1145/3397453.3397455>.

LLERENA, J.; MORáN, N. Álava; ZAMORA-GALINDO, J. Learning analytics for student academic tracking: A comparison between analytics graphs and edwiser reports. In: **Proceedings of the 2021 2nd International Conference on Information Systems and Software Technologies (ICI2ST)**. Virtual, Quito, Ecuador: IEEE, 2021. p. 101–107. Available at: <https://doi.org/10.1109/ICI2ST51859.2021.00022>.

LOCKYER, L.; HEATHCOTE, E.; DAWSON, S. Informing pedagogical action: Aligning learning analytics with learning design. **American Behavioral Scientist**, v. 57, n. 10, p. 1439–1459, 2013. Available at: <https://doi.org/10.1177/0002764213479367>.

MALDONADO-MAHAUAD, J.; PÉREZ-SANAGUSTÍN, M.; BEYLE, C. A questionnaire for measuring self-regulated learning in massive open online courses. In: **Proceedings of the 2020 46th Latin American Computing Conference (CLEI)**. Loja, Ecuador: IEEE, 2020. p. 400–409. Available at: <https://doi.org/10.1109/CLEI52000.2020.00053>.

MANGANELLO, F. *et al.* A dashboard to monitor self-regulated learning behaviours in online professional development. **International Journal of Distance Education Technologies**, v. 19, n. 1, p. 18–34, 2021. Available at: <https://doi.org/10.4018/IJDET.2021010102>.

MANSO-VÁZQUEZ, M.; CAEIRO-RODRÍGUEZ, M.; LLAMAS-NISTAL, M. xapi-srl: Uses of an application profile for self-regulated learning based on the analysis of learning strategies. In: **2015 IEEE Frontiers in Education Conference (FIE)**. El Paso, TX, USA: IEEE, 2015. p. 1–8. Available at: <https://doi.org/10.1109/FIE.2015.7344049>.

MCINNES JOHN HEALY, S. A. L. **The hdbscan Clustering Library**. 2024. Available at: <https://hdbscan.readthedocs.io/en/latest/>. Accessed on: 15/07/2024.

MCKINNEY, W. Data structures for statistical computing in python. In: **Proceedings of the 9th Python in Science Conference**. Austin, TX, USA: SciPy, 2010. p. 51–56. Available at: <https://doi.org/10.25080/Majora-92bf1922-00a>.

MENEZES, C. d.; JUNIOR, A. d. C.; ARAGÓN, R. Arquiteturas pedagógicas para aprendizagem em rede. **Série de livros texto da CEIE-SBC. Disponível em: https://ieducacao. ceie-br. org/arquiteturas-pedagogicas/. Acessado: em**, v. 7, p. 21–08, 2020.

MENEZES, C. S. d. *et al.* Arquiteturas pedagógicas para a aprendizagem em rede no contexto do seminário integrador. **RENOTE: Revista Novas Tecnologias na Educação**, Porto Alegre, v. 11, n. 2, 2013. Available at: <https://seer.ufrgs.br/index.php/renote/article/view/43645>.

MENEZES, R. N. e Maria Dalpiaz e C. Arquitetura pedagógica para construção colaborativa de conceituações. **Anais do Workshop de Informática na Escola**, v. 1, n. 1, p. 1653–1662, 2009. ISSN 2316-6541. Available at: <http://milanesa.ime.usp.br/rbie/index.php/wie/article/view/2150>.

MENEZES, V. B. e José Netto e C. Agent-based conceptual framework for collaborative educational resources adaptation in virtual learning environments. **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação - SBIE)**, v. 28, n. 1, p. 1087, 2017. ISSN 2316-6533. Available at: <http://milanesa.ime.usp.br/rbie/index.php/sbie/article/view/7637>.

MENTARI, N.; SUBCHAN, W.; SUPENO. Self-regulated learning (srl) design to enhance student achievement in the environmental science course facilitating by e-learning. **Journal of Physics: Conference Series**, v. 1521, n. 4, p. 042111, 2020. Available at: <https://doi.org/10.1088/1742-6596/1521/4/042111>.

MOCCOZET, L.; TARDY, C. Open group-sourcing: A social learning model based on peer production and feedback. In: **2014 Eighth International Conference on Complex, Intelligent and Software Intensive Systems**. Birmingham, UK: IEEE, 2014. p. 491–497. Available at: <https://doi.org/10.1109/CISIS.2014.70>.

MOLENAAR, I. *et al.* Personalized visualizations to promote young learners' srl: The learning path app. In: **Proceedings of the Tenth International Conference on Learning Analytics & Knowledge**. New York, NY, USA: Association for Computing Machinery, 2020. (LAK '20), p. 330–339. ISBN 9781450377126. Available at: <https://doi.org/10.1145/3375462.3375465>.

_____. Measuring self-regulated learning and the role of ai: Five years of research using multimodal multichannel data. **Computers in Human Behavior**, v. 139, p. 107540, 2023. ISSN 0747-5632. Available at: <https://www.sciencedirect.com/science/article/pii/S0747563222003600>.

