MILESTONES: The Design and Field Evaluation of a Semi-Automated Tool for Promoting Self-Directed Learning Among Online Learners

Rimika Chaudhury
Computing Science
Simon Fraser University
Burnaby, British Columbia, Canada
rimika_chaudhury@sfu.ca

Courtenay Huffman
Cognitive Science
Simon Fraser University
Burnaby, British Columbia, Canada
chuffman@sfu.ca

Isabelle Kwan
Computing Science
Simon Fraser University
Burnaby, British Columbia, Canada
itk1@sfu.ca

Gurnoor S. Deol
Computing Science
Simon Fraser University
Burnaby, British Columbia, Canada
gsd15@sfu.ca

Supreet Dhillon
Computing Science
Simon Fraser University
Burnaby, British Columbia, Canada
supreet_singh_dhillon@sfu.ca

Parmit K Chilana
Computing Science
Simon Fraser University
Burnaby, British Columbia, Canada
pchilana@sfu.ca

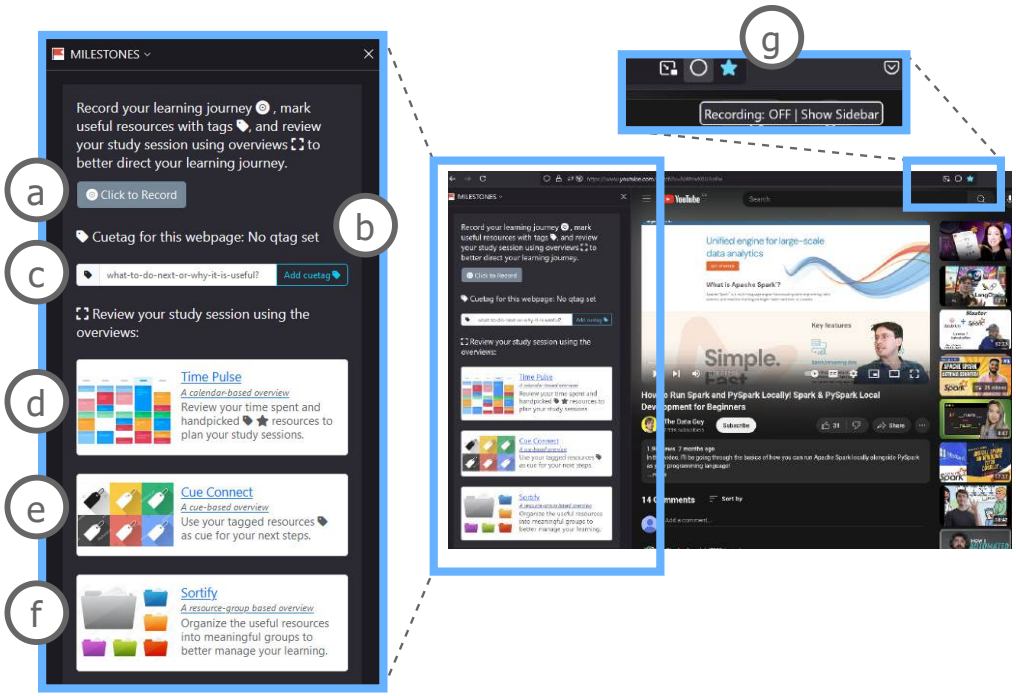

Figure 1: MILESTONES is implemented as a browser plugin and accessed through a side-bar. a) A recording button allows the users to start/stop recording their study sessions (URL, time, tags and bookmarks). b) MILESTONES allows associating tags with currently active web-resources. c) Input for assigning a new tag to the current web-resource. Three interactive visual overviews of the collected data show time spent via Time Pulse (d) (Fig. 4), tagged resources via Cue-Connect (e) (Fig. 4), and automatically categorized resources via Sortify (f) (Fig. 3). g) A page action button allows the side-bar to be opened and closed on demand.

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3714295

Abstract

Self-directed learning of computational skills online poses significant challenges, particularly the lack of effective tools for tracking progress and fostering reflection. To address this, we designed and implemented MILESTONES, a semi-automated self-monitoring tool that tracks online learning sessions and organizes web resources through three visual overviews: Time Pulse, Cue-Connect, and Sortify. In a week-long field deployment study (N=17), learners found MILESTONES intuitive and effective, even without prior experience with self-monitoring. The on-demand visual overviews encouraged learners to pause, reflect, and adjust their learning habits to better align with their goals. These overviews further fostered *micro-reflections* - brief, spontaneous reflections *during* learning. We also explored the role of a companion journal, which, although used inconsistently, helped learners form and reflect on their goals *after* learning sessions. Our findings contribute insights for designing learner-centered semi-automatic self-monitoring tools that can cater to diverse learning needs.

CCS Concepts

• **Human-centered computing → Interactive systems and tools**.

Keywords

Semi-Automated Tracking, Self-Directed Learning, Interactive Visualizations, Micro-Reflections

ACM Reference Format:

Rimika Chaudhury, Courtenay Huffman, Isabelle Kwan, Gurnoor S. Deol, Supreet Dhillon, and Parmit K Chilana. 2025. MILESTONES: The Design and Field Evaluation of a Semi-Automated Tool for Promoting Self-Directed Learning Among Online Learners. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3706598.3714295

1 Introduction

In today's rapidly evolving technological landscape, an increasing number of individuals are seeking to learn new skills on their own either to complement or in lieu of traditional classroom education [52, 69]. Whether pursuing in-demand computational skills [28] such as programming, or other technical expertise, these *self-directed learners* often rely on online resources, peer networks, and on-the-job experiences to guide their learning [10, 15, 66]. However, despite the wealth of information available, self-directed learners pursuing computational skills face unique challenges that differ from those in formal educational settings, including difficulties in locating resources [12, 15, 66] and a lack of reflective opportunities during the learning process [12, 24, 50]. To self-direct effectively, learners must identify knowledge gaps, find relevant resources, and plan strategically to achieve their goals [40].

While there are a growing number of tools such as MOOCs [7, 63] and coding-specific interventions focused on comprehension, debugging, and problem-solving [23, 31, 45], they often overlook the broader challenge of supporting learners in managing their self-directed learning processes related to planning, managing resources and self-assessing progress. Moreover, most of the existing tools assume a structured learning context or generalized use cases, which may not align with the specific learning needs and context of self-directed learners who are often end-user programmers with targeted computational needs [61]. Without a structured approach to learning, learners often feel overwhelmed and unsure of how to consistently achieve meaningful results. [29].

In this work, we complement the ongoing efforts to support the needs of self-directed learners of computational skills by exploring the design of an interactive tool aimed at empowering learners to actively self-monitor their learning progress and foster meaningful self-reflection. However, designing such interventions presents significant challenges, including understanding the diverse motivations and objectives of learners, identifying data, and integrating these into tools that seamlessly blend with the existing practices of the learners [12, 18, 24, 47]. Recent research suggests that learners may benefit from viewing data collected from their learning sessions, such as resource specific time usage information, and interactions with their data may be conducive to fostering reflection [11, 13]. Drawing inspiration from self-tracking tools in domains such as health, well-being, and productivity [18, 38, 48, 54], we adapt several of the existing self-tracking principles to address the unique challenges of self-monitoring learning. Our key guiding research question is: *How can we enhance learners' awareness of the effectiveness of their learning approaches and help them make more informed decisions about their learning practices?* Furthermore, how can we create opportunities for self-reflection during online learning to support their learning approaches? Through these considerations, we aim to bridge the gap between knowledge acquisition and self-awareness, providing learners with the tools needed to take control of their self-directed learning journey.

We present the design and implementation of **MILESTONES** (Monitoring Informal Learning Experiences - Semi-Automated Tool for Observing Online Events in Learning) a system designed to facilitate self-monitoring and self-reflection among learners. MILESTONES introduces meaningful pauses for reflection by automatically collecting information about accessed web resources and the time spent on each. To organize this data, we designed three different temporal and resource-based overviews: 1) *Time-Pulse* (Fig. 4), which provides a calendar-based overview of automatically categorized web-resources accessed during study sessions to help learners reflect on their time spent; 2) *Cue-Connect* (Fig. 4), which enables learners to filter and associate resources by tags, to facilitate reflection on topics aligned with their learning objectives; and 3) *Sortify* (Fig. 3), which allows learners to reflect on their engagement with different types of resource groups, such as text-based tutorials or video-based lectures, to foster awareness of the diversity in their learning materials.

We deployed MILESTONES with 17 participants, with various levels of expertise, learning computational skills, such as programming languages, frameworks, and feature-rich software (see Table 1). Their goals ranged from understanding foundational concepts to implementing complex code. We evaluated how well MILESTONES assisted with promoting reflections and assisted with setting and refining learning objectives. Unlike traditional classroom-based studies, which use materials and assessments pre-determined by the instructors [42, 58], our field study provides critical insights into learners' real-world goals and self-directed online learning behaviors, which are difficult to capture in lab environments.

