

The FLoRA Engine: Using Analytics to Measure and Facilitate Learners' Own Regulation Activities

Xinyu Li¹, Yizhou Fan^{2*}, Tongguang Li³, Mladen Raković⁴, Shaveen Singh⁵, Joep van der Graaf⁶, Lyn Lim⁷, Johanna Moore⁸, Inge Molenaar⁹, Maria Bannert¹⁰ and Dragan Gašević¹¹

Abstract

The focus of education is increasingly on learners' ability to regulate their own learning within technology-enhanced learning environments. Prior research has shown that self-regulated learning (SRL) leads to better learning performance. However, many learners struggle to productively self-regulate their learning, as they typically need to navigate the myriad of cognitive, metacognitive, and motivational processes that SRL demands. To address these challenges, the FLoRA engine is developed to help students, workers, and professionals improve their SRL skills and become productive lifelong learners. FLoRA incorporates several learning tools that are grounded in SRL theory and enhanced with learning analytics (LA), aimed at improving learners' mastery of different SRL skills. The engine tracks learners' SRL behaviours during a learning task and provides automated scaffolding to help learners effectively regulate their learning. The main contributions of FLoRA include (1) creating instrumentation tools that unobtrusively collect intensively sampled, fine-grained, and temporally ordered trace data about learners' learning actions; (2) building a trace parser that uses LA and related analytical techniques (e.g., process mining) to model and understand learners' SRL processes; and (3) providing a scaffolding module that presents analytics-based adaptive, personalized scaffolds based on students' learning progress. The architecture and implementation of the FLoRA engine are also discussed in this paper.

Notes for Practice

- Empirical studies have underscored the positive correlation between self-regulated learning (SRL) proficiency and enhanced learning outcomes. However, the intricacies involved in SRL, spanning cognitive, metacognitive, and motivational processes, present challenges to many learners. The FLoRA engine is designed and developed to detect, measure, and facilitate complicated SRL.
- The main contribution of this paper is the description of the design and implementation of an open-source SRL support engine that can unobtrusively collect fine-grained and temporally ordered trace data, map trace data to corresponding SRL processes, conduct analysis to understand learners' learning progress, and provide corresponding as-needed personalized scaffolds.

Keywords

Learning analytics, self-regulated learning, scaffolding, learning tools, trace data.

Submitted: 14/01/2024 — **Accepted:** 05/12/2024 — **Published:** 23/01/2025

¹ Email: xinyu.li1@monash.edu Address: Centre for Learning Analytics, Faculty of Information Technology, Monash University, Australia. ORCID iD: <https://orcid.org/0000-0003-2681-4451>

^{2*} Corresponding author Email: fyz@pku.edu.cn Address: Graduate School of Education, Peking University, Beijing, China. ORCID iD: <https://orcid.org/0000-0003-2777-1705>

³ Email: tongguang.li@monash.edu Address: Centre for Learning Analytics, Faculty of Information Technology, Monash University, Australia. ORCID iD: <https://orcid.org/0000-0003-4616-5268>

⁴ Email: mladen.rakovic@monash.edu Address: Centre for Learning Analytics, Faculty of Information Technology, Monash University, Australia. ORCID iD: <https://orcid.org/0000-0002-1413-1103>

⁵ Email: shaveen.singh@monash.edu Address: Centre for Learning Analytics, Faculty of Information Technology, Monash University, Australia. ORCID iD: <https://orcid.org/0000-0002-7862-8047>

⁶ Email: joep.vandergaaf@ru.nl Address: Behavioural Science Institute, Radboud University, Nijmegen 6500 HE, Netherlands. ORCID iD: <https://orcid.org/0000-0003-1205-2193>

⁷ Email: lyn.lim@tum.de Address: TUM School of Social Sciences and Technology, Technical University of Munich, Munich 80333, Germany. ORCID iD: <https://orcid.org/0000-0002-0617-5552>

⁸ Email: j.moore@ed.ac.uk Address: School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, United Kingdom. ORCID iD: <https://orcid.org/0000-0001-7247-6823>

⁹ Email: inge.molenaar@ru.nl Address: Behavioural Science Institute, Radboud University, Nijmegen 6500 HE, Netherlands. ORCID iD: <https://orcid.org/0000-0003-4639-2524>

¹⁰ Email: maria.bannert@tum.de Address: TUM School of Social Sciences and Technology, Technical University of Munich, Munich 80333, Germany. ORCID iD: <https://orcid.org/0000-0001-7045-2764>

¹¹ Email: dragan.gasevic@monash.edu Address: Centre for Learning Analytics, Faculty of Information Technology, Monash University, Australia. ORCID iD: <https://orcid.org/0000-0001-9265-1908>

1. Introduction

Self-regulated learning (SRL) is a complex ensemble of cognitive, metacognitive, motivational, and emotional learning processes (Panadero, 2017). Productive self-regulated learners are goal oriented, and they actively oversee their learning processes as they advance toward their learning goals. These learners have been found to outperform their less self-regulated peers in different learning tasks and subjects (Uzir et al., 2020). Moreover, given the dynamically changing requirements in the modern job market, SRL skills have been considered critical for future professionals to successfully adapt to those changes, maintain employment relevance, and facilitate future career transfers (Siadaty et al., 2016). However, despite the well-documented benefits of SRL and the many opportunities provided to students to acquire, practise, and refine SRL skills in different subjects, students often struggle to productively self-regulate their learning (Järvelä et al., 2023). Consequently, more research is needed to help students further improve their proficiency in SRL (Lee et al., 2019). Modern technology-enhanced learning environments (TEs) can be used to advance research and support for SRL, in particular given the fine-grained trace data these environments can gather and the interactive learning support they can provide to learners as they work on different learning tasks using the environment. To date, researchers have developed various TEs that support learners' use and development of SRL. For example, MetaTutor uses four simulated pedagogical agents to nudge learners to adopt effective learning strategies, including summarization, monitoring, and activation of prior knowledge (Azevedo et al., 2011, 2022).

Scaffolding, i.e., external guidance provided to students on a skill they could not enact independently, which is then gradually faded out as they become more proficient in that skill, has been considered a common approach to supporting SRL in classroom and digital learning environments (Hmelo-Silver & Azevedo, 2006; Sharma & Hannafin, 2007). Researchers have demonstrated that scaffolding can be used to facilitate learners' SRL skills, leading to improved learning performance and a better understanding of their SRL (Azevedo et al., 2008; Veenman, 2007). Following the promises of scaffolding for supporting SRL, computer-generated SRL scaffolding in digital learning environments has been increasingly investigated over the past decade. Previous research has shown that automated scaffolding of SRL in digital learning environments typically consists of predefined instructions delivered to all students at set time intervals (also known as standardized, fixed, or generalized scaffolds), rather than scaffolds tailored to each student's learning progress and prior SRL activities (Guo, 2022). In addition, most current scaffolding approaches provide clear learning directions or generic instructions, however, without considering students' learning affordances (the affordances of scaffolds) (Devolder et al., 2012) or offering continuous feedback after scaffolds throughout learning (Wisniewski et al., 2020). Thus, tailoring the content and design of scaffolding to provide adaptive support based on students' unique SRL repertoire is considered more effective in prompting effective use of SRL processes (Guo, 2022).

To identify students' SRL processes, researchers have analyzed educational data using machine learning techniques (Saint et al., 2022). The most commonly used educational data are trace data, such as mouse movement and page navigation (Winne et al., 2019). Numerous innovative data analytic methods (Saint et al., 2022), such as process mining (Ahmad Uzir et al., 2019; Matcha et al., 2019) and epistemic network analysis (ENA) (Shaffer et al., 2016), have been used to identify SRL behaviours from the trace data unobtrusively generated by users in TEs. These SRL behaviours have been demonstrated to be predictive of learning outcomes (Fincham et al., 2018; Saint, Gašević, et al., 2020). Despite progress in the field, researchers have encountered two major challenges (Matcha, Gašević, et al., 2020). Firstly, students' learning activities are tightly connected to the structure of the learning environment, instead of being anchored in well-theorized SRL processes (Matcha, Gašević, et al., 2020). As a result, changes in the learning platform make it challenging to compare the results of various studies and to consolidate the SRL research knowledge base (Fan et al., 2021). Secondly, current SRL models are not sufficiently supported by empirical evidence about contextual factors that may affect the selection and execution of SRL strategies (Winne, 2018). Additionally, there is limited research exploring the relationship between SRL tactics inferred from trace data and other factors such as prior knowledge and metacognitive knowledge (Azevedo & Gašević, 2019). Finally, most current TEL systems lack the ability to detect students' SRL processes during learning. The presentation of scaffolds at fixed time intervals is insufficient to support students at varying levels of achievement (Molenaar et al., 2012). While some systems, such as MetaTutor (Azevedo et al., 2022) and Betty's Brain (Munshi et al., 2023), offer adaptive scaffolds, they are not personalized and do not take into account each student's individual SRL progress.

To address the existing challenges related to SRL measurement and facilitation, we developed FLoRA (short for facilitating learners' own regulation activities), a web-based, micro-service architectural learning and data analyzing engine. The design of the FLoRA engine is grounded in the research on SRL conducted over the past few decades. This includes theoretical models of SRL (Panadero, 2017) and empirical studies that examined SRL support using human and digital tutoring (Chi, 2021),

artificial intelligence (AI) (Biswas et al., 2016), and the principles of multimedia learning (Lajoie, 2021). The most pronounced advantage of the FLoRA engine is that it offers opportunities to use learning analytics (LA) (Gašević et al., 2015) and related data analytic techniques (e.g., process mining) (Matcha et al., 2019) to model and understand learners' use of SRL processes, which informed the design and implementation of adaptive supports (e.g., scaffolding) on SRL. Specifically, the FLoRA engine includes (1) instrumentation tools that facilitate work on a task (e.g., tools for text annotation and writing) and unobtrusively collect fine-grained trace data as learners work on a task; (2) a trace parser that automatically analyzes learner trace data and, based on the predefined set of rules, detects SRL processes that learners enact during the task; and (3) a scaffolding module that generates adaptive and personalized scaffolds to guide learners toward more productive SRL and improve their performance in the task, all of which are theory informed.

2. Background

2.1 SRL

Since Zimmerman (1986) clarified the potential of SRL, a variety of learning models have been developed and evolved rapidly to form a comprehensive research ecosystem. Examples are the COPEs (Winne & Hadwin, 1998) and Dual Processing (Boekaerts & Cascallar, 2006) models. While many current SRL models provide unique insights, they all start from the assumption that SRL is best understood as a cycle with distinct phases that each have their own unique set of activities (Panadero, 2017). Comparison of all the extant SRL models reveals that there are, in fact, three distinct phases to the SRL process: preparatory, performance, and appraisal (Panadero, 2017). In terms of the preparatory phase, the main activities are analyzing the task, planning the learning, and setting goals (Panadero, 2017). Analyzing the task activates prior knowledge and determines the most effective cognitive methods for completing the task. During this process, the goals can be formulated, together with the plan for achieving these goals (Winne, 2018). The actual task is done while monitoring and controlling the progress of performance, which refers to the performance phase. The importance of regulating learning through monitoring and controlling procedures is highlighted by Nelson and Narens (1994). They propose dividing cognition into a "meta-level" (the mental representation of a human learner's cognition) and an "object-level" (human learner's cognition), with the interaction between these two levels reflected by the monitoring and controlling of the learning processes. Controlling procedures, such as rereading the material or terminating the current approach, might be modified based on the meta-level representation (e.g., judgments of their learning) gleaned via monitoring (Bjork et al., 2013). In the last phase—appraisal—students reflect on and regulate their current learning, such as contrasting current progress with the goals set at the first phase (Zimmerman, 2000), leading to adaptation for better performance on the next iteration.

2.2 Detecting and Measuring SRL

According to previous studies (Panadero, 2017), SRL encompasses a spectrum of cognitive and metacognitive activities that learners engage in to control their learning processes. These activities include setting goals, monitoring progress, and adjusting strategies as needed. To comprehensively understand these SRL processes in the computer-based learning environment, log data and LA have been used (Siadaty et al., 2016). For example, a learner's decision to attempt a quiz before moving on to reading materials can be used as a criterion for evaluation (Saint, Whitelock-Wainwright, et al., 2020). The identification of a specific SRL process requires locating the corresponding log entries that represent SRL actions. However, linking specific SRL actions to log data traces is challenging, as not all SRL actions are easily recognized in logs (Papamitsiou & Economides, 2014; Bannert et al., 2014). Furthermore, the methods for recognizing metacognitive-level SRL actions are limited (Clarebout et al., 2013). To address this issue, instrumentation tools embedded within learning environments have been proposed to capture SRL processes that otherwise remain hidden, such as organization and orientation (Marzouk et al., 2016). As learners interact with these instrumentation tools, trace logs that indicate SRL processes are generated. For instance, when learners highlight texts, it may indicate that they are engaging in the process of organizing information. An example of a learning technology with such instrumentation tools is nStudy, proposed by Winne and colleagues (Winne, 2017; Winne et al., 2019). The term "instrumentation tool" is used because the designed tools (e.g., writing or timer tools) not only provide convenience for the reading and writing process but also enable the measurement and collection of learning activities that reflect students' SRL processes. These tools offer students the opportunity to initiate and reflect on their cognitive and metacognitive activities. Meanwhile, the tools provide educators and researchers a means to collect and analyze trace data, which are indicative of students' cognitive and metacognitive engagement. Winne (2018) emphasized the importance of such traces, arguing that they should be a standard feature to understand learners' SRL.