MONTGOMERY, A. P. *et al.* Using learning analytics to explore self-regulated learning in flipped blended learning music teacher education. **British Journal of Educational Technology**, v. 50, n. 1, p. 114–127, 2019. Available at: <https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12590>.

MOODLE. **Analytics graphs**. 2024. Available at: <https://moodle.org/plugins/block_analytics_graphs>. Accessed on: 15/02/2024.

_____. **Configurable Reports**. 2024. Available at: <https://moodle.org/plugins/block_configurable_reports>. Accessed on: 11/02/2024.

_____. **Welcome to the Moodle community**. 2024. Available at: <https://moodle.org>. Accessed on: 10/02/2024.

MOON, T. K. The expectation-maximization algorithm. **IEEE Signal Processing Magazine**, v. 13, n. 6, p. 47–60, 1996. Available at: <https://doi.org/10.1109/79.543975>.

MORENO, J.; PINEDA, A. F. A framework for automated formative assessment in mathematics courses. **IEEE Access**, v. 8, p. 30152–30159, 2020. Available at: <https://doi.org/10.1109/ACCESS.2020.2973026>.

MORGAN, T.; PETTERSEN, I. N.; DAALHUIZEN, J. Exploring socially shared regulated learning in pbl using group digital learning diaries: A study design. In: **Proceedings of the 22nd International Conference on Engineering and Product Design Education (E&PDE 2020)**. Herning, Denmark: The Design Society, 2020. ISBN 978-1-912254-10-1. Available at: <https://doi.org/10.35199/EPDE.2020.14>.

MORRIS, B. *et al.* Comparing badges and learning goals in low- and high-stakes learning contexts. **Journal of Computing in Higher Education**, v. 31, n. 3, p. 573–603, 2019. Available at: <https://doi.org/10.1007/s12528-019-09228-9>.

MUELLER, D.; STROHMEIER, S. Design characteristics of virtual learning environments: State of research. **Computers & Education**, v. 57, n. 4, p. 2505–2516, 2011. ISSN 0360-1315. Available at: <https://doi.org/10.1016/j.compedu.2011.06.017>.

MURTAGH, F.; CONTRERAS, P. **Algorithms for Hierarchical Clustering: An Overview**. 2011. 86–97 p. Available at: <https://doi.org/10.1002/widm.53>.

MôNEGO, L. C. **Monitoramento da Interação e Colaboração no Ambiente Moodle: Um Estudo de Caso**. Dissertation (Master of Science) — Universidade Federal de Santa Maria, 2019. Available at: <https://repositorio.ufsm.br/handle/1/20812>.

NATIONS, U. **Sustainable Development Goals**. 2023. Available at: <https://www.un.org/sustainabledevelopment/>. Accessed on: 10/11/2023.

NEITZEL, S.; RENSING, C.; BELLHÄUSER, H. Concept, design and first evaluation of a mobile learning diary application with access to a learning record store. In: **Proceedings of the 7th Workshop on Awareness and Reflection in Technology-Enhanced Learning (ARTEL) co-located with the 12th European Conference on Technology Enhanced Learning (EC-TEL 2017)**. Tallinn, Estonia: CEUR-WS, 2017. (CEUR Workshop Proceedings, v. 1997). ISSN 1613-0073. Available at: <https://ceur-ws.org/Vol-1997/paper3.pdf>.

OGATA, H. *et al.* Using learning analytics to support computer-assisted language learning. In: **Proceedings of the 25th International Conference on Computers in Education (ICCE 2017)**. Christchurch, New Zealand: Asia-Pacific Society for Computers in Education, 2017. p. 908–913. ISBN 9789869401265. Available at: <https://library.apsce.net/index.php/ICCE/article/view/2196>.

O'LEARY, R.; RAMSDEN, A. J. Virtual learning environments. In: DAVIES, P. (Ed.). **Handbook for Learning and Teaching in Economics**. Bristol, UK: Economics Learning and Teaching Support Network, 2002. p. 1–30. ISBN 1904486061. Available at: <https://www.economicsnetwork.ac.uk/handbook/printable/vle_v5.pdf>.

PANADERO, E. A Review of Self-regulated Learning: Six Models and Four Directions for Research. **Frontiers in Psychology**, v. 8, p. 422, 2017. ISSN 1664-1078. Available at: <https://www.frontiersin.org/article/10.3389/fpsyg.2017.00422>.

PAPAMITSIOU, Z.; ECONOMIDES, A. A. The impact of on-demand metacognitive help on effortful behaviour: A longitudinal study using task-related visual analytics. **Journal of Computer Assisted Learning**, v. 37, n. 1, p. 109–126, 2021. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jcal.12472>.

PARASKEVA, F. *et al.* Applying pbl and srl to enhance learning strategies and collaboration in a blended learning environment. In: **Learning Technology for Education Challenges**. Springer International Publishing, 2017. p. 85–96. Available at: <https://doi.org/10.1007/978-3-319-62743-4_8>.