Our key findings indicate that the learners actively engaged with the on-demand visual overviews as they could monitor their sessions with minimal effort. Overviews like Time Pulse and Sortify simplified self-monitoring through automation, while Cue-Connect provided greater control over resource curation. Initially, most participants associated progress primarily with task completion. However, MILESTONES helped broaden their perspective to include goal adherence, resource evaluation, and topic understanding as meaningful indicators of progress. Cue-tagging and topic grouping in the visual overviews facilitated *micro-reflections*, making learners more aware of their habits and offering opportunities to bridge gaps between learning practices and goals.

In conclusion, our work highlights the important role that semi-automated self-monitoring tools can play in supporting self-directed learning in online environments. Automation reduces the burden of resource management, allowing learners to focus on actual learning tasks. Tools that provide on-demand interactive visual presentations of collected data - such as time spent and resources accessed - can raise awareness among a wide range of learners. The HCI community should further explore the design of self-monitoring tools for learning to enable individuals to make more informed decisions during learning. Our work makes the following contributions to the field of Human-Computer Interaction:

(1) **Design and implementation of MILESTONES**, a semi-automated tool that supports self-monitoring and reflection by collecting and presenting learning data.
(2) **Insights into self-monitoring practices from a field deployment** uncovering the diverse needs and challenges of learners, and demonstrating how self-monitoring can blend with existing learning practices.
(3) **Design implications for creating reflective interventions** to enhance engagement by catering to the immediate learning needs and the nuances of learners' unique reflective practices.

2 Related Work

2.1 Self-Reflections and Self-Directed Learning in Informal Online Settings

In this work, we adopt Malcolm et. al.'s perspective that "formality" and "informality" exist in all learning contexts [51]. For example, informal learning may occur within notionally formal educational contexts. Recent studies of introductory computer science (CS1) students show that they often prioritize quick answers from online forums, or AI-generated responses, bypassing instructor-curated beginner resources [67, 69]. Using Malcolm et al.'s criteria, we classify online learning resources as an *informal learning setting*, characterized by open-endedness, minimal time constraints, and self-directed learning processes [51]. *Self-directed learning* involves identifying goals, selecting resources, and evaluating progress, as outlined in Knowles' framework [40]. However, while self-directed learning depends on reflection and strategy adjustment, learners often struggle to gauge progress [12, 68]. For learners of computational skills, self-direction is common, but practices like self-reflection and evaluation are often neglected [8, 24]. Self-monitoring interventions can address these challenges by helping learners monitor learning sessions and evaluate their progress [11, 13]. Our design and evaluation of MILESTONES demonstrates that offering simple, on-demand visualizations of personal learning data - such as time spent, resource access, and topics covered - can support reflection, session planning, and resource curation. By providing interactive overviews of learning behaviors, such tools empower learners to assess their practices, refine their approaches, and take greater control of their self-directed learning.

2.2 Technologies for Supporting Self-Monitoring and Reflections

Current approaches to monitoring learning in large classrooms often prioritize institutional needs, offering static visualizations of metrics like activity logs and test results to inform instructor and administrator needs [6, 14]. However, while these methods help predict student performance, they offer learners little support for self-reflection or planning across diverse online resources [5, 26]. Without insights into their learning practices or progress, learners struggle to assess their approaches or make informed adjustments. In contrast, self-monitoring tools in health, well-being, and productivity empower individuals by providing interactive, on-demand visualizations that foster reflection and decision-making [19, 25, 37]. By linking data to personal goals and progress, these tools enhance engagement and promote ongoing self-reflection and adaptation. Productivity tools, such as time trackers, session recordings, and app-blockers may be used to self-regulate [32, 33], but are primarily task-focused. In contrast, self-directed learning emphasizes concept understanding and transferability to new contexts. Our work bridges these domains by adapting self-monitoring techniques from health and productivity to support learning. Building on prior research that emphasizes the value of resource access, time spent, and topics studied for assessing self-directed learning [11, 13], we incorporate interactions like resource curation, categorization, and resource-specific time tracking to help learners reflect on past sessions and plan future ones, empowering learners to assess their approaches and make informed decisions.

2.3 Evaluating Self-Monitoring Technologies with Field Deployment

Field deployments of self-monitoring tools are essential for understanding user interactions within real tasks and contexts. Studies in health, well-being, and productivity have used *in-situ* data capture [18, 34], enabling users to reflect on their goals and practices. For example, *"Time for Break"*[49] combined tutorials, questionnaires, and interviews for user feedback, while *"Time Aware"*[37] employed a test-and-withdrawal approach to evaluate its impact. In the context of learning, we shift from classroom-based studies focused on assessments [42, 58] to self-directed online learning, emphasizing instances of reflection rather than test results or task completion. Guided by Schön's theory of reflective practice [62] - distinguishing reflection-in-action (during activity) and reflection-on-action (after activity)—and Fleck and Fitzpatrick's taxonomy [27], we focus on explanatory reflections (interpreting observations), dialogic reflections (identifying patterns or causations), and transformative reflections (reassessing actions to drive change). Prior research also highlights the value of structured reflection, (e.g., journaling [2, 16, 30]), particularly when visualizing

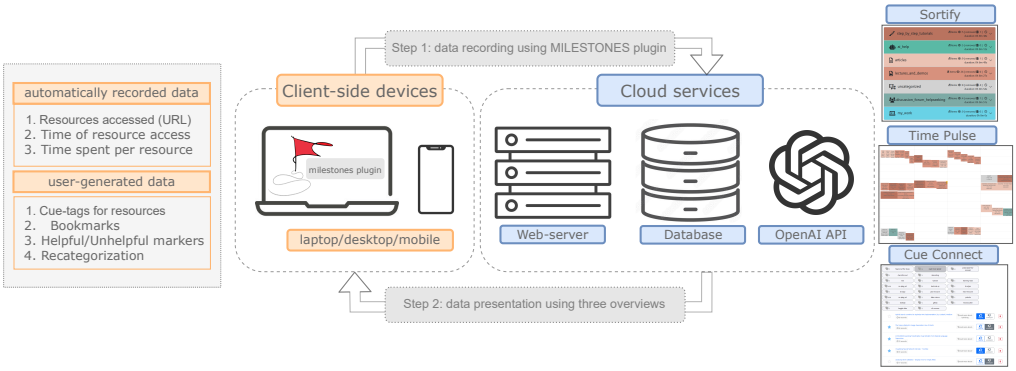

Figure 2: MILESTONES plugin, deployed on users' laptops/desktops sends automatically collected and user-generated data during recording sessions to the cloud database. The user can access the three overviews, which are hosted on a cloud web-server, from any device and view their personal data. The overviews are also connected to OpenAI API to perform automatic resource-type categorizations for Sortify.

the collected data or relying solely on system-driven prompts may be insufficient for users [17, 20]. To explore this, we introduced a non-digital journal as an optional companion, offering participants an open-ended way to incorporate MILESTONES data into their reflections and plan study goals. Our field study contributes to the learning domain by capturing real-world insights into self-directed learners' contexts, self-monitoring approaches, and reflective practices - which are challenging to obtain in lab or classroom settings.

3 Design and Implementation of MILESTONES

Our overarching design goal was to *build a learner-centered tool that creates awareness of and reflection on the effectiveness of learning approaches using a self-monitoring technique and real-time visual feedback.* Research on self-monitoring in health, well-being, and productivity domains highlights key design considerations for technology interventions, including data collection, data presentation, motivational components, material or medium, and interactive features [18, 19, 35, 38, 39, 53, 54, 56]. Drawing from this literature, we developed MILESTONES with the following considerations:

C1: Reduce the burden of data-collection by allowing semi-automatic tracking: Previous research [18, 37] has shown that people value the convenience of automatic tracking but prefer to maintain control over certain aspects of data collection. Therefore, the system should combine automatic tracking with manual controls to allow users to choose what to record.

C2: Enable engagement with self-monitored data by allowing exploration through interactive visual overviews: Visualizations of personal data have been found to be beneficial for awareness and reflection in several domains, particularly when they allow deeper exploration of specific aspects of the data [19, 39].

C3: Support selection and filtering of relevant resources: The self-monitoring data should enable learners to form action items for subsequent study sessions. The system should offer overviews that highlight or support the filtering of meaningful and useful resources from wasteful efforts [65].

C4: Provide cues for progressive goal-setting: Simple progress indicators within overviews, such as learner specified action-items, could help manage learning and provide a sense of progress without burdening the learners to assess themselves while studying [1].

3.1 System Design

After several rounds of brainstorming, iterative prototyping and testing based on the above goals, we arrived at a design for MILESTONES: a semi-automated self-monitoring tool designed to help learners reflect on their online learning activities. By enabling in-situ automatic data collection, the tool encourages learners to be more aware of their time spent and learning resources used. Learners can access 3 overviews on-demand (Sortify, Time Pulse and Cue-Connect) which offer a different way of looking at the same data and create opportunities for learners to pause and reflect on their learning decisions and approach. Below, we list the key design decisions we made to implement MILESTONES:

1) In-situ semi-automatic data recording and overviews: Addressing C1 and C2, MILESTONES enables users to automatically capture their learning activities by tracking web resources, time of access, and time spent on each resource during a recorded study session. The tool, implemented as a browser plugin, addresses C1 by automatically saving any visited, bookmarked or tagged resources, allowing users to focus on relevant learning materials. To ensure control over what is recorded, MILESTONES provides a start-and-stop function, enabling users to exclude non-learning activities from being tracked. A visual indicator lets the user know when recording is active, giving them full awareness and control during their study session. As the session progresses, MILESTONES organizes the collected data into three visual layouts that users can access on their primary device (such as a laptop or desktop) or on their mobile phone.