SRL can be conceptualized as a series of events, highlighting three successively more complex levels—occurrence, contingency, and patterned contingency (Winne et al., 2019). The first level—occurrence—describes the appearance or existence of one specific learning action (e.g., taking notes). This level captures the frequency of learning actions, but it fails to account for the context of the enacted learning action. The second level—contingency—captures a short sequence of learning actions, which adds the context of a single learning action (e.g., taking notes followed by reading, or taking notes followed

by writing). The third level—patterned contingency—is a more complex form of contingency, which adds extra dimensions about the learning pattern (normally involving more than two learning actions) and frequency of the pattern. For example, “it is revealed that a learner takes notes immediately followed by highlighting, and then paraphrases the highlighted information” is a typical example of capturing SRL at the level of patterned contingency. The patterned contingency level is useful not only in understanding the occurrence of learning actions and the flow of learning actions from one to another but also in explaining learners’ SRL patterns. As suggested by Reimann (2009) and Winne (2018), the analysis of SRL within computer-based learning environments can benefit from a trace parser that offers both event-centred and process-based perspectives. As such, there is a need for a learning environment and learning trace data that can unpack the temporal patterns of engagement on learning tasks (Winne & Perry, 2000; Winne et al., 2019). Many recent studies have started using temporal data to recognize learners’ SRL processes (Paans et al., 2019; Cerezo et al., 2020; Huang & Lajoie, 2021; Saint et al., 2022). Recent literature has concluded that there are five major strands using temporally focused analytical methods to reveal learners’ dynamic SRL processes in various learning platforms (Saint et al., 2022). For instance, one of the strands emphasized the use of trace data collected from learning management systems (LMSs) and massive open online courses (MOOCs) and used, for example, the first-order Markov model (FOMM) and process mining to model learners’ tactic and strategy use (Matcha, Gašević, et al., 2020; Saint, Whitelock-Wainwright, et al., 2020).

To further enhance the utility of trace data, a trace parser is developed within the FLoRA engine. This trace parser is designed to map raw trace data (e.g., navigational logs, mouse and keyboard traces) generated by learners into distinct SRL processes (Fan, van der Graaf, et al., 2022). In this way, we ensure real-time measurement of SRL processes. The development of the SRL trace parser in the FLoRA engine was informed by the theoretical framework of SRL processes proposed by Bannert (2007). The Bannert framework originated from Zimmerman’s model. It is an operationalization of existing theoretical work on SRL in the form of a coding scheme for analysis of SRL data. Bannert’s team used this SRL framework and think-aloud coding protocol to research higher education learning tasks (Bannert, 2007, 2009; Bannert & Mengelkamp, 2013; Bannert et al., 2014). According to this framework, SRL processes can be sorted into three broad categories: metacognitive, cognitive, and other processes. The metacognitive processes include *task orientation* (i.e., surveying task requirements and constraints), *planning* (i.e., arranging learning activities for the task), *monitoring* (i.e., overseeing the effectiveness of learning throughout the task), and *evaluation of learning*. For example, a learner orients to understand task instructions and resources available for the task; plans to come up with learning strategies for the task; monitors their progress relative to task requirements and a scoring rubric; and evaluates whether the previously used learning strategies worked well, given the progress achieved. The cognitive category spans several low- and high-cognition processes. As articulated in the Bannert (2007) theoretical framework, the low-cognition processes are *first-time reading* and *rereading*, whereas high-cognition processes are *elaboration* and *organization*, e.g., summarizing or synthesizing the information from source texts in a writing task. Other processes, although not included in the real-time analysis, contain learner motivation (e.g., a learner’s excitement for the topic) and processes enacted due to procedural issues (e.g., a learner asking questions about using the writing tool). Each component of Bannert’s framework can be mapped to the preparation, performance, and appraisal stages, specifically (1) preparation: orientation and planning; (2) performance: low/high cognition, monitoring, and evaluation; and (3) appraisal: evaluation. It is noteworthy that evaluation can occur in both the performance stage and the appraisal stage. More details about the SRL trace parser are provided by Fan, Lim, and colleagues (2022).

2.3 Examining Analogous Systems

Numerous studies have led to the design and development of various tools aimed at supporting SRL from diverse perspectives. However, the majority of these tools predominantly focus on one or two specific phases of SRL (Alvarez et al., 2022). For instance, most tools are oriented toward activities such as planning (Nussbaumer et al., 2014), goal-setting (Alexiou & Paraskeva, 2015; Davis et al., 2018), time management (Yau & Joy, 2009), annotation (Yousef et al., 2015), and LA dashboards (Matcha, Uzir, et al., 2020). Given that SRL involves cyclical processes where different phases can recur multiple times throughout the learning period (Zimmerman, 2013), tools that concentrate solely on a single phase are insufficient for fully fostering and promoting SRL. Therefore, a comprehensive system encompassing all SRL theory-based instrumentation tools is essential to detect and measure all phases of SRL effectively.

Currently, there are only a few comprehensive systems designed to support SRL, including nStudy (Winne et al., 2019), MetaTutor (Azevedo et al., 2022), and Betty’s Brain (Munshi et al., 2023). nStudy is a comprehensive web-based learning tool that supports SRL by providing students with a variety of tools to enhance study strategies, metacognition, and learning progress tracking. It features functionalities such as searching, note-taking, writing, and information organization through concept maps (Winne et al., 2019). MetaTutor is an intelligent tutoring system that supports SRL by guiding learners through complex scientific content using adaptive feedback and scaffolding. It integrates cognitive, metacognitive, and motivational strategies to help learners develop effective learning habits, offering real-time prompts, hints, and progress tracking to facilitate learning through reflection and self-monitoring (Azevedo et al., 2011). Betty’s Brain is an educational software tool designed to teach students about complex systems and scientific concepts through a learning-by-teaching approach. In this interactive

environment, students assist a virtual character named Betty by creating and modifying concept maps that visually represent relationships between different concepts, thereby engaging in SRL activities such as planning, monitoring, and evaluating their understanding (Munshi et al., 2023).

Despite the careful design of these systems to support SRL and promote learning, they exhibit limitations in processing trace data and supporting SRL. Although these systems can capture a range of fine-grained digital raw log data (e.g., clicks, highlights), they do not have well-developed and validated functionalities for mapping these raw log data onto SRL processes. As a result, while they can perform data analysis after the learning task is completed, the scaffolding they provide is not grounded in learners' real-time SRL processes at different stages of learning. Furthermore, these systems are primarily designed for research purposes and have been used predominantly in laboratory settings. This focus raises concerns about their applicability and effectiveness in authentic learning environments, such as traditional classrooms and online learning platforms, where variables are more dynamic and less controlled. The complexities and variabilities inherent in real-world educational settings may significantly impact the systems' ability to support SRL effectively, indicating a need for additional investigation and adaptation. The FLoRA engine has been specifically designed and developed to address these limitations.

3. FLoRA Architecture

FLoRA is an interactive, AI-powered engine, incorporating an instrumentation tools module, a trace parser module, and a scaffolding module, designed to support research and improve SRL skills in complex learning tasks that involve searching, navigating, reading, and interacting with textual or video sources, and developing written responses grounded in those sources. Unlike other learning systems or platforms, which can provide learning materials and conduct learning activities, FLoRA is independent of LMSs, and all three modules are used together with LMSs. Hence, FLoRA can be seamlessly integrated into various web-based learning platforms, such as Moodle. Figure 1 illustrates the integration of the FLoRA engine within the Moodle platform, providing an overview of its appearance. Three main working zones are pointed out in the picture: the Navigation zone, the Reading zone, and the Instrumentation tools zone. Only the Instrumentation tools zone is provided by the FLoRA engine. The other parts of the system are original Moodle components. Because the FLoRA-driven instrumentation tools are injected into the LMS, all of the learners' actions happening on the web pages can be tracked and processed in the back end of the FLoRA engine.

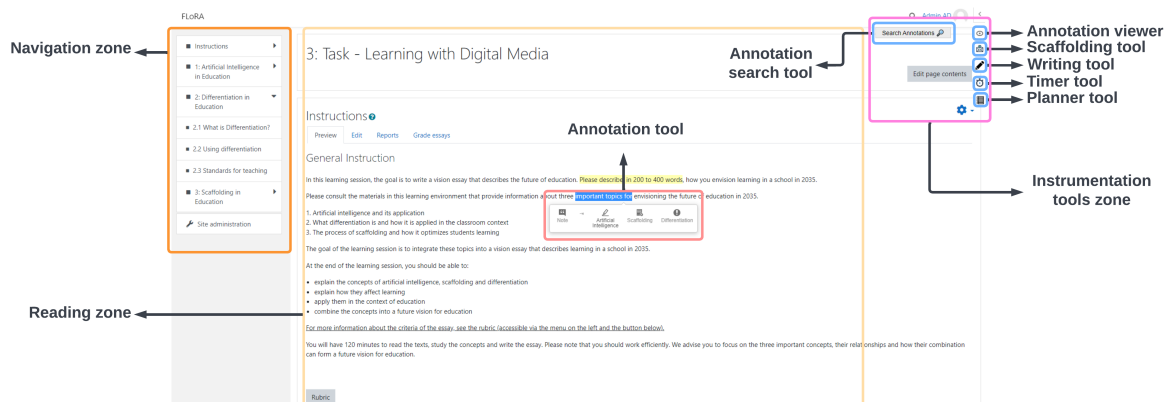


Figure 1. The user interface of FLoRA integrated into the Moodle LMS.

A high-level architecture of the FLoRA engine can be found in Figure 2. The instrumentation tools are designed to be easily integrated into the different LMSs, ensuring a smooth and uninterrupted experience for learners. While students use the learning system and the instrumentation tools, trace data are collected by the engine and systematically analyzed by the trace parser to identify SRL events. As shown in Figure 2, the SRL processes are labelled as metacognition level and cognition level. The details about each code are further explained in Section 3.2. Subsequently, the scaffolding module is used to generate personalized scaffolds. The scaffolds are displayed via the scaffolding display tool and are instrumental in augmenting students' learning performance and fostering productive SRL behaviours, as discussed by Lim and colleagues (2023). This section aims to describe the three primary modules constituting the FLoRA engine, elucidating their individual and collective contributions to the facilitation of advanced learning processes.

3.1 Module I: Instrumentation Tools

As described in Section 2, there remains a significant gap in understanding the specific SRL processes from trace logs. This gap is particularly noteworthy given that instrumentation tools are specifically designed to capture these elusive SRL processes. The

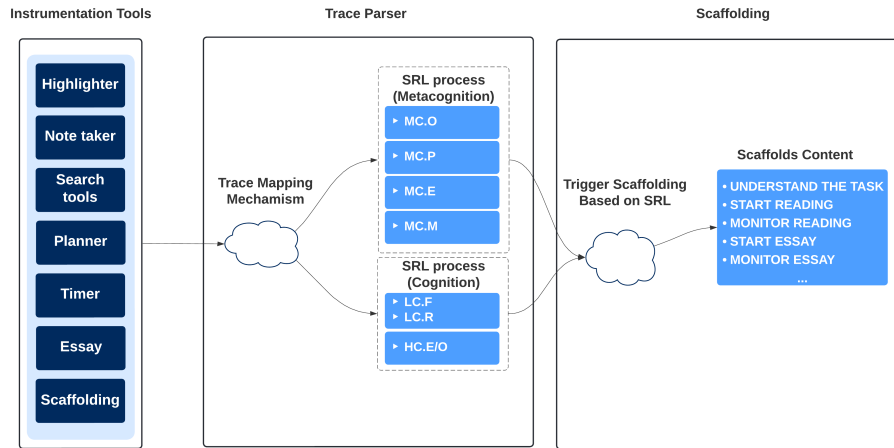


Figure 2. The architecture of the FLoRA engine.

FLoRA engine exemplifies this by offering a suite of instrumentation tools designed to facilitate the detection and measurement of SRL processes. These tools, which include an annotation tool (highlighter and notetaker), an annotation search tool, a timer, a planner, a writing tool, and a scaffolding tool, capture high-quality, granular, and temporally ordered log data, enabling a more nuanced analysis of SRL dynamics.

3.1.1 Annotation Tool

The instrumentation tools feature an annotation tool including a highlighter and a side panel for note taking, shown in Figure 3. When text is selected on a web page, a pane emerges, presenting an array of labelling options for the user’s selection. Post-selection, the annotation is generated in the right-side panel, where users can append notes or assign customized labels. The default label mirrors the choice made during the highlighting phase. Once annotations are completed, users have the option to publish them and, if desired, locate them via the search tools. Additionally, the tool offers the functionality of hiding or displaying all highlights, thereby granting users agency over the annotations’ influence on their reading experience. This tool aims to track students’ highlighting and note-taking behaviours. Previous research has established a positive relationship between the highlighting action and learning performance (van der Graaf et al., 2021), and thus FLoRA leverages this approach to shed light on students’ SRL strategies. In addition to the highlighter tool, a note-taking tool is also provided in the annotation tool. While learners use the highlighter, the highlighted texts are linked with notes, allowing learners to write down extra information. During this process, learners will generate long-term memory for the encoded information from the annotation and save additional information for future checking (Azevedo & Cromley, 2004; Di Vesta & Gray, 1973). This indicates that notes can be used to memorize information (low cognition), organize or elaborate information (high cognition), and monitor the relevance of information (metacognition) (Lim et al., 2021). Beyond the conventional note-taking process, FLoRA also records all keystroke events during note-writing and note-browsing events, providing more comprehensive trace data to help researchers gain a deeper understanding of the SRL processes.

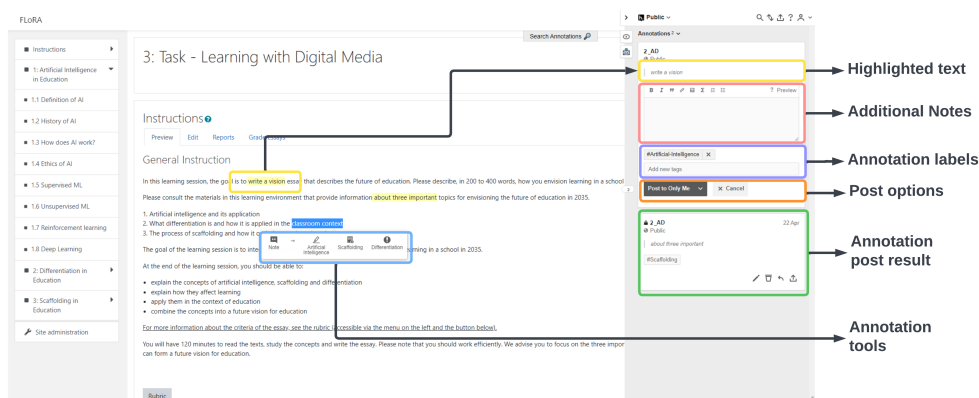


Figure 3. Annotation tool.