PARDO, A.; HAN, F.; ELLIS, R. A. Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance. **IEEE Transactions on Learning Technologies**, v. 10, n. 1, p. 82–92, 2017. Available at: <https://doi.org/10.1109/TLT.2016.2639508>.

PARK, Y.; JO, I.-H. Factors that affect the success of learning analytics dashboards. **Educational Technology Research and Development**, Springer, v. 67, n. 6, p. 1547–1571, 2019. Available at: <https://doi.org/10.1007/s11423-019-09693-0>.

PEACH, R. L. *et al.* Data-driven unsupervised clustering of online learner behaviour. **npj Science of Learning**, Nature Publishing Group, v. 4, n. 1, p. 14, 2019. Available at: <https://doi.org/10.1038/s41539-019-0054-0>.

PÉREZ-ÁLVAREZ, R.; MALDONADO-MAHAUAD, J.; PÉREZ-SANAGUSTÍN, M. Design of a tool to support self-regulated learning strategies in moocs. **Journal of Universal Computer Science**, v. 24, n. 8, p. 1090–1109, 2018. Available at: <https://doi.org/10.3217/jucs-024-08-1090>.

PÉREZ-ÁLVAREZ, R. *et al.* Notemyprogress: A tool to support learners' self-regulated learning strategies in mooc environments. In: LAVOUÉ, É. *et al.* (Ed.). **Data Driven Approaches in Digital Education**. Cham: Springer International Publishing, 2017. p. 460–466. ISBN 978-3-319-66610-5. Available at: <https://doi.org/10.1007/978-3-319-66610-5_43>.

PHODONG, K.; SUPNITHI, T.; KONGKACHANDRA, R. A framework of computer-based learning system based on self-regulated model in english writing. In: **2019 14th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)**. Chiang Mai, Thailand: IEEE, 2019. p. 1–6. Available at: <https://doi.org/10.1109/iSAI-NLP48611.2019.9045562>.

PINTRICH, P. R.; GROOT, E. V. D. Motivational and self-regulated learning components of classroom academic performance. **Journal of Educational Psychology**, v. 82, n. 1, p. 33–40, 1990. Available at: <https://doi.org/10.1037/0022-0663.82.1.33>.

PINTRICH, P. R. *et al.* **A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)**. Ann Arbor, MI, 1991. Available at: <https://files.eric.ed.gov/fulltext/ED338122.pdf>.

PUUSTINEN, M.; PULKKINEN, L. Models of self-regulated learning: A review. **Scandinavian Journal of Educational Research**, v. 45, n. 3, p. 269–286, 2001. Available at: <https://doi.org/10.1080/00313830120074206>.

RABELO, A. *et al.* Educational data mining and learning analytics: A review of educational management in e-learning. **Information Discovery and Delivery**, Emerald Publishing Limited, v. 52, n. 2, p. 149–163, 2024. Available at: <https://doi.org/10.1108/IDD-10-2022-0099>.

RAKOVIć, M. *et al.* Examining the critical role of evaluation and adaptation in self-regulated learning. **Contemporary Educational Psychology**, v. 68, p. 102027, 2022. ISSN 0361-476X. Available at: <https://www.sciencedirect.com/science/article/pii/S0361476X21000862>.

RIAZY, S.; SIMBECK, K.; SCHRECK, V. Fairness in learning analytics: Student at-risk prediction in virtual learning environments. In: **Proceedings of the 12th International Conference on Computer Supported Education (CSEDU 2020) - Volume 1**. SciTePress, 2020. p. 15–25. Available at: <https://doi.org/10.5220/0009324100150025>.

RIBEIRO, A. C. **MP-SocioAVA: Modelo Pedagógico com foco nas interações sociais em um Ambiente Virtual de Aprendizagem**. Thesis (Doctor in Education) — Universidade Federal do Rio Grande do Sul, 2019. Available at: <http://hdl.handle.net/10183/201307>.

ROBAL, T. *et al.* Intellieye: Enhancing mooc learners' video watching experience through real-time attention tracking. In: **Proceedings of the 29th ACM Conference on Hypertext and Social Media**. Baltimore, MD, USA: ACM, 2018. p. 106–114. ISBN 978-1-4503-5427-1. Available at: <https://doi.org/10.1145/3209542.3209547>.

RODRIGUEZ, F. *et al.* Using clickstream data mining techniques to understand and support first-generation college students in an online chemistry course. In: **LAK21:**

**11th International Learning Analytics and Knowledge Conference**. New York, NY, USA: Association for Computing Machinery, 2021. (LAK21), p. 313–322. ISBN 9781450389358. Available at: <https://doi.org/10.1145/3448139.3448169>.

ROMERO, L. *et al.* Supporting self-regulated learning and personalization using eportfolios: A semantic approach based on learning paths. **International Journal of Educational Technology in Higher Education**, v. 16, n. 1, p. 16, 2019. Available at: <https://doi.org/10.1186/s41239-019-0146-1>.

ROUSSEEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. **Journal of Computational and Applied Mathematics**, Elsevier, v. 20, n. 1, p. 53–65, 1987. Available at: <https://doi.org/10.1016/0377-0427(87)90125-7>.