2) Interactive visual overviews of personal learning histories: To address C2, MILESTONES enables users to view three real-time visual overviews generated automatically during

their learning session. These overviews: Sortify, Time Pulse, and Cue-Connect - present data relevant to the user's learning, such as time spent on resources (Fig. 4), key concepts or tasks pursued (Fig. 4), and the types of learning resources chosen by the learner (Fig. 3). Each overview is designed to be easily recognizable and interactive, allowing users to engage with their learning data in a way that enhances reflection and supports effective learning.

a) Sortify: In the Sortify overview (Fig. 3), MILESTONES automatically groups all resources recorded during a study session into two categories: general resources and specific resources, visually distinguished using a complementary color palette. General resources include materials that provide broad information related to the user's topic of interest, such as articles, lectures, demos, and step-by-step tutorials. Specific resources, on the other hand, are more directly tied to the user's work or interests, such as discussion forum entries, AI conversations (e.g., with ChatGPT or Gemini), and project-specific documents created by the user. The shades within each broad group are used to distinguish and discretize between the specific subgroups. Additionally, a third category, labeled uncategorized, includes any resources that do not fit into the predefined groups. Users can expand each category to view the specific resources within it (Fig. 3.b). MILESTONES also enables users to re-categorize resources if they have been misclassified or remove irrelevant ones, giving them full control over how their study materials are organized.

b) Time Pulse: The Time Pulse overview (Fig. 4.1) provides a calendar-based view that automatically displays a curated set of web resources accessed during the study session. To address C3 and help users identify the most useful resources, Time Pulse presents resources that have been either bookmarked or cue-tagged by the user. Opening in a week-view, Time Pulse presents these resources as events on the calendar, starting from the time they were accessed in the most recent session. By default, each resource occupies a one-hour slot to display its title, time, and tag information clearly. However, if more time is spent on a resource, its corresponding slot grows in length. Additionally, resources categorized by Sortify are color-coded, giving users a visual indication of how effectively they've utilized their study time.

c) Cue-Connect: The Cue-Connect overview (Fig. 4.2) organizes resources based on tags, which MILESTONES enables users to associate with resources during a recording session. To address C4, MILESTONES draws from the concept of task-based to-do lists to create *cue-tags*, which are concise, dash-delimited phrases that serve as reminders or cues for action items related to each resource(Fig. 4.a). MILESTONES also provides a default list of default cue-tags, which users can either adopt directly or modify to suit their specific needs. A single cue-tag can be associated with multiple resources, allowing users to group related materials by tasks or topics of interest. By clicking on a specific cue-tag, users can filter and display a subset of resources that share the same tag, making it easier to focus on specific tasks or learning goals.

3) Aggregated Weekly Overviews: In addition to the on-demand overviews, users can access three static, aggregate overviews. Two bar graphs present the time spent and cue-tags created across different resource types. A third aggregate overview shows how frequently the participant switches between different resource types using an arc diagram (Fig. 6).

While Time Pulse is a completely automatically generated calendar-based overview, requiring no interaction, Sortify and Cue-Connect allow users to indicate which resources are helpful/unhelpful and remove extraneous data. All the overviews are updated live and therefore can be accessed during a study session even when a recording is on-going (albeit, requiring to be refreshed for the latest data to reflect).

MILESTONES also includes a mobile version of the three overviews. While the mobile version does not provide recording, it presents the recorded data on demand, with all the interactions available in the desktop version, for a quick review on-the-go.

3.2 Implementation Details

MILESTONES (Fig. 2) is implemented as a browser plugin. The overviews are each implemented as separate websites using various Javascript frameworks and libraries. The Time Pulse utilizes a calendar-based library using the Vue framework, called Vue-Cal [3]. Cue-Connect and Sortify are both implemented as an Express App. Sortify additionally utilizes OpenAI API [55] to perform automatic categorizations. All the overviews are connected to a remote database and hosted on a managed server to make them available on desktop and mobile. When recording starts, the browser plugin automatically records the current tab's URL, time of access, and the duration the tab is active. Additionally, cue-tags and bookmarks associated with the URL are also recorded. This data is then updated live on the database and presented in the overviews on demand, in the browser. This data is then updated on the database live and presented in the overviews on demand, on the browser. When the Sortify overview loads in the browser, it connects with the OpenAI API to automatically categorize recorded web resources into groups (e.g., articles, tutorials, lectures, etc.) before displaying them. Time Pulse reflects this categorization in its presentation of the resources by using the same colors. The aggregated weekly overviews (Fig. 6) are created using Tableau (for bar graphs) and Vega-Lite (for arc diagram). To ensure privacy, we installed the browser plugin on a separate version of Firefox (Developers' Version) that participants did not use for everyday browsing. We also provided unique logins, ensuring participants could only access their own data. Additionally, participants had the ability to explicitly start or stop recording a session, giving them control over what was recorded.

4 Evaluation: Field Deployment

Self-monitoring is shaped by learners' real tasks, contexts and motivations. To understand learners' authentic contexts, we deployed MILESTONES in real-world settings, allowing learners to use the tool on their own devices, at their own convenience. We targeted individuals who were self-learning computational skills using informal online resources and willing to engage in learning for at least 7 days. While the tool facilitated recording and visualizing online learning activities, our deployment aimed to understand the factors influencing learners' self-monitoring decisions during the learning process.

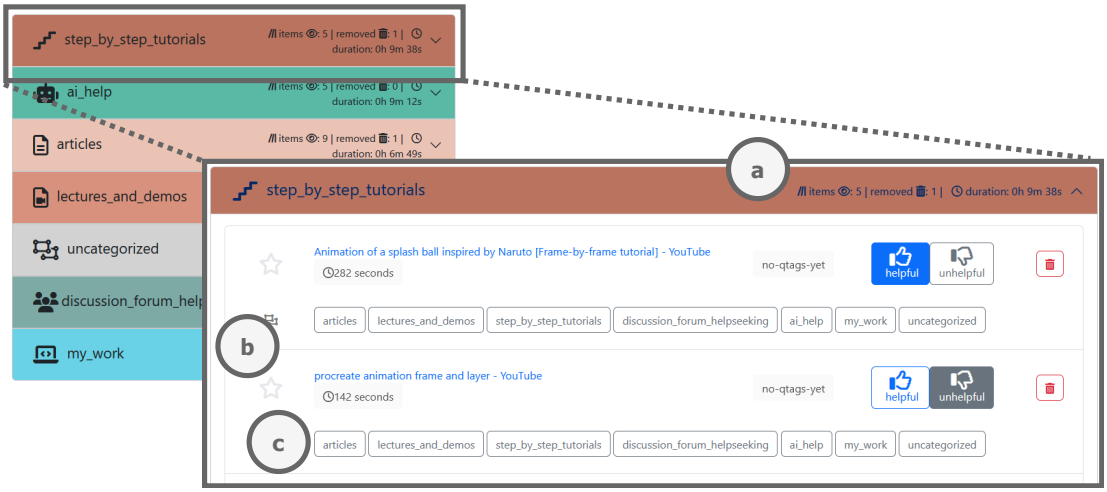

Figure 3: The Sortify Overview automatically categorizes all web-resources visited during a recorded session into broad groups (articles, lectures and videos, step-by-step tutorials) and specific resources (discussion forums, AI help, "my work"). a) The section headers display the aggregate time-spent in that group and dynamically reorganizes the sections in a descending order, giving users a sense of a resource-based view of their time. b) Each section expands to show the user specific web-resources they visited during the recorded session. c) Sortify also provides an option to recategorize the resources and mark their usefulness.

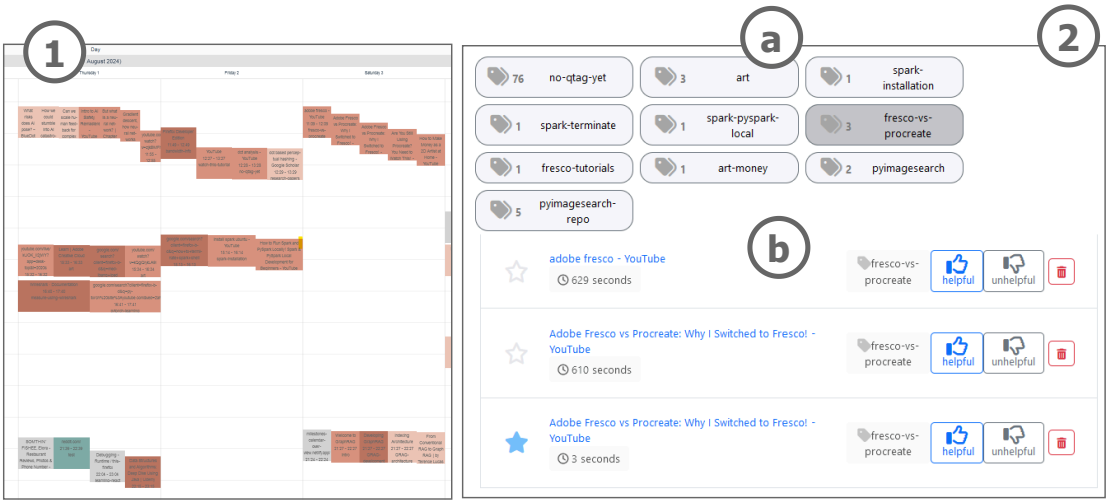

Figure 4: 1) The Time Pulse is a completely automatically generated calendar-based overview which presents the categorized web-resources in the order in which they were accessed. User can see an overview of their week (3 days shown in image), and also view a day at a time. 2) Cue-Connect organizes the cue-tags that users associate with web-resources a) Cue-tags appear as a collection. User can view and visit the collection of web-resources that share the same tag by using the cue-tags as a filter. b) The curated web-resources, based on the selected cue-tag, are displayed alongside the collection of tags allowing users to have an overview of their chosen subtopics and their hand-picked associated learning resources.