3.1.2 Annotation Search Tool

Figure 4 displays the annotation search tool. The interface displays a list of annotations and a text-input box. Users can view public annotations in the search results area or employ the text-input box for fuzzy searches, with filtered outcomes presented. Furthermore, the tool enables users to locate highlighted texts tagged with specific labels, offering a streamlined method to track annotations across various web pages in the learning environment. The use of a search tool is indicative of planning behaviour, as the initiation of a search suggests that the learner has a specific goal in mind and has chosen keywords to direct their search (Greene & Azevedo, 2009). Furthermore, the process of comparing search results aligns with the monitoring phase of SRL, wherein learners critically evaluate the pertinence of the information they have retrieved (Greene & Azevedo, 2009). While the connections between search actions and SRL processes remain an area not yet exhaustively investigated, FLoRA implements a sophisticated annotation search tool. The FLoRA engine tracks a comprehensive array of events associated with the use of the search tool. This includes but is not limited to the opening and closing of the tool, and the specific content of these keywords. Such detailed tracking is instrumental in providing insight into learners' search strategies and behaviours. In our previous studies, the use of an annotation search tool connects with the SRL process *planning and monitoring*. Besides, the frequencies of timer, search, and planner usage were positively correlated, as they share the SRL processes they captured. But the search function is used least often compared with other tools. An explanation for the low usage of the search tool is that the learning materials in our studies were limited to three topics, which might have reduced the need for goal-directed search behaviour (van der Graaf et al., 2021).

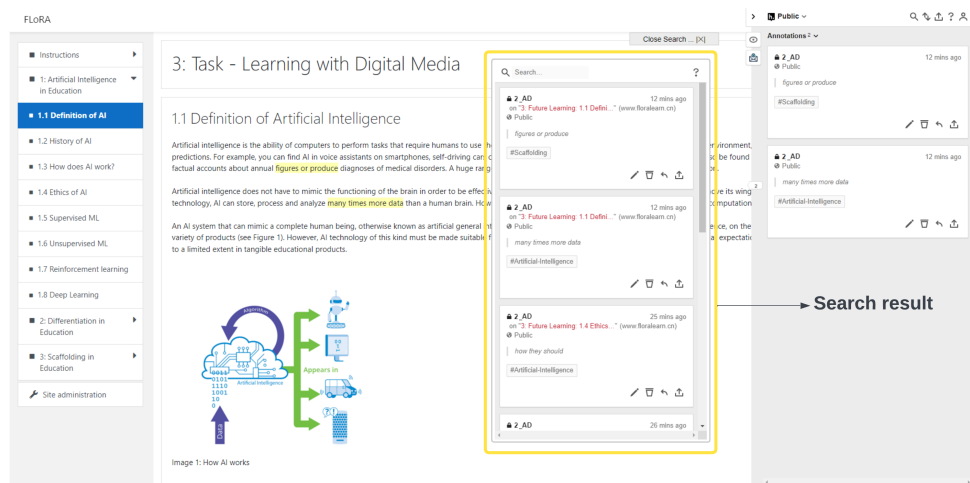


Figure 4. Annotation search tool.

3.1.3 Timer Tool

As shown in Figure 5, the timer tool, triggered by an alarm clock icon, displays a countdown timer with adjustable duration via the configuration file. To mitigate the issue of constant timer monitoring by students, the tool automatically closes after a 2-second display. This feature is crucial in facilitating the detection of student interaction with the timer. The tool remains accessible throughout the task, enabling students to monitor the time remaining for the completion of a given task.

Based on the theoretical model, the checking timer is identified as a metacognitive process: *monitoring* (Bannert, 2007). This feature is instrumental in enabling students to track their learning progress in alignment with their predefined objectives (Winne, 2017). Moreover, the systematic tracking of all actions associated with the use of the timer, including but not limited to the opening and closing of the tool, is essential for capturing a comprehensive and accurate representation of students' SRL behaviours. This level of detailed monitoring is crucial for providing insight into the temporal aspects of students' learning strategies, thereby contributing to the understanding and enhancement of SRL processes in educational settings. In our previous studies, the timer was used frequently in a task and it was used by all students. Since the timer can be used to assess monitoring of time remaining, it can help evaluate progress in relation to the set goals (van der Graaf et al., 2021).

3.1.4 Planner Tool

Given that planning is a principal facet of metacognition (Winne et al., 2019), the FLoRA engine incorporates a planner tool (as shown in Figure 6) designed to foster engagement in planning and goal-setting activities among learners. The significance of such activities is extensively documented in the literature, highlighting their positive impact on motivation and learning

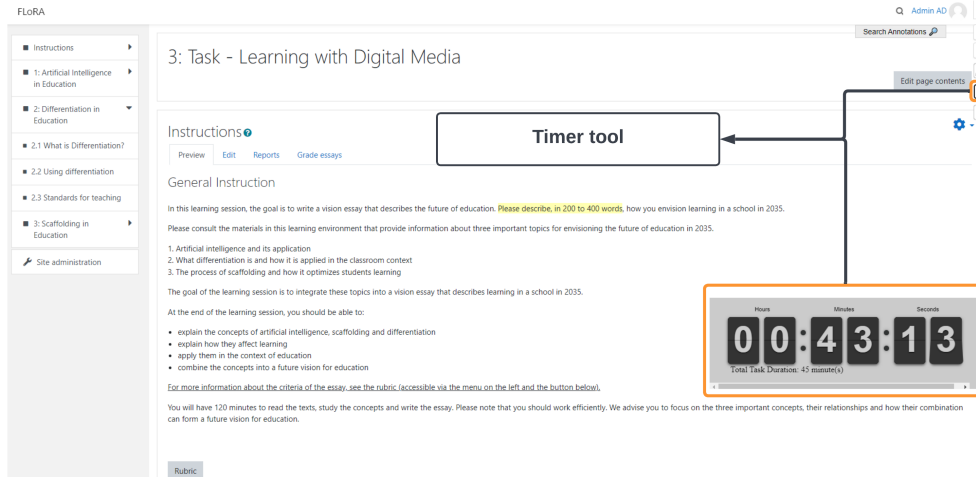


Figure 5. Timer tool.

performance (Bowman et al., 2020). This tool enables learners to methodically organize their tasks in a timeline format, offering the flexibility to adjust the allocated time for each specific task.

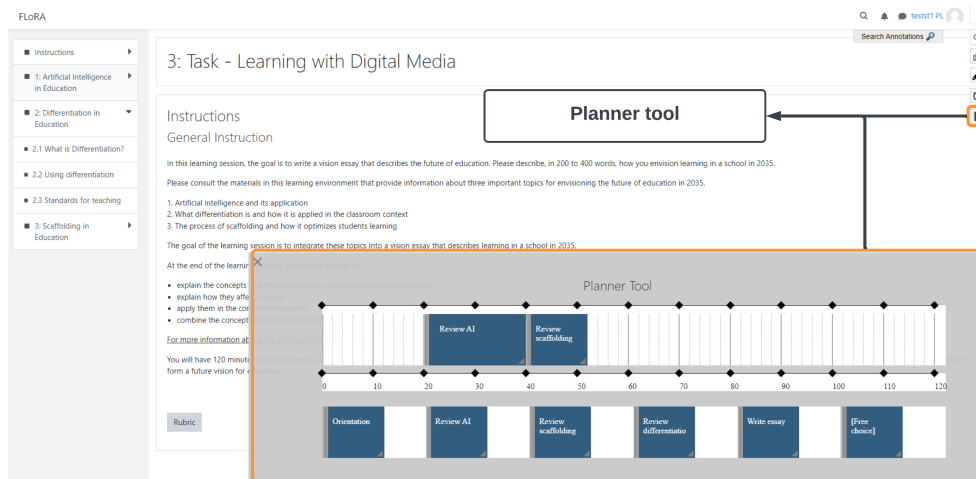


Figure 6. Planner tool.

This tool is engineered to offer a drag-and-drop functionality, thereby significantly reducing the need for manual text or time entry for scheduling tasks. Users can effortlessly relocate predefined plan items to specific time slots, with the added functionality of adjusting the duration by resizing the item block. This intuitive design enhances users' experience by simplifying the task-planning process. Furthermore, FLoRA ensures a comprehensive record of all interaction events associated with the planner tool. This includes but is not limited to opening and closing the planner, as well as detailed tracking of mouse clicks and movements. Such extensive data collection is crucial for analyzing learners' planning behaviours, providing valuable insights into their metacognitive processes. This, in turn, contributes to a deeper understanding of the role of planning in enhancing learning outcomes. The planner was hardly used in our study. An explanation for the low usage of the planner is that the learning task in our study was 45 minutes long, which would reduce the need for a planner (van der Graaf et al., 2021). In another study we conducted in a non-English speaking country that had a task time of 2 hours, planner usage increased significantly (Fan et al., 2023).

3.1.5 Writing Tool

The FLoRA engine is designed to facilitate a diverse array of learning activities, using textual or video-based materials. In prior research endeavours (Fan, Lim, et al., 2022), a heightened focus was placed on the reading and writing task. According to the theoretical framework of SRL processes proposed by Bannert (2007), the actions related to the writing tool, such as the writing text and open tools, can be identified as a metacognitive process—*evaluation*—and a cognition process—*elaboration/organization*. To capture learners' SRL processes during a task, a writing tool was developed and integrated into

the engine, as displayed in Figure 7. This tool features an automatic recording mechanism that captures and stores keystroke events and the corresponding essay content in a database, with each keystroke being automatically logged. This feature is particularly beneficial as it ensures the preservation of the essay text, safeguarding against potential losses due to unforeseen closures of the writing tool. Additionally, the events of opening and closing the writing tool are tracked to further contribute to the comprehensive data collection.

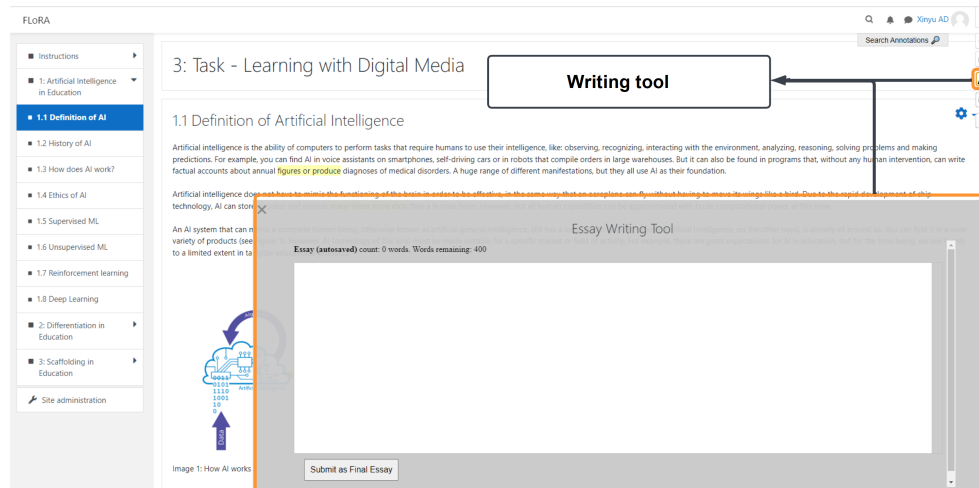


Figure 7. Writing tool.

3.1.6 Scaffolding Display Tool

The scaffolding display tool, as depicted in Figure 8, is designed to activate upon the fulfillment of predefined conditions, detailed in Table 1 and Section 3.3. This tool is presented as a pop-up window set against a shadowed backdrop, providing learners with the opportunity to select various options for the creation of a to-do list. Significantly, the tool's design allows researchers to customize both the timing for the scaffold's activation and its content through modifications in the back-end configuration file. This adaptability ensures that the tool can be tailored to meet specific research requirements and learning contexts.

The scaffolding tool is intricately connected to the scaffolding module. The scaffolding module determines the appropriate scaffold content to present to students based on the detection of various SRL processes. Moreover, all interactions with this tool are captured and recorded. Given the recognized deficiency in SRL proficiency among many learners, this scaffolding display tool, in conjunction with the scaffolding module, has been specifically developed to provide more effective and targeted assistance, tailored to the needs of individual learners. A comprehensive elaboration of the back-end scaffolding module, including its design logic and applications, is presented in Section 3.3.

3.2 Module II: Trace Parser

Within the FLORA trace parser, a sophisticated protocol for the measurement of SRL processes has been established using trace data. This protocol is underpinned by prior scholarly research (Fan, Lim, et al., 2022), and a theoretical framework (Bannert, 2007) has been rigorously developed to delineate specific SRL processes. To be more specific, the learning process is categorized into three major domains: metacognition (MC), low cognition (LC), and high cognition (HC). Each cognitive level encompasses distinct SRL processes, such as *MC.Orientation*, *MC.Planning*, *MC.Monitoring*, *MC.Evaluation*, *LC.First-reading*, *LC.Re-reading*, and *HC.Elaboration/organization*. These processes are elaborated with detailed definitions in Appendix A. Moreover, the protocol encompasses an action library and a process library, which are instrumental in translating raw trace data into meaningful learning actions and their corresponding SRL processes (a comprehensive library is available in Appendix A).

The action library defines a total of 17 learning actions, which correlate with specific trace data, such as *RELEVANT_READING*, *WRITE_essay*, and *EDIT_ANNOTATION*. For example, the *OPEN_essay* label denotes the action of opening the writing tool. When students start typing in the writing tool, all keystroke events are captured and mapped to the *WRITE_essay* action. Further information on the complete action definitions and the labelling procedure is provided in Appendix A.