RUIPÉREZ-VALIENTE, J. A. *et al.* From sensor data to educational insights. **Sensors**, v. 22, n. 21, p. 8556, 2022. Available at: <https://doi.org/10.3390/s22218556>.

SAFSOUF, Y.; MANSOURI, K.; POIRIER, F. Understand the influence of learning analytics dashboards on learner self-regulation and academic success. In: **2022 IEEE Global Engineering Education Conference (EDUCON)**. Tunis, Tunisia: IEEE, 2022. p. 1044–1047. ISBN 978-1-6654-4434-7. Available at: <https://doi.org/10.1109/EDUCON52537.2022.9766741>.

SANTOS, B. G. d. *et al.* A criatividade como um caminho interdisciplinar para inovação na arquitetura pedagógica. **Cuadernos de Educación y Desarrollo**, v. 17, n. 6, p. e8546, jun. 2025. Available at: <https://ojs.cuadernoseducacion.com/ojs/index.php/ced/article/view/8546>.

SAQR, U. F. M.; TEDRE, M. How learning analytics can early predict under-achieving students in a blended medical education course. **Medical Teacher**, Taylor & Francis, v. 39, n. 7, p. 757–767, 2017. PMID: 28421894. Available at: <https://doi.org/10.1080/0142159X.2017.1309376>.

SELVI, S. T.; PANNEERSELVAM, K. A self-regulated learning approach for programming language using cloud-based learning management system. In: **2012 International Conference on Recent Trends in Information Technology (ICRTIT)**. Chennai, Tamil Nadu, India: IEEE, 2012. p. 191–196. ISBN 978-1-4673-1599-9. Available at: <https://doi.org/10.1109/ICRTIT.2012.6206810>.

SHARP, L. A.; SHARP, J. H. Enhancing student success in online learning experiences through the use of self-regulation strategies. **Journal on Excellence in College Teaching**, v. 27, n. 2, p. 57–75, 2016. Available at: <https://celt.miamioh.edu/index.php/JECT/article/view/389>.

SIEMENS, G.; LONG, P. Penetrating the fog: Analytics in learning and education. **EDUCAUSE Review**, v. 46, n. 5, p. 30–40, 2011. Available at: <https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education>.

SILVA, F.; MENEZES, C.; JUNIOR, A. C. Ensino introdutório de programação: Um estudo rumo ao uso das arquiteturas pedagógicas. In: **Anais do XXXIV Simpósio Brasileiro de Informática na Educação**. Porto Alegre, RS, Brasil: SBC, 2023. p. 428–438. ISSN 0000-0000. Available at: <https://sol.sbc.org.br/index.php/sbie/article/view/26682>.

SILVA, F. X. d.; MENEZES, C. S. d.; JUNIOR, A. N. d. C. Arquitetura pedagógica baseada no pensamento computacional para a compreensão do problema & portfólio de aprendizagem. **Revista Novas Tecnologias na Educação**, v. 21, n. 2, p. 1–12, dez. 2023. Available at: <https://seer.ufrgs.br/index.php/renote/article/view/137719>.

SILVA, J. C. S. **Uma Abordagem de Learning Analytics para a Autorregulação da Aprendizagem de Estudantes em Sala de Aula Invertida**. Thesis (Doctor in Philosophy) — Universidade Federal de Pernambuco, Local da Publicação (opcional), Mês de Defesa (opcional) 2018. Informação Adicional (opcional). Available at: <https://repositorio.ufpe.br/handle/123456789/29849>.

SINATRA, A. M. Utilizing the generalized intelligent framework for tutoring to encourage self-regulated learning. In: **Foundations of Augmented Cognition**. Springer International Publishing, 2014. p. 257–264. Available at: <https://doi.org/10.1007/978-3-319-07527-3_24>.

SIROTHEAU, S. *et al.* Aprendizagem de iniciantes em algoritmos e programação: Foco nas competências de autoavaliação. In: **Anais do Simpósio Brasileiro de Informática na Educação (SBIE)**. Aracaju, SE, Brasil: SBC, 2011. p. 754–763. Available at: <http://milanesa.ime.usp.br/rbie/index.php/sbie/article/view/1637>.

SONEGO, A. H. S. **ARQPED-MOBILE: Uma arquitetura pedagógica com foco na aprendizagem móvel**. Thesis (Doctor in Education) — Universidade Federal do Rio Grande do Sul, 2019. Available at: <http://hdl.handle.net/10183/197443>.

SONEGO, A. H. S. *et al.* Formação de professores: uma arquitetura pedagógica com foco na m-learning. **Revista Novas Tecnologias na Educação**, Porto Alegre, v. 16, n. 2, p. 301–310, 2018. Available at: <https://doi.org/10.22456/1679-1916.89279>.

SPEARMAN, C. The proof and measurement of association between two things. **The American Journal of Psychology**, v. 15, n. 1, p. 72–101, 1904. Available at: <https://www.jstor.org/stable/1412159>.