4.1 Participants

We recruited participants through university email lists, word-of-mouth, personal contacts, and online forums like LinkedIn. We prioritized local participants to distribute a companion reflective journal and sought those with limited CS knowledge but at least 1-2 months of experience to avoid overwhelming beginners. After screening, we selected 17 participants (9 men, 8 women) with diverse backgrounds and expertise in CS (Table 1), including students and professionals. During the study, participants pursued various computational skills, ranging from basic programming, such as R fundamentals (P09) and Data Structures (P07), to advanced areas like Deep Learning (P15), LLMs (P13), and iOS development (P03).

ID	Gender	Age (years)	Current Position	Topics chosen for learning
P01	Woman	25-34	MS Student - Research (CS)	Figma, Clip model, Stable diffusion
P02	Man	25-34	Instructor/Researcher (CS)	Prompt Engineering
P03	Man	25-34	UX Engineer (Professional)	iOS development
P04	Woman	25-34	MS Student - Research (CS)	Compute Express Link (CXL) tool
P05	Man	18-24	BS Student (CS)	Web dev (HTML, CSS, JS, React)
P06	Man	18-24	BS Student (CS)	Web development and Data Analysis/Science
P07	Man	18-24	BS Student (CS)	Data Science, Software Development Course
P08	Man	25-34	Data Analyst (Professional)	Python on Coursera
P09	Woman	18-24	MS Student - Biomedical Physiology	Basics of R
P10	Man	18-24	BS Student - Human Geography	Data management e.g., Excel/ArcGIS
P11	Woman	18-24	BS Student (CS)	ML, AI, Web development
P12	Woman	18-24	BS Student (CS)	Leetcode, React
P13	Woman	24-35	PhD Student - Research (CS)	LLM implementation for my research project
P14	Woman	18-24	MS Student - Research (CS)	Fresco, Procreate, Image Processing, ML
P15	Woman	25-34	MS Student - Research (CS)	Deep learning (pytorch) - VR encoding
P16	Man	25-34	MS Student - Research (CS)	Computational Imaging and Image Compression
P17	Man	25-34	PhD student - Research (CS)	Features of Zotero

Table 1: Our participants had a nearly equal distribution of men (N=9) and women (N=8). Participants also had a range of expertise in CS with several students at various stages of undergraduate and graduate studies, as well as two professionals. Participants were pursuing various types of computational skills for their personal interests.

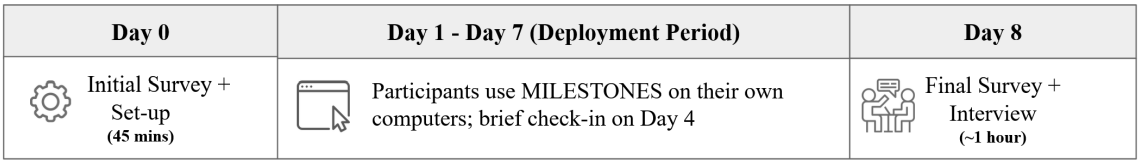

Figure 5: Study protocol: We typically met participants in person to set up MILESTONES on their laptops and mobile phones. Participants answered a pre-study questionnaire indicating their chosen topic for learning. Between Day 1 to Day 7, participants used MILESTONES for tracking their study sessions on a designated browser, with optional use of a paper-based journal to reflect on their learning progress and plan the next study session based on the data in MILESTONES. There was a mid-study check-in on Day 3 or 4 to address technical issues. At the study's end, participants completed an online questionnaire and we had a follow-up interview (45-60 minutes) to discuss their experiences with MILESTONES, and retrieve the journal or transmit images online.

4.2 Study Procedure

We designed the study as a take-home activity that lasted seven days for each participant (Fig. 5). We first had an initial meeting in person or online to deploy the MILESTONES tool. We assured the participants that there was room for flexibility if they could not engage with the tool to their satisfaction during the 7-day period and needed an extension. We scheduled the final interviews usually on Day 8 unless someone requested a later date. All 17 deployments were completed over a two-month period.

4.2.1 Pre-deployment setup: We recruited participants based on their availability and invited them to a 45-minute session in a university meeting room (except one participant who requested a home visit, and two who requested an online meeting). For those who requested to meet online, we provided the journal templates and guided the participants in assembling the journal remotely. During the session, we introduced the study and asked participants to complete a brief questionnaire about their background, existing self-monitoring practices, and learning goals for the upcoming week. We then installed the MILESTONES tool on their personal laptops/desktops and mobile phones. After a brief walkthrough, we asked participants to try out MILESTONES, and create two or three cue-tags and bookmarks based on the learning goals they indicated in the questionnaire. We also guided them through a paper-based companion reflective journal which we describe below.

4.2.2 Study protocol and tasks: In a pre-study questionnaire, we asked participants to state their learning topic of interest for the next 7 days and asked them to engage in their learning activities as usual. We installed MILESTONES on Firefox Developers' version to keep the study context separate from their usual browsing activities. We asked participants to ensure the plugin was turned on while studying and asked them to use it frequently. We also requested that they use the mobile version at least once during the

week (preferably Day 1). Additionally, we provided all learners with an optional paper-based journal to be used in conjunction with the visual overviews. We asked participants to keep the companion journal close to their learning environment to ensure it would be visible, easy to access, and serve as a reminder to use MILESTONES.

Companion Journal: Research suggests that maintaining a journal can foster deeper connections between learners and their information [1]. In particular, the use of journals for structured reflection [2, 16, 30] can be helpful when learners do not find visualizing the collected data or relying on system-driven prompts to be sufficient [17, 20]. To explore how learners would use structured reflections to plan their study goals, we designed an optional paper-based companion journal, to be free from the constraints of desktop or mobile environments and offer learners flexibility in deciding what and how to record their reflections. The journal balanced flexibility with structure to ensure its usefulness without overwhelming participants. The journal included a simple template for reflection to encourage participants to engage in open-ended reflections on their goals, successes, and challenges. It consisted of seven sheets, each representing a day of the study with three daily reflective questions: 1) a "tiny goal" for the day 2) highlights of the day, and 3) areas to work on next. Additionally, we invited participants to reflect on the data collected in MILESTONES, and write one question to ask an expert and/or frame one piece of advice they would offer to a fellow learner.

4.2.3 Midstudy check-in: To encourage engagement and resolve any technical issues with MILESTONES, we checked in with each participant on day 4.

4.2.4 Post-deployment follow-up interview: At the end of the study, participants completed a brief online questionnaire about their self-monitoring experience and we scheduled a follow-up 45-60 minute interview, typically within 1-2 days after the week of deployment. The interview focused on their experiences with MILESTONES, including factors influencing their learning decisions and any significant incidents noted during the mid-study check-in. We explored how participants' interactions with MILESTONES evolved, their views on data privacy and accuracy in self-assessment and goal setting, and their opinions on different data capture methods (using multiple devices and a physical journal). Finally, participants received assistance in uninstalling the plugin and had the option to submit the journal or send images of the responses over email.

4.3 Data Analysis

We audio-recorded the interviews with participants' consent and transcribed them for analysis. Two researchers analyzed the qualitative data using an inductive analysis approach [64], starting with open coding to examine each transcript. We focused on how participants' responses reflected their perceptions of MILESTONES, focusing on how the tool influenced their awareness of learning practices and progress toward personally relevant learning goals. We assigned multiple codes where needed and regularly discussed the coding scheme with the research team to ensure consistency. We then performed axial coding to identify themes aligned with our research questions and synthesized key insights for our results.

5 RESULTS: Insights from MILESTONES' Field Deployment

This section presents key findings from the field deployment study. We first explore participants' prior self-monitoring experiences to contextualize their perceptions of MILESTONES. Next, we examine how MILESTONES' overviews integrated into learners' existing practices, and shaped their views on progress. Finally, we tease apart the levels of reflections that occurred during learning sessions.