The process library employs the SRL process definitions from Appendix A as a framework to translate various actions (as described in the paragraph above) into SRL events (as depicted in Figure 9). For example, when a learner highlights text or edits a note while surfing the general instruction page of the task, the actions are labelled as *GENERAL_INSTRUCTION* and *EDIT_ANNOTATION*, respectively. The combination of these actions forms a pattern "*GENERAL_INSTRUCTION* <

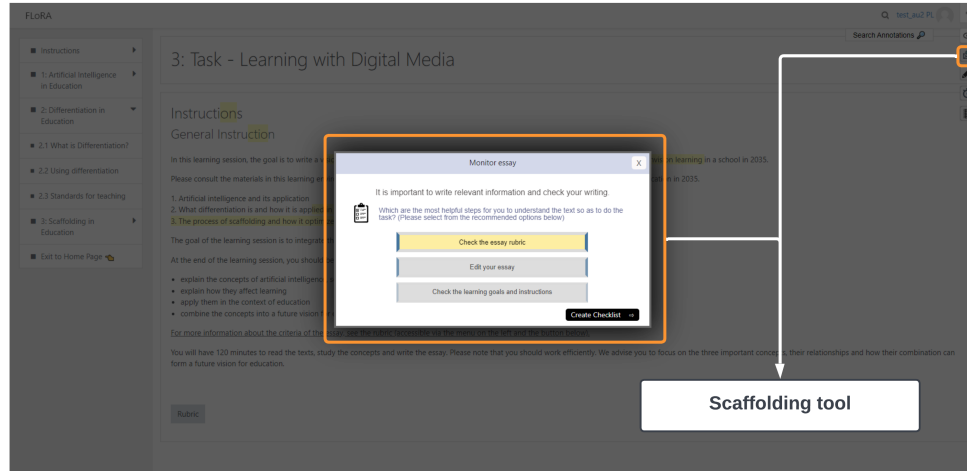


Figure 8. Scaffolding display tool.

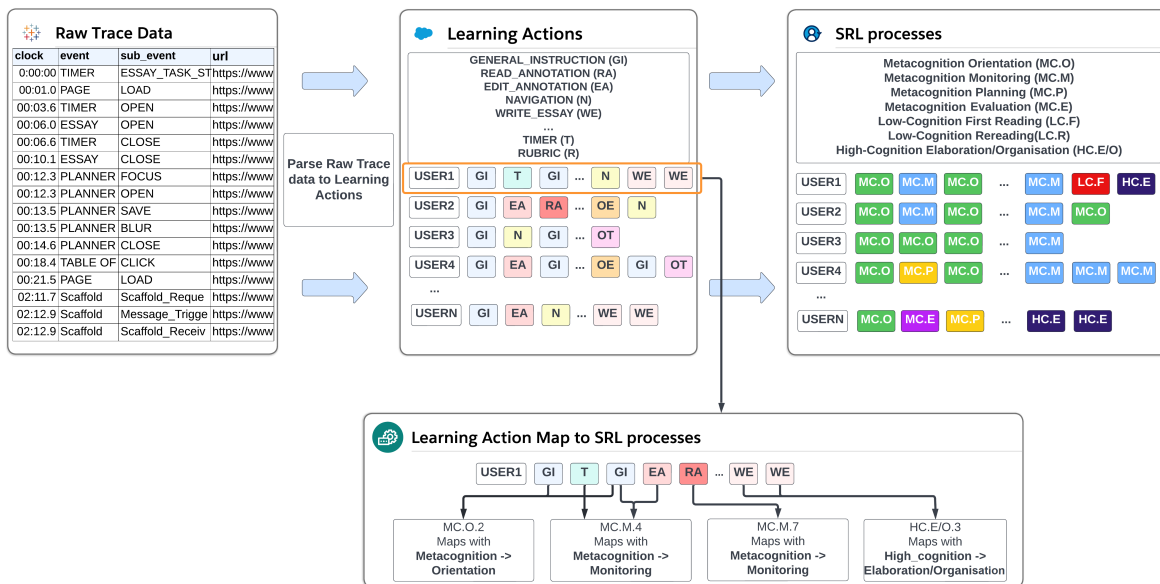


Figure 9. Parsing trace data to SRL process.

- > EDIT_ANNOTATION,” which, according to the process library, is mapped to *MC.Orientation*. The symbol “< - >” indicates that there is no sequential dependency between the actions GENERAL_INSTRUCTION and EDIT_ANNOTATION, allowing for flexibility in the order of these actions. In another scenario, when a learner engages with the timer to check the time, this action is labelled as TIMER and categorized under *MC.Monitoring*, signifying the learner’s effort to monitor the remaining time for the task. In the analysis of trace data, all items labelled as NO_PROCESS are excluded, as they do not correspond to any meaningful sequences in the process library. Finally, if a learner does not interact with the FLoRA tools or learning systems for a certain period of time (5 minutes defined in FLoRA; this time is decided based on our previous studies (Fan, van der Graaf, et al., 2022) and can be changed in the config file), an action named OFF_TASK is recorded. In summary, through the integration of the action and process libraries, all trace logs are systematically transformed into labels indicative of the SRL processes.

To validate the protocol, both trace data and think-aloud data are collected and used. Two previous studies have been conducted to collect the think-aloud data. The first one is from 44 participants ($M_{age} = 26.56$ years, $SD_{age} = 4.18$ years) across various universities in Germany. The second one was conducted at a university in the Netherlands and involved 44 learners ($M_{age} = 21$ years, $SD_{age} = 3$ years). Participants joined the onsite session individually. The learning environment used is an older version of FLoRA, which also contains an annotation tool, a timer tool, a planner tool, and a writing tool. Throughout the task, participants were required to read and think aloud. Whenever they were silent for more than 5 seconds or spoke quietly,

they were prompted verbally by the experimenter. More details can be found in our prior work (Lim et al., 2021). The trace data were collected via the FLoRA engine.

After collecting all the data, the validation process is achieved through an innovative methodology that effectively merges both theory-driven and data-driven approaches, as proposed in prior research (Fan, van der Graaf, et al., 2022). The theory-driven approach starts with the development of a comprehensive coding scheme based on an established theoretical framework of SRL. This involves labelling raw trace data into meaningful learning actions and categorizing these actions into SRL processes according to Bannert's framework (Bannert, 2007). Multiple rounds of expert discussions are conducted to refine the process library and ensure its alignment with theoretical constructs. On the other hand, the data-driven approach leverages empirical evidence by aligning coded think-aloud data with trace data. This involves segmenting trace data based on think-aloud codes and using process mining techniques to identify dominant action patterns within each segment. These patterns are then interpreted and validated against the think-aloud data to construct a data-driven process library. The final step involves integrating empirical evidence with theoretical rationale to resolve any ambiguities and finalize the interpretation of action patterns, resulting in an improved process library. This integrated methodology ensures a robust and validated measurement of SRL processes by combining the strengths of both theoretical and empirical approaches.

This methodology establishes three pivotal alignments between trace data and think-aloud data. The first alignment contrasted the theory-driven process library with coded think-aloud data, highlighting matches, mismatches, and unique aspects of each data type, thereby validating the SRL processes extracted from trace data. The second alignment, adopting a data-driven approach, refined the process library by directly aligning learning actions with think-aloud data, using process mining techniques for more accurate mapping between SRL processes and think-aloud codes. Finally, the third alignment involved an empirical evaluation of this improved process library against think-aloud data from both training and test sets. This comprehensive approach ensured a deeper and more nuanced understanding of the SRL processes, aligning them more closely with the learners' cognitive and metacognitive strategies as expressed in their think-aloud data, thereby significantly enhancing the validity of the measurement. In addition, the protocol employs a quantitative indicator, termed the match rate, to assess the degree of correspondence between SRL processes ascertained from trace data and those derived from think-aloud data. This match rate is further augmented by three supplementary metrics: sensitivity, specificity, and trace coverage. Each of these metrics plays a vital role in providing a comprehensive and robust evaluation of the protocol's overall validity.

3.3 Module III: Scaffolding

The scaffolding module is a pivotal component in the FLoRA engine. The overall design of scaffolding was informed by SRL theory (Winne & Hadwin, 1998) and supplemented with multimedia learning theory (Clark & Mayer, 2016). Within this module, scaffolds are crafted and implemented to provide assistance as needed. With the advanced instrumentation tools and trace parser, log data are collected and processed to detect learners' SRL processes in real time. The triggering mechanism of these scaffolds is complexly designed to be contingent upon the learner's real-time detected SRL processes (Maier & Klotz, 2022).

Informed by previous empirical studies, the scaffolding module incorporates five distinct scaffolds at the moment. This decision is supported by findings from our previous lab studies, which indicate that, during a 2-hour reflective writing task, the use of five distinct scaffolds—each serving a specific purpose and introduced at designated intervals—maximizes learners' overall performance (van der Graaf et al., 2023). However, it was also noted that an excessive number of scaffolds could potentially impede students' learning progress during their learning activities. Consequently, these findings advocate for a moderated use of scaffolds, specifically the incorporation of five, to optimize educational outcomes (van der Graaf et al., 2022).

In terms of the content, each scaffold has a generalized and personalized version. The generalized scaffolds uniformly present identical messages and options to all students, whereas the personalized scaffolds are tailored to individual learners' needs, based on real-time detection of their SRL processes. Each scaffold has a specific purpose. As shown in Table 1, the first scaffold (Orientation) is aligned with the initial stages of SRL, where learners analyze the task and set their learning goals. This scaffold can help learners understand the task requirements, clarify objectives, and establish clear and achievable goals. The second scaffold (Start Reading) corresponds to the phase where learners select and implement strategies for understanding and assimilating information. This scaffold can provide guidance on effective reading techniques, how to identify key information, and methods for active engagement with the text. The third scaffold (Monitor Reading) aligns with the metacognitive aspect of SRL, specifically monitoring one's understanding and effectiveness during the reading process. It can help learners assess their comprehension, identify areas of confusion, and decide whether to continue with the current reading strategy or modify it. The fourth scaffold (Start Writing) maps to the phase where learners implement strategies to perform the task, in this case, writing. This scaffold can assist in organizing thoughts, structuring the writing, and applying effective writing strategies. The fifth scaffold (Monitor Writing) is aligned with the self-reflection phase of SRL. It supports learners in evaluating their written work, reflecting on the writing process, and making necessary adjustments. It can guide learners in revising their work, ensuring that it meets the objectives, and adjusting strategies for future tasks.

These scaffolds align with the SRL cycle definition (Winne & Hadwin, 1998; Zimmerman, 2000). Each scaffold provides

four options from which learners can choose to create a to-do list that can help them plan, take, and monitor further learning actions. In our design, we recommend that learners read before writing. This is because our previous field study has confirmed that learners with different adoption of learning strategies demonstrated significant differences in medium effect size regarding the transfer test, with significant differences between “read first then write” and “write intensively, read selectively” strategies (Srivastava et al., 2022). As such, with the current FLoRA design used in reflective writing tasks, both our lab study and field study confirmed that “writing after reading” should be promoted. This is not to say that FLoRA can only be used for reflective writing tasks which promote the writing after reading strategy, but rather emphasizing that scaffolding design should be contextualized and empirically evidenced to support effective SRL processes.

As for triggering of scaffolding, the FLoRA engine employs a request/response communication model (Oluwatosin, 2014), where the client dispatches a scaffold request encapsulating client parameters (user identification, scaffolding condition, elapsed time) at predetermined scaffold intervals. Subsequently, the scaffolding module conducts an analysis of trace actions, responding with a scaffold imbued with pertinent message text and options. In existing research, FLoRA has mostly been used in studies that support tasks requiring students to produce essays based on a given set of readings. In that context, the triggering of scaffolding was informed by the work of Fan, van der Graaf, and colleagues (2022). Their study used learners’ think-aloud data to identify when a specific SRL process should be executed, with learners categorized into three groups based on their essay score (poor/average/good). All of the essays are marked by two expert researchers. The scoring passed the reliability test—inter-annotator agreement, measured using Cohen’s kappa coefficient (Cohen, 1960). For example, compared to the average and poor essay groups, the high-performing essay group tended to begin the orientation process earlier and complete it sooner than the other two groups. This finding suggests the importance of prompting orientation processes at the early stages of the task. Subsequently, based on how the good essay group learned compared to the others, in a 45-minute task, the scaffold’s trigger timing is determined: orientation at 2 minutes since the start of the task, reading at 7 minutes, monitoring of reading at 16 minutes, writing at 21 minutes, and monitoring of writing at 35 minutes. For learners for whom the language they are using to work on the task is not their first, the scaffold triggering time can be adjusted accordingly. Table 1 lists the contents of the five scaffolds that have been used in the FLoRA engine for essay writing tasks with a given set of reading materials. As the FLoRA engine is designed to be easily reusable, the scaffold triggering time and the content of each scaffold are adjustable for different studies. Furthermore, we repeatedly collect learners’ opinions (through post-task interviews) regarding the perceived usefulness and adaptivity of scaffolding to understand how the scaffolding content can be further improved (Li et al., 2023).

An example of the scaffold triggering mechanism for both generalized and personalized groups at the 2-minute time point is illustrated in Figure 10. This representation underscores the dependence of the prompting process on three critical factors: the student’s progression (time elapsed during the task); study condition (generalized or personalized); and, in the case of personalized scaffolding, the student’s identity to authenticate individual action labels. According to our scaffolding framework, orientation-related scaffolds are activated upon the student reaching the 2-minute threshold in the writing activity. For students in the generalized condition, the scaffold content for orientation (including message text and all four options) is sourced from the database, formatted as a JSON object, and transmitted to the client browser for display. In the context of personalized scaffolds, the module assesses the learner’s SRL process concerning scaffold 1 in Table 1 before generating the corresponding message and options. For instance, in Figure 10, at the 2-minute mark, a learner is exhibiting adequate engagement (exceeding 15 seconds) with the general instruction page and demonstrating reading behaviours (a sequence categorized as *ORIENTATION - MC.O.1* in the process library, involving “GENERAL_INSTRUCTION – > NAVIGATION – > RELEVANT_READING”), which causes the deactivation of the option to reorient (i.e., “Check the learning goal and instruction”) in the first personalized scaffold. This leaves the learner with only three options. Should all SRL processes be detected, all options within the scaffold are deactivated, leading to the omission of the entire scaffold. This contrasts with generalized scaffolds, where users receive all scaffolds complete with all options, regardless of their divergent learning behaviours.

4. Implementation

This section explores the practical implementation of the FLoRA engine across diverse learning activities, using advanced and scalable technologies. Developed in accordance with the architecture outlined in the previous section, the FLoRA engine is available for download on GitHub (<https://github.com/Xinyu-Li/flora-lms>) under the BSD-2-Clause licence. As open-source software, FLoRA aims to significantly contribute to the advancement of SRL research. The subsequent sections clarify several key aspects of FLoRA’s deployment and technological choices, specifically (1) the FLoRA engine, (2) the log server, (3) the infrastructure, and (4) case studies.