SPILIOTOPOULOS, D. *et al.* A mixed-reality interaction-driven game-based learning framework. In: **Proceedings of the 11th International Conference on Management of Digital EcoSystems**. New York, NY, USA: Association for Computing Machinery, 2019. (MEDES '19), p. 229–236. ISBN 9781450362382. Available at: <https://doi.org/10.1145/3297662.3365802>.

SU, J.-M. A rule-based self-regulated learning assistance scheme to facilitate personalized learning with adaptive scaffoldings: A case study for learning computer software. **Computer Applications in Engineering Education**, v. 28, n. 3, p. 536–555, 2020. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cae.22222>.

TAN, J. P.-L. *et al.* A collaborative video annotation and analytics environment (covaa) intervention: User experiences and reflections of teacher-practitioners. In: **Proceedings of the 26th International Conference on Computers in Education (ICCE 2018)**. Asia-Pacific Society for Computers in Education, 2018. Available at: <https://doi.org/10.13140/RG.2.2.28861.15843>.

TAN, P.-N. *et al.* **Introduction to Data Mining**. 2. ed. Boston, MA: Pearson, 2016. ISBN 9780133128901. Available at: <https://www.pearson.com/en-us/subject-catalog/p/Tan-Introduction-to-Data-Mining-2nd-Edition/P200000003204>.

TAN, T. Y. *et al.* What can completion time of quizzes tell us about students' motivations and learning strategies? **Journal of Computing in Higher Education**, v. 32, n. 1, p. 1–17, 2020. Available at: <https://doi.org/10.1007/s12528-019-09239-6>.

TANG, L. M.; KAY, J. Gamification: Metacognitive scaffolding towards long term goals? In: **Proceedings of the 4th Workshop on Personalization Approaches for Learning Environments (PALE 2014)**. [s.n.], 2014. (CEUR Workshop Proceedings, v. 1181), p. 63–68. Available at: <https://ceur-ws.org/Vol-1181/pale2014_paper_08.pdf>.

TANG, Y.; FAN, A. An integrated approach to self-regulated learning platform enhanced with web 2.0 technology. In: **2011 IEEE 3rd International Conference on Communication Software and Networks (ICCSN)**. Xi'an, China: IEEE, 2011. p. 236–239. Available at: <https://ieeexplore.ieee.org/document/6013817/>.

TZIMAS, D. E.; DEMETRIADIS, S. N. Impact of learning analytics guidance on student self-regulated learning skills, performance, and satisfaction: A mixed methods study. **Education Sciences**, MDPI, v. 14, n. 1, p. 92, 2024. Available at: <https://doi.org/10.3390/educsci14010092>.

URDAN, T. **Statistics in Plain English, Third Edition**. Taylor & Francis, 2010. ISBN 9780203851173. Available at: <http://books.google.com.br/books?id=kNNZiszH6GMC>.

WANG, T.-H. Developing web-based assessment strategies for facilitating junior high school students to perform self-regulated learning in an e-learning environment. **Computers & Education**, v. 57, n. 2, p. 1801–1812, 2011. ISSN 0360-1315. Available at: <https://www.sciencedirect.com/science/article/pii/S036013151100011X>.

WENTROBA, J.; VOGT, P.; BOTELHO, L. d. L. R. Objetivos do desenvolvimento sustentável e o contexto educacional brasileiro. **Revista de Estudos Interdisciplinares**, v. 5, n. 2, p. 110–124, 2023. Available at: <https://revistas.ceeinter.com.br/revistadeestudosinterdisciplinar/article/view/412>.

WINNE, P. H.; HADWIN, A. F. Studying as self-regulated learning. In: HACKER, D. J.; DUNLOSKY, J.; GRAESSER, A. C. (Ed.). **Metacognition in Educational Theory and Practice**. Mahwah, NJ: Erlbaum, 1998. p. 277–304. Available at: <https://doi.org/10.4324/9781410602350-12>.

WONG, J. *et al.* Exploring sequences of learner activities in relation to self-regulated learning in a massive open online course. **Computers & Education**, v. 140, p. 103595, 10 2019. Available at: <https://doi.org/10.1016/j.compedu.2019.103595>.

World Health Organization. **WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 11 March 2020**. Geneva, Switzerland: World Health Organization, 2020. Accessed: 2024-03-17. Available at: <https://www.who.int/news-room/speeches/item/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>.

WU, T.-T. *et al.* Promoting self-regulation progress and knowledge construction in blended learning via chatgpt-based learning aid. **Journal of Educational Computing Research**, SAGE Publications, 2024. Available at: <https://doi.org/10.1177/07356331231191125>.

XIA, Q. *et al.* The mediating effects of needs satisfaction on the relationships between prior knowledge and self-regulated learning through artificial intelligence chatbot. **British Journal of Educational Technology**, v. 54, n. 4, p. 967–986, 2023. Available at: <https://doi.org/10.1111/bjet.13305>.

YAMADA, M. *et al.* Learning analytics of the relationships among self-regulated learning, learning behaviors, and learning performance. **Research and Practice in Technology Enhanced Learning**, v. 12, n. 1, p. 13, 2017. Available at: <https://doi.org/10.1186/s41039-017-0053-9>.