5.1 Prior Experiences with Self-Tracking and Learners' Expectations from MILESTONES

5.1.1 Automatic and Passive Tracking in Non-Learning Contexts. Most participants (16/17) had experience with self-monitoring tools that required minimal interaction, mainly for tracking sleep (6/17) and fitness (6/17). Apps such as *Strava* [60] were used, along with smartwatches, to automatically track quantifiable performance metrics. For example, P12 used *"Strava to track workouts, biking, running, etc., along with a smartwatch that tracks heart rate and speed."* Some participants monitored data to maintain consistency, such as P02, who used an Apple Watch daily to track *"the number of hours slept and minutes active."* A few participants (2/17) manually tracked nuanced aspects such as calories and food intake to identify problem foods and establish habits by reviewing data regularly.

5.1.2 Manual Approaches for Gauging Progress in Learning Contexts. Few participants who reported self-monitoring their progress in learning, primarily relied on task completion or topic coverage, employing both manual and digital methods. Manual tools included to-do lists and project notes (5/17), with P03 using notes to maintain focus on larger goals and P08 favoring a physical whiteboard for immediate tasks, explaining, *"I create a checklist of the topics on a whiteboard where I mark whether I have done it or not."* Digital tools were used to manage time on tasks, although participants found them challenging to use consistently. For instance, P17 used *Windows Clock* but found it cumbersome, stating, *"It's annoying that I forget to turn it off when I'm done or switch projects."* P11 used Opal to block distractions and measure productivity, noting, *"[it tracks] how often I use productive or unproductive apps."* Some participants combined techniques, such as the Pomodoro method [22] with apps like Habitica [57], or ClickUp and Asana [4], to bring more structure to their learning. Participants primarily gauged progress in binary terms, to check whether tasks were completed or if they engaged in a task on a given day. However, these methods lacked record-keeping, making it difficult to track progress over time.

When discussing expectations for the MILESTONES tool, participants were enthusiastic about understanding their learning processes and gauging progress. They envisioned the tool as a way to determine remaining tasks while others intended to use MILESTONES to manage time and resource usage, and build on their existing habits with other tools and techniques: *"I will plan my to-do list, monitor my productivity (how much I have learned and implemented), [and identify] any other resource I need to use"* (P13). Additionally, Participants expected MILESTONES to help maintain motivation and consistency by providing an overarching view of their learning tasks and topics. P05 noted, *"Having a clear picture of*

Overall Usage and Type of Data collected by MILESTONES	**Total**
Number of days of MILESTONES use per person (median)	5 days
Duration of study sessions recorded on MILESTONES per person (average)	6.06 hrs
Total number of cue-tags generated	100
Total number of bookmarks recorded during study sessions	87
Total number of unique web-resources across top 4 categories (below)	**1193**
Category 1: Step-by-step tutorials	*611*
Category 2: Lectures and demos (videos)	*268*
Category 3: Articles	*220*
Category 4: Discussion Forums	*94*

Table 2: Minimum recorded study hours and usage statistics for MILESTONES across 17 participants over an average of 5 days. On average, participants identified close to 5 topics of interest using cue-tags, and mostly referred to tutorials and lectures.

my progress may motivate me to study harder." Participants also expected MILESTONES to foster reflection and iterative planning. For instance, P17 expressed a desire to *"reflect on my learning process and update my plan accordingly."*

5.2 Usage and Usability: How MILESTONES Blended with Everyday Learning Practices

This section presents how MILESTONES supported learners' existing approaches across diverse topics (Table 1). Participants adopted various learning approaches (e.g., following tutorials or courses, or deconstructing example projects), that shaped their practices of resources and time usage and strategies to optimize these practices. Most participants (14/17) were beginners with under four months of experience, while three returned after gaps of 1, 3, and 7 years.

5.2.1 Usage of MILESTONES Overviews. Every participant used MILESTONES for one week, recording a total of 103.02 hours of study time across all participants. However, this represents the minimum recorded time, as MILESTONES only logs the most recent session for revisited web resources. We only report usage of the desktop-based overviews as most participants (except P01 and P03), did not have a significant interaction with the mobile-based overviews. On average, participants used MILESTONES for 5 days, logging an average of 6.06 hours throughout the week, with the least usage recorded being 1 day for 1 participant. All participants actively engaged with the cue-tagging and bookmarking features on the desktop browser, generating a total of 100 unique cue-tags and 87 bookmarks. In terms of resources consulted, participants most frequently referred to step-by-step tutorials, recording 611 web resources, followed by video-based lectures (268 websites) and text-based articles (220 websites). These resources were automatically categorized with minimal participant intervention or correction.

To understand the participants' preferences for visual overviews, we asked them to compare MILESTONES' overviews with three alternatives: static, aggregated weekly bar charts showing time spent by resource category (Fig 6.a), cue-tags and resource group (Fig 6.b), and an arc diagram depicting switches between resource groups (Fig 6.c). Participants appreciated the bar charts for their familiarity and found the arc diagram intriguing for uncovering hidden patterns in resource access. However, many suggested that integrating these alternatives within MILESTONES' existing visualizations would be more useful. P04 explained: *"From the bar-graphs I know how many keynotes and papers I have referred to, and videos I have watched. But I would also like to know what resources I refer to more and where I ike to study from. [MILESTONES] helps me narrow down [my resources] for the next time I want to study."*

5.2.2 Usage of MILESTONES companion journal. Six out of 17 participants did not use the companion journal, citing it as unnecessary as they had the same goal every day. Those who tried the journal appreciated its simplicity but used it sparingly, typically making brief notes, to revise goals, before or after study sessions. We observed a range in the types of learning goals set by the participants, related to tasks, topics, tools, learning resources, and time allocation. For example, participants who were focused on gaining foundational understanding indicated broader resource-management-focused goals like *"gather resources for AI, ML, and web development"* (P11) or topic-focused goals such as *"learn about functions in ArcGIS"* (P10). While participants who were more focused on skill development or practical applications set more targeted goals by specifying more than one criterion. For example, P05 specified a broad task and tool: *"build a cafe menu using CSS"*, and P17 specified a task to *"rethink which aspects of Zotero might be helpful"*, narrowing down the resource to *"mainly by reading its wiki."*

We asked participants to refer to their data on MILESTONES to frame a question to seek expert help or offer advice to others, in the journal. Journal reponses indicated that the questions mostly pertained to the overall learning approaches. For example, participants sought general directions on time management: *"What is the best way to manage my limited time and maximize my learning outcomes?"* (P05). Others questioned the relevance of learning content: *"Should I even learn syntax these days for different programming languages - we can just chatGPT or search online."* (P06). They also sought clarity to gauge learning depth and understanding: *"How much do I need to know for the basics?"* (P09). These concerns reflect a desire for support beyond task-specific needs, highlighting the need for guidance on improving overall learning approaches.

5.2.3 Meeting Learners' Needs: How MILESTONES' Overviews were Used in Daily Learning Sessions. Participants appreciated being able to monitor self-curated resources through Cue-Connect and track time using Time Pulse (Fig. 7). While Sortify's automatic categorization of resources helped participants quickly assess their use of time, it was more effective for identifying long-term patterns in resource use, compared to the immediate insights provided by

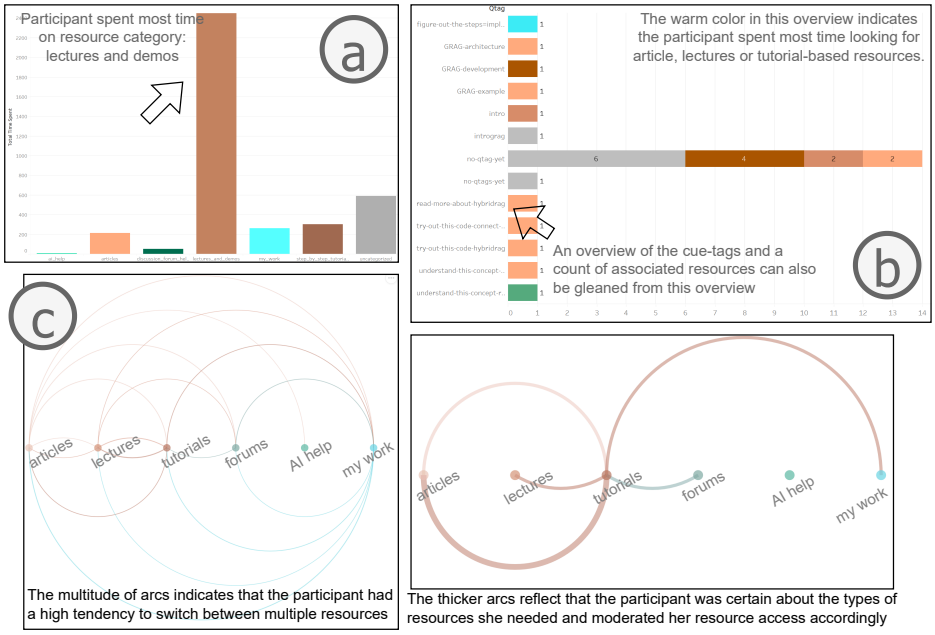

Figure 6: We showed participants three static overviews at the end of the study, including: a) aggregated time spent across resource categories, b) distribution of cue-tags across categories of resources, and c) arc diagrams comparing the frequency and pattern in moving across different groups of resources - the higher the frequency of switches, the thicker the arc. All the overviews reflected the resource categorization done by Sortify. While participants found these overviews interesting, they preferred having these as add-on features as having access to specific learning resources was perceived to be more useful for better guiding their learning decisions.