4.1 FLoRA Engine

The FLoRA engine is designed for integration into various LMSs, such as Moodle. In alignment with this integration, its interface has been crafted to resonate with the thematic elements of the LMSs and can be easily adjusted for different LMSs. The primary interface elements presented to students are the instrumentation tools as well as the LMS interface (i.e., web pages).

Table 1. Example of five scaffolds that are currently implemented in FLoRA: content of suggested actions and examples of times when scaffolds are triggered in some of the previous studies where tasks were 45 minutes long and tasks focused on essay writing based on a given set of readings.

Targeted perspective and timing	Prompt message	Learning suggestions
Scaffold 1: Orientation (2nd minute)	Accurate understanding of the content and requirements of the literacy task is critical. Based on your learning behaviour so far, we recommend the following steps:	(a) Use the table of contents to get an overview and skim text. (b) Check the essay rubric carefully. (c) Make sure you understand the learning goals and instructions. (d) Process information by taking notes.
Scaffold 2: Start reading (7th minute)	Efficient and high-quality reading of information on different topics in the material is essential. Based on your learning behaviour so far, we recommend the following steps:	(a) Note down important information. (b) Select what to read next using the table of contents. (c) Check time and monitor your reading progress. (d) Search for (specific) information.
Scaffold 3: Monitor reading (16th minute)	Making sure you only read task-related pages and thinking about the relationship between reading and writing is the key to learning success. Based on your learning behaviour so far, we recommend the following steps:	(a) Review annotations and check what you have learned so far. (b) Review the learning goals and instructions to focus on relevant content. (c) Check your essay structure to determine what to read next. (d) Check the essay rubric.
Scaffold 4: Start writing (21st minute)	Starting writing early and conscientiously writing a high-quality essay is central to the success of this task. Based on your learning behaviour so far, we recommend the following steps:	(a) Check the remaining time for writing. (b) Check and reread the essay rubric. (c) Draft the essay by using your own language and transferring learning to main points. (d) Write the essay with help from notes.
Scaffold 5: Monitor writing (35th minute)	At the end of the task, complete the essay according to the task instructions and rubric to get a high score. Based on your learning behaviour so far, we recommend the following steps:	(a) Check the essay rubric to revise your essay. (b) Edit your essay to make sure it's complete. (c) Check the learning goals and instructions to avoid digressions. (d) Check the timer to manage your time.

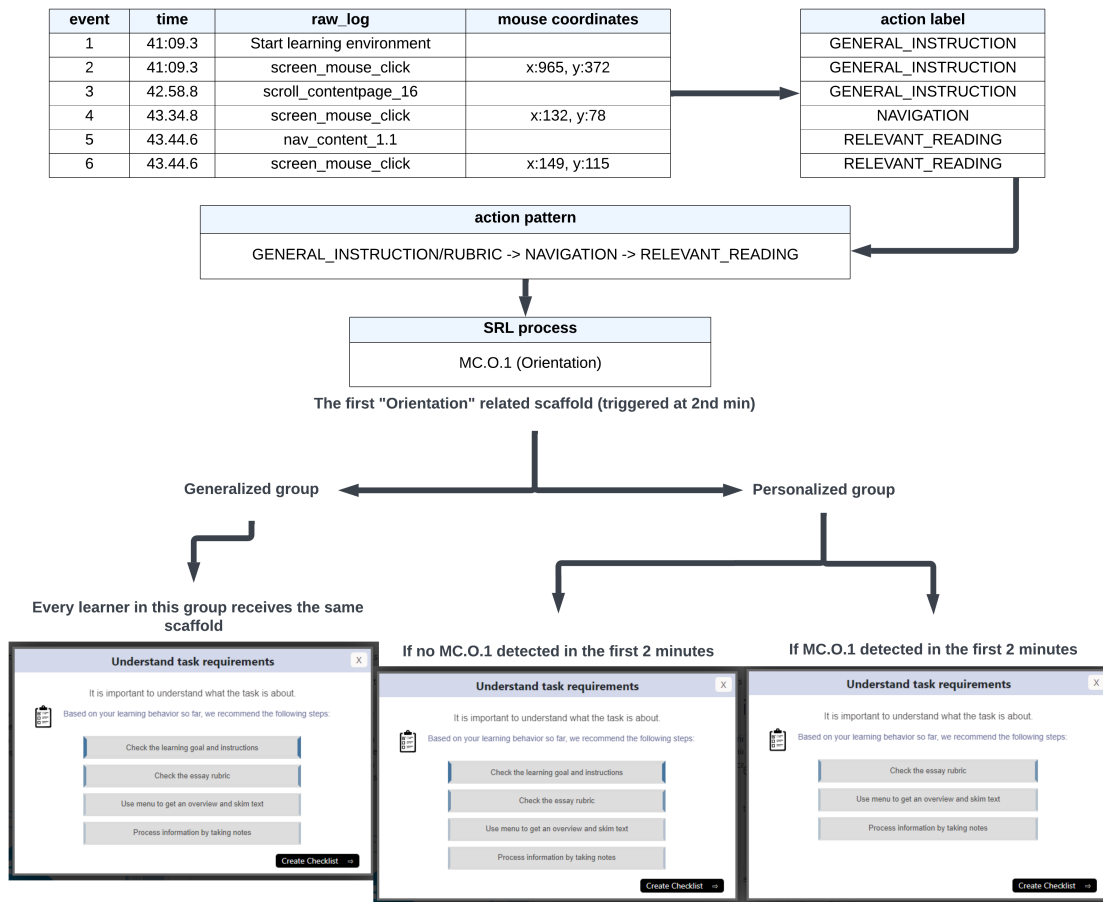


Figure 10. Data parsing example and one scaffold rule.

These tools are predominantly developed using JavaScript, supplemented with PHP programming, ensuring a seamless and efficient user experience. In this integrated environment, both learners and educators engage not only with the instrumentation tools but also with the broader LMSs' interface. This strategic design decision significantly reduces the learning curve for users, as LMSs such as Moodle are extensively used across educational institutions globally. Within this ecosystem, the LMSs serve as repositories and management systems for educational content created by teachers. As learners engage with these materials, the FLoRA instrumentation tools unobtrusively record all interaction logs to a back-end database.

The technological choices made in the development of FLoRA underscore its commitment to leveraging robust, widely recognized open-source tools to enhance functionality and scalability. By integrating the Hypothesis¹ JavaScript library, FLoRA benefits from a well-established platform for collaborative annotation and interaction within learning materials. The annotation search tool employs Elasticsearch (Gormley & Tong, 2015) as its back-end service provider, offering powerful, real-time search and analytics capabilities that are essential for handling large datasets and ensuring quick, accurate information retrieval. The timer tool is built using the open-source FlipDown.js² library, which provides a visually appealing and highly customizable countdown timer. Additionally, the planner tool is developed with the JQuery-UI library, known for its rich set of user interface interactions, effects, widgets, and themes built on top of the jQuery JavaScript library. These technological choices not only enhance the performance and user experience of FLoRA but also ensure that it remains adaptable and scalable, facilitating continuous improvements and contributions from the open-source community.

4.2 Log Server

An advanced log portal has been established to facilitate researchers' access to and retrieval of student trace data. This log interface, as illustrated in Figure 11, has a well-structured layout. The header section of the interface serves as a navigation area,

¹<https://github.com/hypothesis>

²<https://github.com/PButcher/flipdown>

offering links to various categories of logs, encompassing both processed and raw (unprocessed) data. The processed logs are presented with action labels and process labels, as delineated in Section 3.2 and Appendix A, offering a clear and systematic view of the data. A search tool embedded within the interface enables the efficient location of logs through participant ID or specific keywords. This enhances the accessibility and usability of the log data for research purposes. The lower segment of the page is dedicated to displaying the logs, providing a comprehensive view of the data collected. Additionally, the portal is equipped with an export tool, allowing for the straightforward extraction of all log data. While it is feasible to retrieve student logs directly from the back-end database, this approach necessitates advanced technical skills and poses a potential risk to the integrity of the original data. Therefore, the implementation of the log portal serves as a vital measure to safeguard the raw data against inadvertent modification or damage, ensuring the preservation of data integrity for research purposes.

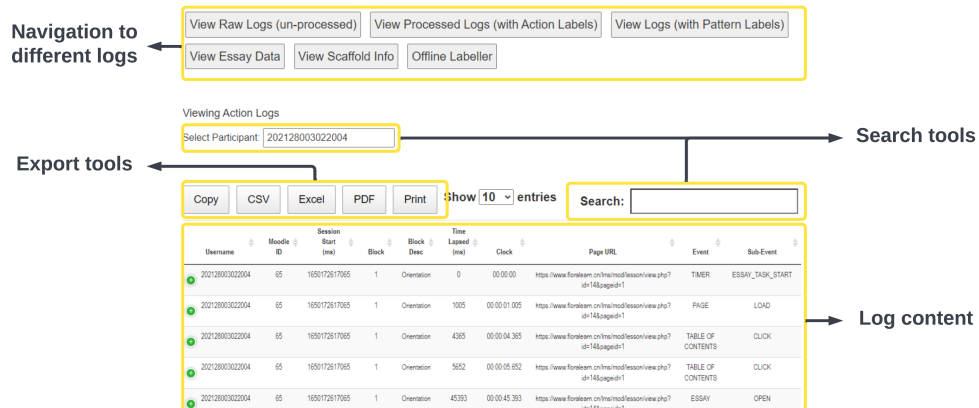


Figure 11. Log server interface.

4.3 Infrastructure

The FLoRA engine, now implemented in over 10 universities worldwide and serving a significant student demographic, requires a robust, scalable, and reliable infrastructure (as displayed in Figure 12). As previously described, the architecture comprises three core modules: instrumentation tools, a trace parser, and scaffolding. The instrumentation tools are developed using JavaScript, HTML, and CSS, enhanced by key JavaScript libraries such as jQuery and Hypothesis. Conversely, the trace parser and scaffolding module are crafted with PHP, aligning with modern programming norms to ensure reliability and robustness. The trace parser and scaffolding module are deployed on an Nginx server, known for its excellent performance, security, and scalability. This infrastructure choice highlights a strategic commitment to optimizing the engine’s functionality and user experience, ensuring FLoRA’s effective and stable operation in educational settings.

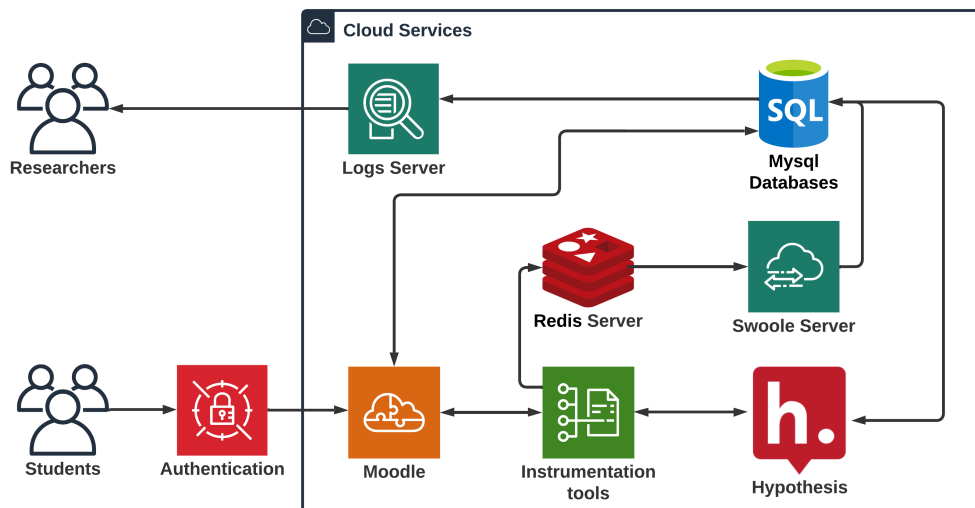


Figure 12. FLoRA engine infrastructure.

In addressing the challenges posed by the substantial volume of trace data generated during user interaction with FLoRA—including keystrokes, mouse movements, clicks, and URL visits—it becomes evident that a database that is merely fast is insufficient. To effectively manage this influx of data, the engine employs two additional services, Redis and Swoole, as middleware systems to facilitate the communication between the front end and the database.

Redis serves as an open-source in-memory data store, functioning as a data message queue. Apart from Redis, other common caching middleware are Memcached, Ehcache, Apache Ignite, and Hazelcast. Each has its own strengths and weaknesses. Memcached is simple and high performing but lacks data persistence and supports only basic data types. Ehcache integrates well with Java and supports data persistence, but its distributed caching features are less mature. Apache Ignite offers in-memory computing and SQL support but is complex to configure. Hazelcast provides strong distributed caching and rich data structures but requires significant memory and can be complex to set up. Redis stands out due to its support for diverse data structures, high performance, data persistence, robust distributed capabilities, and rich feature set, making it a preferred choice for many applications. Its large community and extensive ecosystem further enhance its flexibility and usability in various scenarios. This configuration allows for the efficient caching of data prior to its insertion into a MySQL database.

Meanwhile, the Swoole server, functioning as an HTTP server, is used to asynchronously process database insertions. There are some other similar options, including ReactPHP, Workerman, and Amp. ReactPHP is known for its event-driven architecture and simplicity but lacks built-in co-routine support, leading to potential performance bottlenecks. Workerman offers high performance and ease of use but is not as feature rich and lacks native co-routine support. Amp provides a powerful asynchronous framework with good co-routine support but has a steeper learning curve and less community support. Swoole distinguishes itself due to its built-in co-routine support, high performance, and extensive feature set, including an integrated HTTP server, WebSocket server, task workers, and timer support. Its robust ecosystem and active community make it a superior choice for building scalable, high-performance PHP applications. This multi-faceted approach to data management underscores the FLoRA engine's adaptability and capacity to handle extensive data loads with efficiency.