YE, C.; BISWAS, G. Early prediction of student dropout and performance in moocs using higher granularity temporal information. **Journal of Learning Analytics**, v. 1, n. 3, p. 169–172, 2014. Available at: <https://doi.org/10.18608/jla.2014.13.14>.

YE, D.; PENNISI, S. Using trace data to enhance students' self-regulation: A learning analytics perspective. **The Internet and Higher Education**, v. 54, p. 100855, 2022. ISSN 1096-7516. Available at: <https://www.sciencedirect.com/science/article/pii/S1096751622000112>.

YOT-DOMÍNGUEZ, C.; MARCELO, C. University students' self-regulated learning using digital technologies. **International Journal of Educational Technology in Higher Education**, v. 14, n. 1, p. 38, 2017. Available at: <https://doi.org/10.1186/s41239-017-0076-8>.

YUN, H.; FORTENBACHER, A.; PINKWART, N. Improving a mobile learning companion for self-regulated learning using sensors. In: **Proceedings of the 9th International Conference on Computer Supported Education - Volume 1: CSEDU**. Porto, Portugal: SciTePress, 2017. p. 531–536. Available at: <https://doi.org/10.5220/0006375405310536>.

ZAR, J. H. Spearman rank correlation. In: **Encyclopedia of Biostatistics**. Wiley, 2005. Available at: <https://doi.org/10.1002/0470011815.b2a15150>.

ZHANG, X.; CHENG, H. N. H. Mining the patterns of graduate students' self-regulated learning behaviors in a negotiated online academic reading assessment. In: **Proceedings of the 2019 3rd International Conference on E-Society, E-Education and E-Technology**. New York, NY, USA: Association for Computing Machinery, 2019. (ICSET 2019), p. 109–114. ISBN 9781450372305. Available at: <https://doi.org/10.1145/3355966.3355977>.

ZHAO, Z. *et al.* Studentlyzer for analyzing and visualizing e-learning data. In: **Proceedings of the International MultiConference of Engineers and Computer Scientists 2019**. Hong Kong: [s.n.], 2019. (Lecture Notes in Engineering and Computer Science, v. 2239), p. 147–152. Available at: <https://www.iaeng.org/publication/IMECS2019/IMECS2019_pp147-152.pdf>.

ZIMMERMAN, B. J. Becoming a self-regulated learner: Which are the key subprocesses? **Contemporary Educational Psychology**, v. 11, n. 4, p. 307–313, 1986. ISSN 0361-476X. Available at: <http://www.sciencedirect.com/science/article/pii/0361476X86900275>.

ZIMMERMAN, B. J.; MARTINEZ-PONS, M. Development of a structured interview for assessing student use of self-regulated learning strategies. **American Educational Research Journal**, v. 23, n. 4, p. 614–628, 1986. Available at: <https://doi.org/10.3102/00028312023004614>.

ZORRILLA, M.; GARCÍA, D.; ÁLVAREZ, E. A decision support system to improve e-learning environments. In: **Proceedings of the EDBT/ICDT 2010 Workshops**. Lausanne, Switzerland: ACM, 2010. p. 1–8. Available at: <https://doi.org/10.1145/1754239.1754252>.

# Appendix

# APPENDIX A

# Adapted Questionnaire

**Motivação:**

- ❏ Motivação Intrínseca

  1. Num curso como esse, eu prefiro o material de estudo que desperte a minha curiosidade, mesmo que seja difícil de aprender.

- ❏ Motivação Extrínseca

  2. Se puder, eu quero obter notas melhores neste curso que a maioria dos outros estudantes.

- ❏ Valorização da Atividade

  3. Eu acho que serei capaz de usar o que aprendi neste curso em outros cursos.
  4. Eu estou muito interessado no conteúdo deste curso.

- ❏ Autoeficácia para Aprendizado

  5. Tenho certeza que posso fazer um bom trabalho nas tarefas e testes neste curso.

**Estratégias de Aprendizagem Cognitivas:**

- ❏ Repetição

  6. Quando estudando para este curso, eu leio as minhas anotações de aula e materiais de estudo várias vezes.

- ❏ Elaboração

  7. Quando eu estudo para este curso, eu extraio informações de diferentes fontes, tais como aulas, leituras e discussões.

- ❏ Organização

  8. Quando eu estudo as leituras para este curso disciplina, eu faço tópicos, anotações, figuras, gráficos dentre outros do material para me ajudar a organizar minhas ideias.

**Estratégias de Aprendizagem Metacognitivas:**

- ❏ Planejamento Estratégico

  9. Estabeleço metas para me ajudar a gerenciar o tempo que gasto aprendendo.

- ❏ Definição de Metas

  10. Eu cumpro as metas que estabeleci para mim nas minhas atividades.

- ❏ Autoavaliação

  11. Eu me comunico com meus colegas e professores através de fóruns para refletir sobre o conteúdo que estou aprendendo no curso.