Time Pulse and Cue-Connect. We present findings on participants' use of MILESTONES' visual overviews, highlighting aspects that aligned with learners' existing learning practices.

a) Managing self-curated resources with Cue-Connect: The ability to select default cue-tags, create and re-use cue-tags from the dynamically updating list provided in MILESTONES (Fig. 1.c) integrated smoothly into participants' learning practices, encouraging them to be more deliberate in selecting resources. Participants found Cue-Connect useful for streamlining resource management, helping them group and select relevant resources for future study sessions. For instance, P10 noted, *"I found it helpful to organize the different types of resources myself to identify the relevant ones for later."* Additionally, participants found that adding cue-tags helped them focus on chosen topics during learning sessions without being disruptive. For example, P03 appreciated the ease of locating relevant resources in their recent browsing history, and shared: *"It really helped that something was recording, and I could just scroll back to [the cue-tagged resource]."*

b) Tracking time on and off-task with Time Pulse: Overall, participants found Time Pulse effective for tracking their time and gaining insights for improving time management. For example, P02 appreciated the ability to switch to *"week mode"* that gave him *"a high level overview"* of his learning journey. Participant P12 highlighted that Time Pulse allowed her to assess the effectiveness of her study habits by visualizing *"time on and off-task"*. As P15 observed *"Time Pulse shows the time I spend on each website and when I'm productive during the day"* giving her an insight on when she is most likely to study during the day. Some participants appreciated tracking time on specific topics such as P07 who stated *"I could actually look at my time on a specific day and see which topic I worked on."* Additionally, Time Pulse facilitated reflection on the relevance of study materials, as P11 pointed out *"It gives me a timestamp for each resource, so I can pinpoint when I visited it and reflect on its relevance"*. Time Pulse also helped reveal inefficiencies in learning approaches such as switching contexts too frequently, as noted by P06 who shared: *"On the first day, I was reading articles and watching the video simultaneously."*

c) Reflecting on and revising learning tasks with Sortify: Although Sortify was the least favored overview for immediate study sessions, it stood out for its automatic categorization and insights into learning patterns. For example, P16 appreciated that: *"[Sortify] updates automatically, categorizing resources into articles, lectures, and so on"*, allowing more focus on learning, rather than monitoring. Participants used Sortify to track unused resources and assess time spent: *"I could see what resources I used, what I hadn't, and for how long" (P10).* Additionally, P01 found the tool helpful for identifying challenging topics: *"I could see which category took the most time, which helps me identify the toughest areas."* P13 leveraged

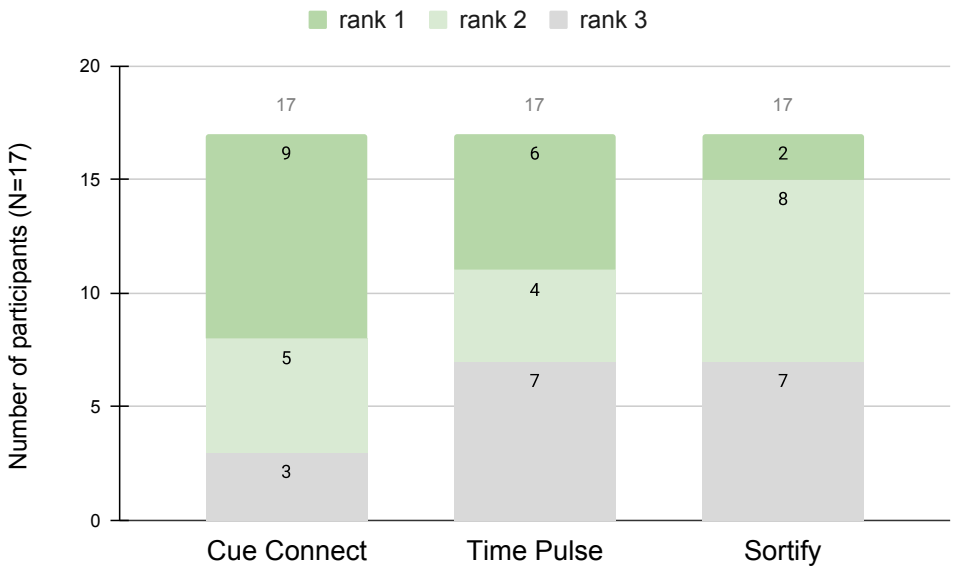

Figure 7: Most participants (14/17) ranked Cue-Connect among their top two preferences because it allowed easy access to personally curated resources. Time Pulse was the second most preferred overview (10/17) as the fine-grained time-tracking overview helped participants reflect on and strategize their time use. Sortify typically ranked second to Cue-Connect or Time Pulse. While participants appreciated Sortify's automation, its resource-based categorization was less relevant to immediate learning goals, leading to lower engagement.

Sortify to optimize her learning, explaining, *"It helped me identify unnecessary time spent on lectures when I needed specific articles for implementing a prototype."*

Participants' reflections on their use of visual overviews shows how MILESTONES integrated with existing learning practices. The three visual overviews combined supported learners address both immediate learning tasks and long-term learning goals.

5.3 MILESTONES Helped Identify Personally Meaningful Learning Indicators

We anticipated that MILESTONES would help learners widen their perspective of progress as initially most participants defined progress as task completion or time spent on a task. Through the field study, we observed how MILESTONES fostered a more nuanced understanding of progress by raising participants' awareness of deeper aspects of learning, such as topic understanding, resource quality assessment, and adherence to learning goals (Table 3).

5.3.1 Cue-tagging helped evaluate relevance of resources to goals. Learners found bookmarking and cue-tagging helpful for organizing and reflecting on resource quality and relevance to learning goals. P07, who initially tracked progress by timing his study sessions and breaks, later prioritized resources using Sortify *"during the actual study [by checking] the time spent on reading journals and articles relevant to study [goals] versus the time I spent looking at the [coding] challenge, which is additional study, not required to be done right now."* Similarly, P02 found cue-tagging helpful for grouping resources such as videos or articles that he intended to revisit: *"Adding the cue-tags to the resources I really like, such as 'read-more', or 'watch-this-video' - they filtered [relevant resources] through all the things I was looking at."*

5.3.2 Reviewing overviews supported adherence to learning goals. Participants' experiences with MILESTONES helped them become aware that progress in learning includes maintaining focus, and adhering to goals. Before using MILESTONES, P03 perceived progress as completing tasks determined through expert guidance, such as by *"talking to people who are already iOS developers."* After using MILESTONES, P03 identified inefficiencies in his learning practices, noting how cue-tagging resources helped him avoid distractions: *"I don't go into rabbit holes; I can stay focused on what I'm doing at the moment"*, and considered that as progress. Similarly, P01 who initially relied solely on clarification of misconceptions as a way to gauge progress, shared that staying on topic was a recurring challenge for her. She reflected, *"[Reviewing Overviews] helped me realize I didn't stick with my starting goal... I actually never get back to it"*, and adhering to her goals would count as progress.

5.3.3 Reviewing resource groups helped confirm topic understanding. Cue-Connect's aggregated overview of resources, organized into learner-defined groups and Sortify's automatic categorization by resource type, encouraged participants to assess their confidence and decide whether to revisit material or move on to new goals. For example, P17, who previously gauged learning by task completion, found that Cue-Connect helped him feel more engaged with his learning: *"Seeing that I spent 26 minutes to internalize that knowledge, enables me to review my learning process in a more direct, intuitive*

Personal Learning Progress Indicators	Examples from Participants
1) Evaluating resource relevance	*I began to check milestone [Sortify Overview] each time during the actual study [to check] the time spent on reading journals and articles relevant to study [goals] versus the time I spent looking at the [coding] challenge, which is additional study, not required to be done right now. (P07)*
2) Evaluating adherence to learning goals	*[Reviewing Overviews] helped me realize I didn't stick with my starting goal… I actually never get back to it. (P01)*
3) Self-checking topic understanding	*I can go into [a specific] cuetag to see how much time I spent on a specific web page, and then if I feel I didn't really understand that page, I can reopen it. (P17)*
4) Becoming aware of learning tendencies	*Usually when I go online to look up information, I prioritize videos over text; I did not know that I've been doing this. (P10)*
5) Closing the gap between goals and approaches	*My plan was to complete a step-by-step tutorial, and then try working with the data sets I already have in R, but realized after day one that it would not work, and I needed to continue with more step-by-step tutorials. (P09)*

Table 3: Using MILESTONES, learners expanded their view of progress from productivity-focused measures (e.g., task completion) to more personally meaningful learning progress indicators (e.g.,understanding, goal alignment, and self-reflection).

way.. I can go into [a specific] cue-tag to see how much time I spent on a specific web page, and then if I feel I didn't really understand that page, I can reopen it." Similarly, P16, a research student who gauged progress while planning his tasks, valued Sortify's automatic organization of resources and emphasized the importance of topic understanding: *"Organizing the articles and the topics is very hard... because when you're doing research, you do not have a particular and clear path to take, but using this kind of tool can help in the long run."* Next, we elaborate on how MILESTONES features allowed participants to assess and alter their approach to studying.