As illustrated in Figure 12, the operational flow within the FLoRA engine initiates with the aggregation of user-generated trace data at the front-end interface. In a strategic move to circumvent the inefficiencies associated with direct database interactions, all collected data are initially routed to a Redis server, using a queue data structure. This server temporarily stores the data in memory, thus optimizing data handling processes. Subsequently, the Swoole server is used to execute asynchronous tasks, which include retrieving data from the Redis server and conducting batch insertions into the MySQL database. The integration of this middleware architecture significantly enhances overall system performance, ensuring both speed and efficiency in data processing. Furthermore, the FLoRA engine, in conjunction with the LMS, boasts a versatile deployment capability. It can be easily deployed into various server environments, including but not limited to Amazon Web Services, Alibaba Cloud, and Microsoft Azure. This flexibility in deployment options not only extends the system's accessibility to a broader range of educational institutions but also allows for customization based on specific infrastructural and resource requirements. The FLoRA engine's ability to adapt to different server environments underscores its suitability for a wide array of educational contexts and its potential for global application.

4.4 Case Studies

The FLoRA engine was validated in laboratory settings and implemented as supplementary material in various authentic courses at universities in Germany, the Netherlands, and China. For example, in several studies, the task involved a 45-minute learning session during which students were required to study three topics: (1) AI, focusing on its basic concepts; (2) differentiation in the classroom, which addresses how teachers can manage differences among learners and the concept of adaptive learning; and (3) scaffolding, an essential method for supporting learners during the learning process. The learning content comprised 5,237 words and six figures. A text readability analysis was conducted to ensure that the material was appropriately challenging for students (Lim et al., 2021). The learning task required students to integrate the three topics into an essay of 300 to 400 words, describing learning in schools in 2035. The participants were undergraduate university students. In these sessions, trace data (lab and authentic classroom settings) and think-aloud data (lab settings only) were collected. Several key research findings have emerged from these studies: (1) instrumentation tools are effective in capturing SRL, as validated by think-aloud data (van der Graaf et al., 2021); (2) a novel approach that combines theory-driven and data-driven perspectives validates that the SRL processes extracted from trace data are promising and effective (Fan, van der Graaf, et al., 2022); (3) guidelines for designing and evaluating personalized scaffolds for SRL were developed (van der Graaf et al., 2023); (4) students often diverge from the plans they make at the start of their learning sessions, and this inconsistency can be attributed to influences on regulatory processes shaped not only by pre-existing learning conditions (e.g., learning resources) but also by dynamic contextual changes as learners proceed with the task (Li et al., 2024); and (5) while the personalized scaffolds induced SRL activities, they had no significant effect on learning performance (Lim et al., 2023). Further research related to the FLoRA studies is ongoing, with more papers expected to be published in the future.

Additional studies have been conducted, modifying the reading materials to include content provided by course coordinators and tutors. This content extends beyond the scope of the original tasks that were used during the project in which the FLoRA

engine was developed. These studies aimed to investigate the applicability of FLoRA in diverse learning contexts. For example, reading materials were changed to “Academic English Writing” for a postgraduate course in China and the task length was changed to 2 hours, while in Australia, Finland, Germany, and the Netherlands, reading materials included “AI in education,” “AI in medicine,” and “Biology” and the task time was limited to 60 minutes for secondary school students (12- to 15-year-olds), who used FLoRA in their classrooms. These studies also enabled us to compare SRL processes between secondary and higher education students, which is an overlooked topic in the literature (Cheng et al., 2025). Educators and researchers have provided valuable feedback on the use of the FLoRA engine, emphasizing its capacity to unobtrusively and automatically measure and diagnose competency levels in large-scale student cohorts. This capability is primarily attributed to the integration of the FLoRA engine with Moodle. Courses and studies that frequently use FLoRA often incorporate supplementary quizzes or questionnaires designed to gather learners’ pre- and post-study knowledge and collect learners’ feedback. These quizzes or questionnaires were set up in the Moodle environment. The FLoRA engine can access the Moodle database in the back end to retrieve the results of these quizzes or questionnaires, enabling the use of data collected by the Moodle LMS. This functionality of FLoRA has the potential to facilitate personalized scaffolding by incorporating data-driven insights so that students can follow the guidance provided by the scaffolding (Lim et al., 2024). Ongoing research is focused on examining how FLoRA influences learning outcomes across different learning contexts and whether SRL processes vary during learning (Li et al., 2024; Lim et al., 2024).

5. Key Comparisons and Limitations

The FLoRA engine propels the advancement of research in SRL through its specialized design, which enables the detection, measurement, and facilitation of students’ SRL processes across a wide range of learning activities. This paper provides a detailed illustration of the design logic and theoretical foundation underlying all of the instrumentation tools within FLoRA. Previous sections include examples demonstrating how the FLoRA engine has been used in various studies, with research outcomes confirming the utility of these instrumentation tools in capturing SRL activities. For instance, the timer was used for monitoring, the highlighter and note-taker were associated with both low and high cognitive activities, and the search and planner tools were primarily used for planning and monitoring, aligning well with their intended purposes (van der Graaf et al., 2021).

5.1 Comparison Analysis

For the analogous SRL support systems/engines, including nStudy (Winne et al., 2019), MetaTutor (Azevedo et al., 2022), and Betty’s Brain (Munshi et al., 2023), a comparative analysis was conducted to evaluate them across various dimension, such as program portability and reusability, the theoretical model/framework employed, and the instrumentation tools supported.

Program portability is a critical aspect where FLoRA and nStudy demonstrate superior adaptability. FLoRA’s design, as a suite of JavaScript components, allows for seamless integration with other LMSs or learning environments. Similarly, nStudy (Winne et al., 2019), functioning as a Google Chrome plugin, is versatile and can operate across various websites. In contrast, MetaTutor (Azevedo et al., 2022) and Betty’s Brain (Munshi et al., 2023) are comprehensive, self-contained systems tailored for specific tasks, with integrated instrumentation tools. This design necessitates the use of the entire system, which may present a learning curve for new users. However, with FLoRA and nStudy, learners can continue working within familiar environments, easing the transition and enhancing user experience.

Regarding reusability, FLoRA stands out due to its high customizability. It employs a configuration file allowing modifications to the tools’ content, scaffolding content, scaffold triggering timing, and task duration. Different studies can use the FLoRA engine by incorporating simple configuration files, obviating the need for program modification. Similarly, nStudy (Winne et al., 2019) also offers tool content customization. MetaTutor (Azevedo et al., 2022) and Betty’s Brain (Munshi et al., 2023) can also be employed in various learning scenarios; however, they do not offer a straightforward method for configuring all support instrumentation tools for reuse (such as modifying tool content, enabling or disabling tools, altering scaffolding content, adjusting the timing for scaffold triggers, and changing task duration) and may require additional development. All systems support textual and image resources, but only FLoRA and nStudy (Winne et al., 2019) extend support to video resources.

Theoretical underpinnings are essential for SRL engines, with FLoRA informed by the Bannert (2007) framework, while the others align with Winne and Hadwin’s COPES model (Winne & Hadwin, 1998). The choice of theoretical underpinning influences the development of supported tools, ensuring that they capture the corresponding SRL processes effectively. FLoRA, in particular, encompasses a comprehensive set of tools, as detailed in Table 2, sufficient for capturing a range of SRL processes. However, specific tools like the bookmark tool may become redundant in environments with a navigation area, as in the case of FLoRA. If the learning task requires using a large volume of websites or web pages, the bookmark tool would be useful. Additionally, the concept map tool, designed for specific tasks, is not commonly employed in reading and writing tasks.

Table 2. Comparison between different SRL support engine/systems.

	FLoRA	nStudy	MetaTutor	Betty’s Brain
Portability	Compatible with various LMSs	Compatible with any websites	Limited for specific tasks	Limited for specific tasks
Reusability	Configurable for tools content, scaffolding content, scaffold trigger time, and task length	Configurable for tools content	Development required	Development required
Support Resources	Textual, images, video	Textual, images, video	Textual, images	Textual, images
Theoretical Model/Framework	Bannert (2007)	Winne and Hadwin (1998)	Winne and Hadwin (1998)	Winne and Hadwin (1998)
Support Tools	Bookmark Tool	no	yes	no
	Annotation Tool	yes	yes	yes
	Search Tool	yes	yes	no
	Timer Tool	yes	no	yes
	Planner Tool	yes	no	no
	Writing Tool	yes	yes	no
	Scaffolding Tool	yes	no	yes
	Discussion Tool	yes	yes	no
Concept Map Tool	no	no	no	yes

In summary, the comparison of these SRL support systems reveals that while each has its strengths and unique features, FLoRA exhibits greater portability and reusability, making it more adaptable to various learning environments and tasks. The strengths of other systems will be implemented in future development, such as converting the FLoRA engine to a browser plugin and supporting more tools like the concept map tool.

5.2 FLoRA Engine Limitations

The current limitations of the FLoRA engine have been identified through our previous research outcomes, primarily stemming from design constraints that will be addressed in future development. For example, the search function and planner were underutilized, possibly due to the structured nature of the learning materials and the relatively short duration of the tasks. Besides, the FLoRA engine did not have the functionality to analyze the text selected for highlights or notes, which could provide deeper insights into processes and constructs of relevance to SRL. Furthermore, FLoRA currently lacks the capability to assess learners’ academic backgrounds and prior knowledge.

The trace parser in the FLoRA engine employs the Bannert (2007) framework to categorize SRL processes. Future research could benefit from incorporating additional models, such as COPEs (Winne & Hadwin, 1998), to analyze the same trace data. This comparative approach would facilitate a more comprehensive evaluation of the effectiveness of different models in recognizing SRL processes, potentially leading to more nuanced understandings of SRL dynamics and enabling theory development.

The current scaffolding support in FLoRA has limitations in terms of generalizability. The existing design could not easily be applied to other studies without integrating the specific learning context into the scaffolding framework. The scaffolding design is evidence based and must be tailored to meet the unique needs of different learning contexts. Although the FLoRA engine can be conveniently configured for tasks of varying duration and scaffold content, the distinctive characteristics of each learning condition must be considered. This includes determining the appropriate time intervals for scaffold triggers and the specific content to be displayed to learners.

In summary, while the FLoRA engine has the potential to significantly promote SRL, its current limitations necessitate further attention. Addressing these limitations through future research and development will be essential for enhancing the engine’s effectiveness and applicability across diverse learning contexts.

6. Conclusion

The FLoRA engine offers useful tools for researchers, enabling the acquisition of comprehensive data on students’ learning behaviours. By examining these data, researchers can link learning stages to metacognitive and cognitive actions, providing critical insights into student development in SRL. Furthermore, FLoRA offers customized and diminishing scaffolding to support the evolving competence of learners, promoting autonomy. To enhance collaborative and open resource sharing in the

academic community, FLoRA is available under an open-source licence, allowing for extensive customization across various educational settings and tasks.

Looking ahead, the development agenda for the FLoRA engine is poised to undergo significant expansion, particularly in its data collection capabilities. This expansion is envisioned to be realized through the integration of additional, sophisticated tools, such as a video recording tool, an academic writing advisor tool, a grammar check tool, and a collaborative writing tool, enhancing the engine's ability to gather and analyze educational data. The video recording tool will capture students' facial expressions, allowing for the use of computer vision techniques to determine whether students are focused on their learning tasks and to identify signs of confusion. This capability will provide valuable insights into students' engagement and cognitive states, which are critical factors in understanding their SRL processes. The academic writing advisor tool will offer detailed academic writing suggestions, helping students improve their writing skills. By providing real-time feedback and tailored advice, this tool aims to enhance students' ability to produce well-structured and coherent academic texts. The grammar check tool will automatically identify and suggest corrections for grammatical errors in students' writing. This tool will support students in refining their writing, ensuring clarity, and maintaining high linguistic standards. The collaborative writing tool will facilitate writing tasks in collaborative learning environments. By enabling multiple students to work together on writing assignments, this tool will support the development of collaborative writing skills and foster a deeper understanding of tasks through peer interaction. The primary objective of integrating these tools is to enhance the detection and analysis of students' SRL processes, as well as to monitor their development across various task scenarios. By leveraging these sophisticated tools, the FLoRA engine aims to provide educators and researchers with deeper insights into how students manage and regulate their learning strategies over time. This integration not only enhances the ability to assess the effectiveness of different pedagogical approaches but also enables the customization of learning experiences to better meet individual student needs.

In addition to these developments, a significant focus of future development will be the incorporation of chatbots (e.g., ChatGPT) based on large language models into the FLoRA engine. This integration represents a pioneering step in investigating the intricate relationship between a human's SRL and AI (Gašević et al., 2023). The primary objectives of this initiative include exploring whether AI-powered agents can match or surpass human teachers in effectiveness, understanding the nuances of human-AI interaction for SRL, and evaluating the potential of ChatGPT in enhancing SRL (Järvelä et al., 2023).

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

FLoRA is funded by DFG, ESRC and NWO as part of the Open Research Area (Call 5) under grant number BA 2044/10-1 — GA 2739/1-1 — MO 2698/1-1. This research is also partially funded by the Australian Government through the Australian Research Council (project number DP240100069 and DP220101209) and the Jacobs Foundation (CELLA 2 CERES).