- ❏ Autossatisfação

  12. Eu tento entender como o que aprendi impacta meu trabalho/prática ou dinâmica de estudos.

- ❏ Autoeficácia

  13. Quando sou confrontado com um problema no meu curso, geralmente consigo encontrar várias soluções.

**Estratégias de gerenciamento de recursos:**

- ❏ Procurar Ajuda

  14. Eu procuro alguém que tenha conhecimento do conteúdo do curso para que eu possa consultá-lo quando precisar de ajuda.

  15. Compartilho meus problemas com meus colegas de classe online, para que saibamos com o que estamos trabalhando e como resolvê-los.

- ❏ Gerenciamento de Tempo

  16. Embora não tenhamos que frequentar as aulas diárias, ainda tento distribuir meu tempo de estudo uniformemente entre os dias.

- ❏ Regulação do Esforço

  17. Eu trabalho duro para fazer tudo bem feito neste curso, mesmo que eu não goste do que estamos fazendo.

- ❏ Gerenciamento do Ambiente de Estudo

  18. Encontro um lugar confortável para estudar e escolho um horário de estudo para evitar muita distração.

# APPENDIX B

# Focus Group Interview Script

The following questions were used as a semi-structured interview protocol during the focus group sessions. The purpose of this guide was to explore students' perceptions regarding their interaction with the VLE, the resources they considered most relevant, the challenges they faced, and the perceived impact of these features on their learning process. This instrument was applied as part of the qualitative stage of the research methodology, complementing the quantitative analyses and contributing to the triangulation of results.

1. How do you evaluate your interaction with the Moodle platform during the course? Did you encounter any difficulties using it?

2. Which resources did you find most interesting within the platform?

3. Do you believe these resources contributed to the learning process of the course content?

4. Did the self-assessment process (quiz) provided in the tool help you in understanding the content studied during this period?

5. Were the collaboration resources among students within Moodle important for learning the content?

# Annex

# ANNEX A

# Motivated Strategies for Learning Questionnaire

❏ Motivação Intrínseca

1. Numa matéria como essa, eu prefiro materiais de estudo que realmente me desafiam, assim aprenderei novas coisas.
2. Numa matéria como essa, eu prefiro o material de estudo que desperte a minha curiosidade, mesmo que seja difícil de aprender.
3. A coisa mais satisfatória para mim nesta disciplina é tentar entender o conteúdo da forma mais completa possível.
4. Quando tenho oportunidade nesta matéria, eu escolho realizar tarefas de onde eu possa aprender, mesmo que elas não garantam uma boa nota.

❏ Motivação Extrínseca

1. Conseguir uma boa nota nesta matéria é a coisa mais satisfatória para mim agora.
2. A coisa mais importante para mim agora é melhorar a minha média geral, portanto, minha principal preocupação nessa matéria é conseguir uma boa nota.
3. Se puder, eu quero obter notas melhores nesta matéria que a maioria dos outros estudantes.
4. Eu quero ir bem nessa matéria porque é importante mostrar minha capacidade para minha família, meus amigos, meu chefe, ou outros.

❏ Valorização da Atividade

1. Eu acho que serei capaz de usar o que aprendi nesta disciplina em outras disciplinas.

2. É importante para mim aprender o material de estudo nesta matéria.
3. Eu estou muito interessado no conteúdo desta disciplina.
4. Eu acho que o material de estudo nesta matéria é útil para eu aprender.
5. Eu gosto dos assuntos abordados desta disciplina.

❏ Controle do Aprendizado

1. Se eu estudar da forma apropriada, então serei capaz de aprender o material de estudo desta disciplina.
2. A culpa é somente minha se não aprender o material de estudo nesta disciplina.
3. Se eu me esforçar o suficiente, então eu entenderei o material de estudo desta disciplina.
4. Se eu não entender o material de estudo, é porque eu não me esforcei o suficiente.

❏ Autoeficácia para Aprendizado

1. Eu acredito que ganharei uma ótima nota nesta matéria.
2. Tenho certeza de que posso entender o mais difícil material de estudo apresentado nas leituras para esta disciplina.
3. Tenho certeza que posso aprender conceitos básicos ensinados nesta disciplina.
4. Tenho certeza de que posso entender o material de estudo mais complexo apresentado pelo professor nesta disciplina.
5. Tenho certeza que posso fazer um bom trabalho nas tarefas e testes nesta disciplina.
6. Eu espero ir bem nessa matéria.
7. Tenho certeza que posso dominar as habilidades ensinadas nesta matéria.
8. Considerando a dificuldade desta disciplina, o professor e minhas habilidades, eu acho que irei bem nessa matéria.

❏ Ansiedade em Testes

1. Quando eu faço um teste, eu penso o quanto eu me saio mal em comparação aos outros estudantes.
2. Quando eu faço um teste, penso nas questões das outras partes do teste que não consigo responder.
3. Ao fazer testes, eu penso nas consequências de não ir bem.