5.4 MILESTONES Facilitated Micro-Reflections During Learning Sessions

One of our goal was to promote awareness of learners' previously unrecognized tendencies through reflections on MILESTONES overviews before or after study sessions. Surprisingly, we observed occurrences of *spontaneous reflections while using MILESTONES during the study sessions*. We term these events as ***micro-reflections***. Examples include assessing relevance of resources while recording or cue-tagging, and reviewing resource access sequences in Time Pulse. These micro-reflections often led to confirmation of existing learning tendencies and/or reassessment of learning approaches to improve effectiveness. While prior research on personal informatics suggests that simply presenting data rarely fosters higher-level reflections [17, 27], our study revealed that, beyond fostering awareness through lower-level reflections, MILESTONES also supported higher-level reflections, enabling participants to reassess practices and adopt more effective strategies aligned with their goals. For our analysis, we grouped the levels of micro-reflections based on Fleck and Fitzpatrick's taxonomy [27]:

a) Level 1 Micro-Reflections which encompasses **explanatory** and **dialogic reflections**, where participants confirmed existing behaviors or gained new insights, without change of strategy.

b) Level 2 Micro-Reflections which involves **transformative reflections**, where participants reassessed and modified their learning practices with respect to time or resource use.

5.4.1 Level 1 Micro-Reflections: Becoming aware of and validating existing learning tendencies. Several participants highlighted the usefulness of tracking time spent on specific tasks and topics as opposed to tracking overall time usage. For example, P15 remarked, *"The amount of the time that you spend on each topic, you may feel like you have spent a lot of time, but you have no insight. After using MILESTONES, you can quantify how much you spend on that specific topic and categorize them. This is helpful."*. Participants also became aware of specific preferences or habits that were not consciously acknowledged before. For example, P10 shared, *"Usually when I go online to look up information, I prioritize videos over text; I did not know that I have been doing this."* Similarly, P11 reflected, *"I guess I study mostly at night."* A related theme in these micro-reflections was the realization of discrepancies between planned goals and actual behavior. For instance, P04 admitted, *"I realized that my goals didn't exactly match what I did for the day sometimes... I got distracted by some other aspect of my learning, rather than focusing on what I wanted to do that day."* Similarly, P09 noted how the visual feedback from the tool revealed unmet goals, stating, *"I set a goal of working on it for 20 minutes, but [noticed] I only worked on it for maybe 10 or 12."* These observations suggest that MILESTONES could foster explanatory and dialogic reflections, helping learners understand and confirm their tendencies during learning, thus providing a chance for adjusting their learning behaviors to more effectively achieve their learning objectives.

5.4.2 Level 2 Micro-Reflections: Closing the gap between learning goals and learning approaches. Participants demonstrated Level 2 micro-reflections, showing emerging awareness of how to align their learning practices with their learning goals, based on insights gleaned from MILESTONES' Overviews. For example, P17 shifted from passive bookmarking to intentionally tagging and rating resources, explaining, *"If I add a tag and a thumbs up, but [see that] I only spent a minute, then I know I should go back."* P16 used Sortify to prioritize resources based on their current learning goal, noting, *"Sometimes you need to watch a YouTube video to understand the concept. Sometimes you [refer to] a website, just to get an idea."*. P13 refined her approach to learning, focusing on specific articles with

integrated implementation examples instead of general introductory materials: *"Sortify helped improve my productivity by changing my approach to learning. I prioritized specific articles that were most helpful for integrating JavaScript front-end implementations."* P09 reassessed her strategy for learning R language, saying, *"My plan was to complete a step-by-step tutorial, and then try working with the data sets I already have in R, but realized after day one that it would not work, and I needed to continue with more step-by-step tutorials."*. On the other hand, P06 recognized a tendency to focus on unnecessary details in HTML, subsequently resolving to spend more time on implementation instead: *"[Sortify] could give me an insight about the time am I spending on watching videos, and the time I am spending on actually making the stuff."* These examples show how participants used the overviews to assess and improve their learning practices, leading to more goal-directed behaviors.

6 Discussion

Our design, implementation, and field study of MILESTONES, a self-monitoring tool that integrates directly into learners' contexts, complements and extends the literature on self-directed learning. The key takeaway from our work is that MILESTONES' minimal semi-automatic recording and on-demand visual overviews successfully engaged learners of computational skills in actively monitoring their learning. In particular, the automation in Time Pulse and Sortify simplified self-monitoring time and resource usage, while Cue-Connect's learner-driven curation provided control over resource management. Importantly, learners expanded their definition of progress beyond task completion, valuing goal adherence, assessing resource quality, and gauging topic understanding, as meaningful indicators of progress. Furthermore, learners demonstrated a growing awareness of their own habits, and adjusted their learning practices to better align them with their goals. Based on these findings, we derive several design implications for creating effective semi-automated self-monitoring tools that support the learning of computational skills and beyond.

6.1 Enhancing the Usefulness of Self-Monitoring in Learning

Although our participants did not commonly engage in self-monitoring practices for learning prior to the study [24, 41, 46], they appreciated the ability to record their learning sessions using MILESTONES and access on-demand visual overviews for review. Cue-Connect simplified access to curated resources, thus eliminating the need to search through extensive lists of resources accumulated over time. Building on this, cue-tagging could be improved by suggesting relevant cue-tags from learners' previously used tags when they visit new resources. Additionally, enabling learners to apply multiple cue-tags as filtering criteria could further streamline resource organization. To address the challenge of evaluating learners' understanding of new topics, which is difficult to measure without testing, self-monitoring tools, such as MILESTONES, could link curated online resources to the corresponding artifacts created by the learner (e.g., work-in-progress documents, code, or output) or their defined learning goals. Participants also valued the automation of tasks like resource categorization in Sortify and time tracking in Time Pulse. However, the recording feature, while offering control over data collection, was occasionally cumbersome as learners forgot to turn it on or off. This issue could be mitigated by introducing an automatic reminder when learners visit resources similar to those they have previously tagged or when they interact with overviews.

6.2 Leveraging Micro-Reflections to Support Progressive Goal Setting

Reflective writing is known to positively impact learning outcomes, especially for computational skills [2, 16, 30]. Based on our findings, brief structured reflective journaling could be seamlessly integrated into MILESTONES to ensure engagement without disrupting learners' focus. For instance, reminders at the start or end of sessions could prompt learners to reflect retrospectively on aggregated metrics, such as total time spent, the number of resources tagged or visited, or outputs created (e.g., documents, code files). Responses could include satisfaction ratings on a sliding scale (e.g., quality of time spent, satisfaction with progress) with an option to provide reasons. Learners should also have flexibility in choosing the frequency of journaling, whether daily, weekly, or on-demand, allowing personalized reflection habits. Interventions could also capture micro-reflections that arise spontaneously during study sessions. These reflections could be captured through short text notes, annotated screenshots, or time-lapse videos of ongoing tasks [32, 33]. Additionally, off-desktop reflections, such as conversations with peers or handwritten notes, could be captured using a mobile interface. These reflections should be integrated with MILESTONES' visual overviews for context. Learners must be able to interact with reflective responses in the context of their time and resource overviews with features to search, filter, and link reflections to goals or sessions, to support iterative goal-setting.

6.3 Considering the Impact of Self-Monitoring on Motivation and Learning

While learning is challenging due to both conceptual difficulties and responsibilities of self-direction, we saw that the emotional impact of self-monitoring tools on learners can also be significant. Although self-monitoring interventions for regulating emotions exist in the context of general well-being [21], there is a need to address the emotional impact within self-directed learning, where each stage carries an inherent emotional load [59]. It is important to factor in whether learners would feel discouraged if the data does not reflect favorable outcomes, or feel frustrated by the information overload or interruptions to their learning. In the context of our work, several aspects of MILESTONES positively influenced learners emotionally. Time Pulse motivated participants to maintain learning streaks, while Sortify reassured them of progress even when specific goals weren't met, by accruing time across resource categories. Future designs should carefully consider how the self-monitoring experience affects emotions, avoiding discouragement or frustration from data viewed as unfavorable, rather designed to improve motivation by identifying and reinforcing positive learning behaviors (e.g., learning streaks, achievements) [9, 43].

6.4 Considering How Self-Monitored Data Can Support Learners in Seeking Help

When designing tools for sharing self-monitored data and obtaining feedback, several key considerations emerge, such as determining how much information is appropriate to share [44], establishing trust between parties, and ensuring privacy around personal data in social contexts. Feedback from experts is crucial for learning [36, 70], but barriers exist in facilitating this process. Our study explored learners' initial reactions to sharing self-monitored data for seeking advice and offering help. We found that participants, guided by their self-monitored data, often wanted to ask broad questions about their learning approaches and standards, rather than specific conceptual issues. Future research should investigate how self-monitored data can be effectively integrated into conversations with experts. There is a need for better methods to represent meaningful learning patterns that can be shared with trusted individuals to aid help-seeking. Key questions include whether feedback should be provided live or asynchronously, whether the feedback should come from experts, peers, or any trusted individual of the learner's choice, and how feedback should be shared.