References

- Ahmad Uzir, N., Gašević, D., Matcha, W., Jovanović, J., Pardo, A., Lim, L.-A., & Gentili, S. (2019). Discovering time management strategies in learning processes using process mining techniques. In M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider (Eds.), *Transforming learning with meaningful technologies. EC-TEL 2019. Lecture notes in computer science* (pp. 555–569, Vol. 11722). Springer. https://doi.org/10.1007/978-3-030-29736-7_41
- Alexiou, A., & Paraskeva, F. (2015). Managing time through a self-regulated oriented eportfolio for undergraduate students. In G. Conole, T. Klobučar, C. Rensing, J. Konert, & E. Lavoué (Eds.), *Design for teaching and learning in a networked world. EC-TEL 2015. Lecture notes in computer science* (pp. 547–550, Vol. 9307). Springer. https://doi.org/10.1007/978-3-319-24258-3_56
- Alvarez, R. P., Jivet, I., Pérez-Sanagustin, M., Scheffel, M., & Verbert, K. (2022). Tools designed to support self-regulated learning in online learning environments: A systematic review. *IEEE Transactions on Learning Technologies*, 15(4), 508–522. <https://doi.org/10.1109/TLT.2022.3193271>
- Azevedo, R., Bouchet, F., Duffy, M., Harley, J., Taub, M., Trevors, G., Cloude, E., Dever, D., Wiedbusch, M., Wortha, F., & Cerezo, R. (2022). Lessons learned and future directions of MetaTutor: Leveraging multichannel data to scaffold self-regulated learning with an intelligent tutoring system. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.813632>
- Azevedo, R., Bouchet, F., Harley, J., Behnagh, R. F., Trevors, G., Duffy, M., Taub, M., Pacampara, N., Agnew, L., Griscom, S., Mudrick, N., Stead, V. E., & Yang, W. (2011). MetaTutor: An intelligent multi-agent tutoring system designed to detect, track, model foster self-regulated learning. *Proceedings of the Fourth Workshop on Self-Regulated Learning in Educational Technologies*. <https://doi.org/10.13140/RG.2.1.1334.6640>

- Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523. <https://doi.org/10.1037/0022-0663.96.3.523>
- Azevedo, R., & Gašević, D. (2019). Analyzing multimodal multichannel data about self-regulated learning with advanced learning technologies: Issues and challenges. *Computers in Human Behavior, 96*, 207–210. <https://doi.org/10.1016/j.chb.2019.03.025>
- Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development, 56*(1), 45–72. <https://doi.org/10.1007/s11423-007-9067-0>
- Bannert, M. (2007). *Metakognition beim lernen mit hypermedia*. Waxmann.
- Bannert, M. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie, 23*(2), 139–145. <https://doi.org/10.1024/1010-0652.23.2.139>
- Bannert, M., & Mengelkamp, C. (2013). Scaffolding hypermedia learning through metacognitive prompts. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 171–186). Springer. https://doi.org/10.1007/978-1-4419-5546-3_12
- Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning, 9*(2), 161–185. <https://doi.org/10.1007/s11409-013-9107-6>
- Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From design to implementation to practice a learning by teaching system: Betty's Brain. *International Journal of Artificial Intelligence in Education, 26*(1), 350–364. <https://doi.org/10.1007/s40593-015-0057-9>
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. <https://doi.org/10.1146/annurev-psych-113011-143823>
- Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review, 18*, 199–210. <https://doi.org/10.1007/s10648-006-9013-4>
- Bowman, N. A., Jang, N., Kivlighan, D. M., Schneider, N., & Ye, X. (2020). The impact of a goal-setting intervention for engineering students on academic probation. *Research in Higher Education, 61*, 142–166. <https://doi.org/10.1007/s11162-019-09555-x>
- Cerezo, R., Bogarín, A., Esteban, M., & Romero, C. (2020). Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education, 32*(1), 74–88. <https://doi.org/10.1007/s12528-019-09225-y>
- Cheng, Y., Guan, R., Li, T., Raković, M., Li, X., Fan, Y., Jin, F., Tsai, Y.-S., Gašević, D., & Swiecki, Z. (2025). Self-regulated learning processes in secondary education: A network analysis of trace-based measures (submitted). In *Proceedings of the 15th International Conference on Learning Analytics and Knowledge (LAK 2025)*, 3–7 March 2025, Dublin, Ireland. <https://arxiv.org/pdf/2412.08921>
- Chi, M. T. (2021). Translating a theory of active learning: An attempt to close the research-practice gap in education. *Topics in Cognitive Science, 13*(3), 441–463. <https://doi.org/10.1111/tops.12539>
- Clarebout, G., Elen, J., Collazo, N. A. J., Lust, G., & Jiang, L. (2013). Metacognition and the use of tools. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 187–195). Springer. https://doi.org/10.1007/978-1-4419-5546-3_13
- Clark, R. C., & Mayer, R. E. (2016). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. John Wiley & Sons. <https://doi.org/10.1002/9781119239086>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Davis, D., Triglianios, V., Hauff, C., & Houben, G.-J. (2018). SRLx: A personalized learner interface for MOOCs. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachslers, R. Elferink, & M. Scheffel (Eds.), *Lifelong technology-enhanced learning. EC-TEL 2018. Lecture notes in computer science* (pp. 122–135, Vol. 11082). Springer. https://doi.org/10.1007/978-3-319-98572-5_10
- Devolder, A., van Braak, J., & Tondeur, J. (2012). Supporting self-regulated learning in computer-based learning environments: Systematic review of effects of scaffolding in the domain of science education. *Journal of Computer Assisted Learning, 28*(6), 557–573. <https://doi.org/10.1111/j.1365-2729.2011.00476.x>
- Di Vesta, F. J., & Gray, G. S. (1973). Listening and note taking: II. Immediate and delayed recall as functions of variations in thematic continuity, note taking, and length of listening-review intervals. *Journal of Educational Psychology, 64*(3), 278–287. <https://doi.org/10.1037/h0034589>
- Fan, Y., Lim, L., van der Graaf, J., Kilgour, J., Raković, M., Moore, J., Molenaar, I., Bannert, M., & Gašević, D. (2022). Improving the measurement of self-regulated learning using multi-channel data. *Metacognition and Learning, 17*, 1025–1055. <https://doi.org/10.1007/s11409-022-09304-z>

- Fan, Y., Matcha, W., Uzir, N. A., Wang, Q., & Gašević, D. (2021). Learning analytics to reveal links between learning design and self-regulated learning. *International Journal of Artificial Intelligence in Education*, 31(4), 980–1021. <https://doi.org/10.1007/s40593-021-00249-z>
- Fan, Y., Rakovic, M., van Der Graaf, J., Lim, L., Singh, S., Moore, J., Molenaar, I., Bannert, M., & Gašević, D. (2023). Towards a fuller picture: Triangulation and integration of the measurement of self-regulated learning based on trace and think aloud data. *Journal of Computer Assisted Learning*, 39(4), 1303–1324. <https://doi.org/10.1111/jcal.12801>
- Fan, Y., van der Graaf, J., Lim, L., Raković, M., Singh, S., Kilgour, J., Moore, J., Molenaar, I., Bannert, M., & Gašević, D. (2022). Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, 17, 949–987. <https://doi.org/10.1007/s11409-022-09291-1>
- Fincham, E., Gašević, D., Jovanović, J., & Pardo, A. (2018). From study tactics to learning strategies: An analytical method for extracting interpretable representations. *IEEE Transactions on Learning Technologies*, 12(1), 59–72. <https://doi.org/10.1109/TLT.2018.2823317>
- Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 59(1), 64–71. <https://doi.org/10.1007/s11528-014-0822-x>
- Gašević, D., Siemens, G., & Sadiq, S. (2023). Empowering learners for the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, 4, 100130. <https://doi.org/10.1016/j.caeai.2023.100130>
- Gormley, C., & Tong, Z. (2015). *Elasticsearch: The definitive guide: A distributed real-time search and analytics engine*. O'Reilly Media, Inc. <https://www.oreilly.com/library/view/elasticsearch-the-definitive/9781449358532/>
- Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology*, 34(1), 18–29. <https://doi.org/10.1016/j.cedpsych.2008.05.006>
- Guo, L. (2022). Using metacognitive prompts to enhance self-regulated learning and learning outcomes: A meta-analysis of experimental studies in computer-based learning environments. *Journal of Computer Assisted Learning*, 38(3), 811–832. <https://doi.org/10.1111/jcal.12650>
- Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding complex systems: Some core challenges. *The Journal of the Learning Sciences*, 15(1), 53–61. https://doi.org/10.1207/s15327809jls1501_7
- Huang, L., & Lajoie, S. P. (2021). Process analysis of teachers' self-regulated learning patterns in technological pedagogical content knowledge development. *Computers & Education*, 166, 104169. <https://doi.org/10.1016/j.compedu.2021.104169>
- Järvelä, S., Nguyen, A., & Hadwin, A. (2023). Human and artificial intelligence collaboration for socially shared regulation in learning. *British Journal of Educational Technology*, 54(5), 1057–1076. <https://doi.org/10.1111/bjet.13325>
- Lajoie, S. P. (2021). Student modeling for individuals and groups: The BioWorld and HOWARD platforms. *International Journal of Artificial Intelligence in Education*, 31(3), 460–475. <https://doi.org/10.1007/s40593-020-00219-x>
- Lee, D., Watson, S. L., & Watson, W. R. (2019). Systematic literature review on self-regulated learning in massive open online courses. *Australasian Journal of Educational Technology*, 35(1). <https://doi.org/10.14742/ajet.3749>
- Li, T., Fan, Y., Srivastava, N., Zeng, Z., Li, X., Khosravi, H., Tsai, Y.-S., Swiecki, Z., & Gašević, D. (2024). Analytics of planning behaviours in self-regulated learning: Links with strategy use and prior knowledge. In *Proceedings of the 14th International Conference on Learning Analytics and Knowledge (LAK 2024)*, 18–22 March 2024, Kyoto, Japan (pp. 438–449). ACM. <https://doi.org/10.1145/3636555.3636900>
- Li, T., Lin, J., Iqbal, S., Swiecki, Z., Tsai, Y.-S., Fan, Y., & Gašević, D. (2023). Do learners appreciate adaptivity? an epistemic network analysis of how learners perceive adaptive scaffolding. In G. Arastoopour Irgens & S. Knight (Eds.), *Advances in quantitative ethnography. ICQE 2023. Communications in computer and information science* (pp. 3–17, Vol. 1895). Springer. https://doi.org/10.1007/978-3-031-47014-1_1
- Lim, L., Bannert, M., van der Graaf, J., Fan, Y., Rakovic, M., Singh, S., Molenaar, I., & Gašević, D. (2024). How do students learn with real-time personalized scaffolds? *British Journal of Educational Technology*, 55(4), 1309–1327. <https://doi.org/10.1111/bjet.13414>
- Lim, L., Bannert, M., van der Graaf, J., Molenaar, I., Fan, Y., Kilgour, J., Moore, J., & Gašević, D. (2021). Temporal assessment of self-regulated learning by mining students' think-aloud protocols. *Frontiers in Psychology*, 12, 749749. <https://doi.org/10.3389/fpsyg.2021.749749>
- Lim, L., Bannert, M., van der Graaf, J., Singh, S., Fan, Y., Surendrannair, S., Rakovic, M., Molenaar, I., Moore, J., & Gašević, D. (2023). Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Computers in Human Behavior*, 139, 107547. <https://doi.org/10.1016/j.chb.2022.107547>
- Maier, U., & Klotz, C. (2022). Personalized feedback in digital learning environments: Classification framework and literature review. *Computers and Education: Artificial Intelligence*, 3, 100080. <https://doi.org/10.1016/j.caeai.2022.100080>

- Marzouk, Z., Rakovic, M., Liaqat, A., Vytasek, J., Samadi, D., Stewart-Alonso, J., Ram, I., Woloshen, S., Winne, P. H., & Nesbit, J. C. (2016). What if learning analytics were based on learning science? *Australasian Journal of Educational Technology*, 32(6). <https://doi.org/10.14742/ajet.3058>
- Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). Detection of learning strategies: A comparison of process, sequence and network analytic approaches. In M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider (Eds.), *Transforming learning with meaningful technologies. EC-TEL 2019. Lecture notes in computer science* (pp. 525–540, Vol. 11722). Springer. https://doi.org/10.1007/978-3-030-29736-7_39
- Matcha, W., Gašević, D., Uzir, N. A., Jovanovic, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Perez-Sanagustin, M., & Tsai, Y.-S. (2020). Analytics of learning strategies: Role of course design and delivery modality. *Journal of Learning Analytics*, 7(2), 45–71. <https://doi.org/10.18608/jla.2020.72.3>
- Matcha, W., Uzir, N. A., Gašević, D., & Pardo, A. (2020). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, 13(2), 226–245. <https://doi.org/10.1109/TLT.2019.2916802>
- Molenaar, I., Roda, C., van Boxtel, C., & Slegers, P. (2012). Dynamic scaffolding of socially regulated learning in a computer-based learning environment. *Computers & Education*, 59(2), 515–523. <https://doi.org/10.1016/j.compedu.2011.12.006>
- Munshi, A., Biswas, G., Baker, R., Ocumpaugh, J., Hutt, S., & Paquette, L. (2023). Analysing adaptive scaffolds that help students develop self-regulated learning behaviours. *Journal of Computer Assisted Learning*, 39(2), 351–368. <https://doi.org/10.1111/jcal.12761>
- Nelson, T., & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). MIT Press. <https://doi.org/10.7551/mitpress/4561.003.0003>
- Nussbaumer, A., Kravcik, M., Renzel, D., Klamma, R., Berthold, M., & Albert, D. (2014). A framework for facilitating self-regulation in responsive open learning environments. *arXiv preprint arXiv:1407.5891*. <https://doi.org/10.48550/arXiv.1407.5891>
- Oluwatosin, H. S. (2014). Client-server model. *IOSR Journal of Computer Engineering*, 16(1), 67–71. <https://doi.org/10.9790/0661-16195771>
- Paans, C., Molenaar, I., Segers, E., & Verhoeven, L. (2019). Temporal variation in children's self-regulated hypermedia learning. *Computers in Human Behavior*, 96, 246–258. <https://doi.org/10.1016/j.chb.2018.04.002>
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 422. <https://doi.org/10.3389/fpsyg.2017.00422>
- Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4), 49–64. <https://www.jstor.org/stable/jeductechsoci.17.4.49>
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4, 239–257. <https://doi.org/10.1007/s11412-009-9070-z>
- Saint, J., Fan, Y., Gašević, D., & Pardo, A. (2022). Temporally-focused analytics of self-regulated learning: A systematic review of literature. *Computers and Education: Artificial Intelligence*, 3, 100060. <https://doi.org/10.1016/j.caeai.2022.100060>
- Saint, J., Gašević, D., Matcha, W., Uzir, N. A., & Pardo, A. (2020). Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. In *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK 2020)*, 23–27 March 2020, Frankfurt, Germany (pp. 402–411). ACM. <https://doi.org/10.1145/3375462.3375487>
- Saint, J., Whitelock-Wainwright, A., Gašević, D., & Pardo, A. (2020). Trace-SRL: A framework for analysis of microlevel processes of self-regulated learning from trace data. *IEEE Transactions on Learning Technologies*, 13(4), 861–877. <https://doi.org/10.1109/TLT.2020.3027496>
- Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, 3(3), 9–45. <https://doi.org/10.18608/jla.2016.33.3>
- Sharma, P., & Hannafin, M. J. (2007). Scaffolding in technology-enhanced learning environments. *Interactive Learning Environments*, 15(1), 27–46. <https://doi.org/10.1080/10494820600996972>
- Siadat, M., Gašević, D., & Hatala, M. (2016). Trace-based micro-analytic measurement of self-regulated learning processes. *Journal of Learning Analytics*, 3(1), 183–214. <https://doi.org/10.18608/jla.2016.31.11>
- Srivastava, N., Fan, Y., Rakovic, M., Singh, S., Jovanovic, J., Van Der Graaf, J., Lim, L., Surendrannair, S., Kilgour, J., Molenaar, I., Bannert, M., Moore, J., & Gasevic, D. (2022). Effects of internal and external conditions on strategies of self-regulated learning: A learning analytics study. In *Proceedings of the 12th International Conference on Learning*