4. Eu fico ansioso e preocupado quando faço uma prova.
5. Eu sinto meu coração batendo rápido quando eu faço uma prova.

**Estratégias de Aprendizagem Cognitivas:**

❑ Repetição

1. Quando eu estudo para esta matéria, eu repito o material de estudo para mim mesmo várias vezes.
2. Quando estudando para esta matéria, eu leio as minhas anotações de aula e materiais de estudo várias vezes.
3. Eu memorizo palavras-chave para me lembrar de conceitos importantes para essa matéria.
4. Eu faço listas de itens importantes para esta disciplina e memorizo as listas.

❑ Elaboração

1. Quando eu estudo para esta matéria, eu extraio informações de diferentes fontes, tais como aulas, leituras e discussões.
2. Eu tento relacionar ideias deste conteúdo aos conteúdos de outras disciplinas sempre que possível.
3. Quando lendo para essa matéria, eu tento relacionar o material de estudo com o que eu já sei.
4. Quando eu estudo para esta disciplina, eu escrevo pequenos resumos das principais ideias das minhas leituras e anotações de aula.
5. Eu tento entender o material de estudo nesta matéria fazendo conexões entre as leituras e os conceitos das aulas.
6. Eu tento aplicar ideias das leituras desta disciplina em outras atividades da aula, tais como palestra e discussão.

❑ Organização

1. Quando eu estudo as leituras para esta disciplina, eu faço tópicos do material para me ajudar a organizar minhas ideias.
2. Quando estudo para esta matéria eu vou nas leituras e as minhas anotações de aulas e tento achar as ideias mais importantes.
3. Eu faço gráficos, diagramas ou tabelas simples para me ajudar a organizar o material de estudo.

4. Quando eu estudo para esta disciplina, eu vou para as minhas anotações de aula e faço uma lista de tópicos dos conceitos importantes.

**Estratégias de Aprendizagem Metacognitivas:**

❑ Planejamento Estratégico

1. Ao planejar meu aprendizado, adapto estratégias que funcionaram no passado.
2. Estabeleço metas para me ajudar a gerenciar o tempo que gasto aprendendo.
3. Estabeleço metas de longo prazo (mensais ou anuais) para mim mesmo, a fim de direcionar minhas atividades de aprendizado.

❑ Definição de Metas

1. Eu cumpro as metas que estabeleci para mim nas minhas atividades.
2. Sinto-me preparado para a maioria das exigências do meu curso.
3. Estabeleço padrões pessoais de desempenho no meu curso.

❑ Auto Avaliação

1. Eu sei o quanto aprendi quando terminei uma tarefa.
2. Eu me faço muitas perguntas sobre o material do curso quando estou estudando para um curso online.
3. Eu me comunico com meus colegas para saber como estou indo nas minhas aulas online.
4. Eu me comunico com meus colegas para descobrir se o que estou aprendendo é diferente do que eles estão aprendendo.

❑ Auto Satisfação

1. Eu tento entender como o que aprendi impacta meu trabalho/prática.
2. Considero como o que aprendi se relaciona com meus colegas.

❑ Auto Eficácia

1. Consigo permanecer calmo quando enfrento dificuldades no meu curso porque posso confiar em minhas habilidades.
2. Quando sou confrontado com um problema no meu curso, geralmente consigo encontrar várias soluções.
3. O que quer que aconteça no meu curso, geralmente consigo lidar com isso.

**Estratégias de gerenciamento de recursos:**

- ❏ Procurar Ajuda

  1. Eu procuro alguém que tenha conhecimento do conteúdo do curso para que eu possa consultá-lo quando precisar de ajuda.
  2. Compartilho meus problemas com meus colegas de classe online, para que saibamos com o que estamos trabalhando e como resolvê-los.
  3. Se necessário, tento encontrar meus colegas de forma presencial.
  4. Eu sou persistente em obter ajuda do professor por e-mail.

- ❏ Gerenciamento de Tempo

  1. Eu aloco tempo extra de estudo para meus cursos online porque sei que isso demanda muito tempo.
  2. Eu tento agendar o mesmo horário todos os dias ou toda semana para estudar para meus cursos online, e observo o horário.
  3. Embora não tenhamos que frequentar as aulas diárias, ainda tento distribuir meu tempo de estudo uniformemente entre os dias.

- ❏ Regulação do Esforço

  1. Eu geralmente me sinto tão entediado e com tanta preguiça quando estudo para esta matéria, que eu desisto antes de terminar o que havia planejado fazer.
  2. Eu trabalho duro para fazer tudo bem feito nesta matéria, mesmo que eu não goste do que estamos fazendo.
  3. Quando o trabalho na disciplina está difícil, eu desisto ou estudo somente as partes fáceis.
  4. Mesmo quando o material de estudo está chato e desinteressante, eu administro isso para me manter estudando até terminar.

- ❏ Gerenciamento do Ambiente de Estudo

  1. Escolho o local onde estudo para evitar muita distração.
  2. Encontro um lugar confortável para estudar.
  3. Eu sei onde posso estudar com mais eficiência para cursos online.
  4. Escolho um horário com poucas distrações para estudar para meus cursos online.