- Analytics and Knowledge* (LAK 2022), 21–25 March 2022, online (pp. 392–403). ACM. <https://doi.org/10.1145/3506860.3506972>
- Uzir, N. A., Gašević, D., Jovanović, J., Matcha, W., Lim, L.-A., & Fudge, A. (2020). Analytics of time management and learning strategies for effective online learning in blended environments. In *Proceedings of the 10th International Conference on Learning Analytics and Knowledge* (LAK 2020), 23–27 March 2020, Frankfurt, Germany (pp. 392–401). ACM. <https://doi.org/10.1145/3375462.3375493>
- van der Graaf, J., Lim, L., Fan, Y., Kilgour, J., Moore, J., Bannert, M., Gašević, D., & Molenaar, I. (2021). Do instrumentation tools capture self-regulated learning? In *Proceedings of the 11th International Conference on Learning Analytics and Knowledge* (LAK 2021), 12–16 April 2021, Irvine, California, USA (pp. 438–448). ACM. <https://doi.org/10.1145/3448139.3448181>
- van der Graaf, J., Lim, L., Fan, Y., Kilgour, J., Moore, J., Gašević, D., Bannert, M., & Molenaar, I. (2022). The dynamics between self-regulated learning and learning outcomes: An exploratory approach and implications. *Metacognition and Learning*, 17(3), 745–771. <https://doi.org/10.1007/s11409-022-09308-9>
- van der Graaf, J., Raković, M., Fan, Y., Lim, L., Singh, S., Bannert, M., Gašević, D., & Molenaar, I. (2023). How to design and evaluate personalized scaffolds for self-regulated learning. *Metacognition and Learning*, 18, 783–810. <https://doi.org/10.1007/s11409-023-09361-y>
- Veenman, M. V. (2007). The assessment and instruction of self-regulation in computer-based environments: A discussion. *Metacognition and Learning*, 2(2), 177–183. <https://doi.org/10.1007/s11409-007-9017-6>
- Winne, P. H. (2017). Learning analytics for self-regulated learning. In C. Lang, G. Siemens, A. F. Wise, & D. Gašević (Eds.), *Handbook of learning analytics* (pp. 241–249). SoLAR. <https://doi.org/10.18608/hla17.021>
- Winne, P. H. (2018). Theorizing and researching levels of processing in self-regulated learning. *British Journal of Educational Psychology*, 88(1), 9–20. <https://doi.org/10.1111/bjep.12173>
- Winne, P. H., & Hadwin, A. (1998). Studying as self-regulated learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice*. Erlbaum. <https://psycnet.apa.org/record/1998-07283-011>
- Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50045-7>
- Winne, P. H., Teng, K., Chang, D., Lin, M. P.-C., Marzouk, Z., Nesbit, J. C., Patzak, A., Raković, M., Samadi, D., & Vytasek, J. (2019). nStudy: Software for learning analytics about processes for self-regulated learning. *Journal of Learning Analytics*, 6(2), 95–106. <https://doi.org/10.18608/jla.2019.62.7>
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10, 3087. <https://doi.org/10.3389/fpsyg.2019.03087>
- Yau, J.-K., & Joy, M. (2009). A mobile context-aware framework for supporting self-regulated learners. In Kinshuk, D. G. Sampson, J. M. Spector, P. Isaías, & D. Ifenthaler (Eds.), *Proceedings of the IADIS International Conference on Cognition and Exploratory Learning in Digital Age* (CELDA 2009), 22–24 November 2009, Rome, Italy (pp. 415–418). IADIS Digital Library. <https://www.iadisportal.org/digital-library/a-mobile-context-aware-framework-for-supporting-self-regulated-learners>
- Yousef, A. M. F., Chatti, M. A., Danoyan, N., Thijs, H., & Schroeder, U. (2015). Video-mapper: A video annotation tool to support collaborative learning in MOOCs. *Proceedings of the Third European MOOCs Stakeholders Summit* (EMOOCs 2015), 18–20 May 2015, Mons, Belgium, 131–140. https://www.researchgate.net/profile/Mattias-Mano/publication/319402795_A_Tale_of_two_MOOCs_Analyzing_long-term_course_dynamics/links/59a82831a6fdcc2e1bdfbb17/A-Tale-of-two-MOOCs-Analyzing-long-term-course-dynamics.pdf#page=131
- Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, 11(4), 307–313. [https://doi.org/10.1016/0361-476X\(86\)90027-5](https://doi.org/10.1016/0361-476X(86)90027-5)
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50031-7>
- Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, 48(3), 135–147. <https://doi.org/10.1080/00461520.2013.794676>

Appendices

Appendix A: Library of Action Labels and Processing Labels

Table 3. Action library table.

#	Action	Description
1	GENERAL_INSTRUCTION	Learners read or reread the general instructions and learning goals.
2	RUBRIC	Learners read or reread the rubric for essay writing.
3	RELEVANT_READING	Learners read and learn learning content for the first time.
4	RELEVANT_RE-READING	Learners reread and review for learning content that they have read before.
5	IRRELEVANT_READING	Learners read the pages that are not relevant to the learning goal and learning task.
6	IRRELEVANT_RE-READING	Learners reread the pages that are not relevant to the learning goal and learning task.
7	NAVIGATION	(1) Learners quickly navigate through pages by quick reading and clicking, (2) learners look/scroll at catalogue area or page, or (3) learners open the table of contents page.
8	OPEN_ESSAY	Learners (1) open, (2) read, and (3) think about the essay.
9	WRITE_ESSAY	Learners (1) write and (2) edit or rewrite.
10	EDIT_ANNOTATION	On all pages, learners (1) create, (2) delete, or (3) edit the annotation.
11	READ_ANNOTATION	On all pages, learners (1) read, (2) reread, or (3) look at annotations.
12	LABEL_ANNOTATION	Learners create tags for annotations.
13	SEARCH_ANNOTATION	On all pages, learners use the search function to search notes or highlights.
14	TIMER	On all pages, learners check the timer during learning.
15	SEARCH_CONTENT	On all pages, learners use the search function to search contents.
16	PLANNER	On all pages, learners (1) create planner, (2) edit planner, or (3) check planner.
17	OFF_TASK	Learners do not have any action for a relatively long time.
18	SCAFFOLDING	Learners interact with scaffoldings.

Table 4. Scaffolding sub-action library table.

#	Action	Sub-action
1	SCAFFOLDING	Message_Triggered
2	SCAFFOLDING	Message_Displayed
3	SCAFFOLDING	Notification_Clicked
4	SCAFFOLDING	Message_Closed
5	SCAFFOLDING	MessageOption_Checked
6	SCAFFOLDING	MessageOption_UnChecked
7	SCAFFOLDING	CreateChecklist
8	SCAFFOLDING	CurrToDoList_Displayed
9	SCAFFOLDING	PrevToDoList_Displayed
10	SCAFFOLDING	CurrToDoList_Edit
11	SCAFFOLDING	PrevToDoList_Edit
12	SCAFFOLDING	ToDoList_Closed
13	SCAFFOLDING	CurrToDoListItem_Checked
14	SCAFFOLDING	CurrToDoListItem_UnChecked
15	SCAFFOLDING	PrevToDoListItem_Checked
16	SCAFFOLDING	PrevToDoListItem_UnChecked
17	SCAFFOLDING	CurrToDoList_Re-Ordered
18	SCAFFOLDING	PrevToDoList_Re-Ordered
19	SCAFFOLDING	PrevToDoListItem_ClickedLink
20	SCAFFOLDING	NextToDoListItem_ClickedLink

Main Category	Subcategory	Code	Priority	Pattern No.	Process Pattern	Description	
Metacognition	Orientation	MC.O	3	MC.O.1	GENERAL_INSTRUCTION/RUBRIC->(NAVIGATION*)-> RELEVANT_READING	After reading the general instruction page, learners read through the catalogue to get an overview of what topics they need to learn, then read some pages.	
			16	MC.O.2	GENERAL_INSTRUCTION/RUBRIC-> (GENERAL_INSTRUCTION/RUBRIC)	Learners read general instruction and then check the rubric for essay, or the other way around.	
			12	MC.O.3	GENERAL_INSTRUCTION/RUBRIC<-> EDIT_ANNOTATION	Learners read the general instruction and create/edit annotations during reading.	
				15	MC.O.4	GENERAL_INSTRUCTION/RUBRIC<-> NAVIGATION	Learners read the general instruction and navigate through pages during reading instruction.
	Planning	MC.P	4	MC.P.1	PLANNER->(NAVIGATION*)-> RELEVANT_READING	Learners make plan or check plan using planner tool, then navigate to read something.	
			10	MC.P.2	GENERAL_INSTRUCTION/RUBRIC<-> PLANNER* (during first 15 mins)	Learners open planner tool and make personal plan after reading instruction.	
			21	MC.P.3	PLANNER* (during first 15 mins)	Learners open planner tool and make personal plan in the first 15 minutes.	
			24	MC.P.4	SEARCH_CONTENT*	Learners search certain contents according to instruction and plan.	
	Evaluation	MC.E	1	MC.E.1	READING->(NAVIGATION*)-> GENERAL_INSTRUCTION*/RUBRIC*/READ_ANNOTATION*->(NAVIGATION*)-> READING	Learners check instruction/rubric when they read some irrelevant pages then move on to read some relevant pages, which showed evaluation pattern of these irrelevant materials.	
			2	MC.E.2	OPEN_ESSAY->(NAVIGATION*)-> GENERAL_INSTRUCTION*/RUBRIC*/READ_ANNOTATION*->(NAVIGATION*)-> READING/WRITE_ESSAY	Learners check instruction/rubric when they read their own writing then move on to write or read something else, which showed evaluation pattern of these writings.	
			13	MC.M.1	NAVIGATION<-> READ_ANNOTATION	Learners read previous notes during navigation.	
			11	MC.M.2	GENERAL_INSTRUCTION/RUBRIC<-> PLANNER* (after the first 15 mins)	Learners check progress according to instruction and plan.	
			14	MC.M.3	WRITE_ESSAY/OPEN_ESSAY<-> PLANNER/GENERAL_INSTRUCTION/RUBRIC (after the first 15 mins)	Learners check planner, while writing the essay, to monitor if there is enough time to finish the essay.	
25			MC.M.4	TIMER*	Learners check timer during reading or writing as monitoring the learning time.		
Monitoring	MC.M	22	MC.M.5	PLANNER* (after the first 15 mins)	Learners open planner tool and check or change personal plan after the first 15 minutes.		
		23	MC.M.6	SEARCH_ANNOTATION*	Learners search certain annotations.		
		26	MC.M.7	READ_ANNOTATION*	Learners read certain annotations.		
		33	MC.M.8	OPEN_ESSAY-> READING	Learners read content after checking their own essay.		

Figure 13. Process label library: Metacognition.

Main Category	Subcategory	Code	Priority	Pattern No.	Process Pattern	Description	
Cognition	Low_cognition	L.C.F	7	L.C.F.1	(IR)RELEVANT_READING -> EDIT_ANNOTATION* -> (IR)RELEVANT_READING	Learners read and highlight new learning materials.	
			9	L.C.F.2	(IR)RELEVANT_READING -> (NAVIGATION*) -> (IR)RELEVANT_READING	Learners read something new and then navigate to next page to continue reading.	
			19	L.C.F.3	(IR)RELEVANT_READING <-> EDIT_ANNOTATION/READ_ANNOTATION	During reading, learners take notes or highlight something; they might also read their notes or highlight.	
			20	L.C.F.4	(IR)RELEVANT_READING <-> (IR)RELEVANT_READING	Learners read several pages continuously.	
			31	L.C.F.5	IRRELEVANT_READING -> (IRRELEVANT_READING)	Learners read new irrelevant learning materials.	
			32	L.C.F.6	RELEVANT_READING -> (RELEVANT_READING)	Learners read new relevant learning materials for some minutes.	
	High_cognition	Elaboration/Organization	H.C.E/O	29	L.C.R.1	RELEVANT_RE-READING -> (RELEVANT_RE-READING)	Learners review some relevant learning materials they have been reading before.
				30	L.C.R.2	IRRELEVANT_RE-READING -> (IRRELEVANT_RE-READING)	Learners review some relevant learning materials they have been reading before.
				6	H.C.E/O.1	(IR)RELEVANT_RE-READING -> (NAVIGATION*) -> WRITE_ESSAY -> (WRITE_ESSAY)	After rereading some learning materials, learners write their essay.
				5	H.C.E/O.2	GENERAL_INSTRUCTION/RUBRIC -> (NAVIGATION*) -> WRITE_ESSAY -> (WRITE_ESSAY)	After reading the rubric, learners write their essay.
				18	H.C.E/O.3	WRITE_ESSAY -> (WRITE_ESSAY)	Learners write the essay continuously.
				17	H.C.E/O.4	WRITE_ESSAY/OPEN_ESSAY <-> READ_ANNOTATION	While writing the essay, learners read their highlights or notes.
			27	H.C.E/O.5	LABEL_ANNOTATION*	Learners create or choose labels for the highlights they made during learning.	
			28	H.C.E/O.6	EDIT_ANNOTATION*	Learners take notes during learning.	

Figure 14. Process label library: Cognition.