6.5 Limitations

While the field study of MILESTONES provided valuable insights into learners' perceptions of self-monitoring and its role in self-directed learning, several limitations should be addressed in future research. First, the 7-day duration of the study, though sufficient for our goals, may benefit from long-term engagement to better understand habit formation and compare different levels of data automation. Second, the absence of shared learning contexts (e.g., peers, mentors) limited our ability to explore how learners may use personal data to seek or provide help. Future studies should investigate the use of self-monitoring tools in collaborative learning environments. Lastly, since our study focused on individuals pursuing computational skills, future work should expand the participant pool and study learners of non-computational skills to improve the generalizability of the findings.

7 Conclusions

Through the design, implementation, and field deployment of MILESTONES, we showed how self-directed learners can leverage personal data from learning sessions to strategize learning computational skills. On-demand visual overviews seamlessly integrated into learners' practices, fostering reflection, and encouraging learners to move beyond task completion to focus on understanding topics, evaluating resource relevance, and enhancing self-awareness. To advance self-monitoring tools for learning, the HCI community should prioritize semi-automatically capturing micro-reflections during learning, supporting iterative goal-setting, enabling learners to seek expert guidance, and addressing the emotional and motivational aspects of self-monitoring. By supporting learners' immediate needs and fostering a sense of achievement, learner-centered tools can help learners reach their desired outcomes.

Acknowledgments

We thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding this research.

References

[1] Parastoo Abtahi, Victoria Ding, Anna C Yang, Tommy Bruzzese, Alyssa B Romanos, Elizabeth L Murnane, Sean Follmer, and James A Landay. 2020. Understanding Physical Practices and the Role of Technology in Manual Self-Tracking. *Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (December 2020), 1–24. https://dl.acm.org/doi/10.1145/3432236

[2] Huda Alrashidi, Thomas Daniel Ullman, and Mike Joy. 2020. An Empirical Evaluation of a Reflective Writing Framework (RWF) for Reflective Writing in Computer Science Education. In *2020 IEEE Frontiers in Education Conference (FIE)*. IEEE, IEEE, 1–9.

[3] Antoni Andre. 2024. Vue Cal - Date Picker & Calendar Component - Made with Vue.js. https://madewithvuejs.com/vue-cal. Accessed: 2024-09-12.

[4] Shivani Arya and Mugdha Shailendra Kulkarni. 2024. A Comprehensive Study of Agile Project Management Tools. In *2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM)*. 1–6. https://doi.org/10.1109/ICIPTM59628.2024.10563875

[5] Aneesha Bakharia, Kirsty Kitto, Abelardo Pardo, Dragan Gašević, and Shane Dawson. 2016. Recipe for Success: Lessons Learnt from Using xAPI within the Connected Learning Analytics Toolkit. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (LAK '16)*. 378–382. http://dl.acm.org/citation.cfm?doid=2883851.2883882

[6] Robert Bodily and Katrien Verbert. 2017. Review of Research on Student-Facing Learning Analytics Dashboards and Educational Recommender Systems. *IEEE Transactions on Learning Technologies* 10, 4 (2017), 405–418.

[7] Curtis J. Bonk, Mimi Miyoung Lee, Xiaojing Kou, Shuya Xu, and Feng-Ru Sheu. 2015. Understanding the self-directed online learning preferences, goals, achievements, and challenges of MIT OpenCourseWare subscribers. *Journal of Educational Technology & Society* 18, 2 (2015), 349–368.

[8] Joel Brandt, Mira Dontcheva, Marcos Weskamp, and Scott R. Klemmer. 2010. Example-centric programming: integrating web search into the development environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Atlanta Georgia USA, 513–522. https://doi.org/10.1145/1753326.1753402

[9] Scott Brave and Cliff Nass. 2007. Emotion in Human-Computer Interaction. In *The Human-Computer Interaction Handbook*. CRC Press, 103–118.

[10] Carrie J. Cai and Philip J. Guo. 2019. Software Developers Learning Machine Learning: Motivations, Hurdles, and Desires. In *VL/HCC*. 25–34.

[11] Rimika Chaudhury and Parmit K. Chilana. 2025. Designing Visual and Interactive Self-Monitoring Interventions to Facilitate Learning: Insights from Informal Learners and Experts. *IEEE Transactions on Visualization and Computer Graphics* 31, 2 (2025), 1542–1556. https://doi.org/10.1109/TVCG.2024.3366469

[12] Rimika Chaudhury, Philip J. Guo, and Parmit K. Chilana. 2022. "There's no way to keep up!": Diverse Motivations and Challenges Faced by Informal Learners of ML. In *2022 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 1–11. https://doi.org/10.1109/VL/HCC53370.2022.9833100

[13] Rimika Chaudhury, Taha Liaqat, and Parmit Chilana. 2023. Exploring the needs of informal learners of computational skills: Probe-based elicitation for the design of self-monitoring interventions. In *Graphics Interface 2023*.

[14] Li Chen, Min Lu, Yoshiko Goda, Atsushi Shimada, and Masanori Yamada. 2021. Learning Analytics Dashboard Supporting Metacognition. In *Balancing the Tension between Digital Technologies and Learning Sciences*, Dirk Ifenthaler, Demetrios G. Sampson, and Pedro Isaías (Eds.). Cham, 129–149. https://doi.org/10.1007/978-3-030-65657-7_8

[15] Parmit K. Chilana, Rishabh Singh, and Philip J. Guo. 2016. Understanding Conversational Programmers: A Perspective from the Software Industry. In *CHI Conference on Human Factors in Computing Systems*. ACM, 1462–1472. https://dl.acm.org/doi/10.1145/2858036.2858323

[16] Sue Inn Chng. 2018. Incorporating Reflection into Computing Classes: Models and Challenges. *Reflective Practice* 19, 3 (2018), 358–375.

[17] Janghee Cho, Tian Xu, Abigail Zimmermann-Niefield, and Stephen Voida. 2022. Reflection in Theory and Reflection in Practice: An Exploration of the Gaps in Reflection Support among Personal Informatics Apps. In *CHI Conference on Human Factors in Computing Systems*. 1–23. https://doi.org/10.1145/3491102.3501991

[18] Eun Kyoung Choe, Saeed Abdullah, Mashfiqui Rabbi, Edison Thomaz, Daniel A. Epstein, Felicia Cordeiro, Matthew Kay, Gregory D. Abowd, Tanzeem Choudhury, James Fogarty, Bongshin Lee, Mark Matthews, and Julie A. Kientz. 2017. Semi-Automated Tracking: A Balanced Approach for Self-Monitoring Applications. *IEEE Pervasive Computing* 16, 1 (jan 2017), 74–84. https://doi.org/10.1109/MPRV.2017.18

[19] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. 2017. Understanding Self-reflection: How People Reflect on Personal Data through Visual Data Exploration. In *Proceedings of the 2017 Pervasive Computing Technologies for Healthcare (PervasiveHealth '17)*. ACM, New York, NY, USA, 173–182.

[20] Eun Kyoung Choe, Nicole B Lee, Bongshin Lee, Wanda Pratt, and Julie A Kientz. 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1143–1152.

[21] Kevin Chow, Thomas Fritz, Liisa Holsti, Skye Barbic, and Joanna McGrenere. 2023. Feeling Stressed and Unproductive? A Field Evaluation of a Therapy-Inspired Digital Intervention for Knowledge Workers. *ACM Transactions on Computer-Human Interaction* 31, 1 (2023), 1–33.

[22] Francesco Cirillo. 2018. *The Pomodoro Technique: The Acclaimed Time-Management System That Has Transformed How We Work*. Currency.

[23] Kathryn Cunningham, Barbara J. Ericson, Rahul Agrawal Bejarano, and Mark Guzdial. 2021. Avoiding the Turing Tarpit: Learning Conversational Programming by Starting from Code's Purpose. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (May 2021), 1–15. https://doi.org/10.1145/3411764.3445571

[24] Brian Dorn and Mark Guzdial. 2010. Learning on the Job: Characterizing the Programming Knowledge and Learning Strategies of Web Designers. In *CHI Conference on Human Factors in Computing Systems*. 703. http://portal.acm.org/citation.cfm?doid=1753326.1753430

[25] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A Lived Informatics Model of Personal Informatics. In *UbiComp*. 731–742. http://dl.acm.org/citation.cfm?doid=2750858.2804250

[26] Clayton Feustel, Shyamak Aggarwal, Bongshin Lee, and Lauren Wilcox. 2018. People Like Me: Designing for Reflection on Aggregate Cohort Data in Personal Informatics Systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)* (2018). https://www.microsoft.com/en-us/research/publication/people-like-me-designing-for-reflection-on-aggregate-cohort-data-in-personal-informatics-systems/ Full text available: https://www.microsoft.com/en-us/research/publication/people-like-me-designing-for-reflection-on-aggregate-cohort-data-in-personal-informatics-systems/.

[27] Rowanne Fleck and Geraldine Fitzpatrick. 2010. Reflecting on reflection: framing a design landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*. 216–223.

[28] Joe Foley. 2024. Ability to work with AI identified as "top skill of the future" for Creatives. https://www.creativebloq.com/news/autodesk-ai-creative-skills-study

[29] Gao Gao, Finn Voichick, Michelle Ichinco, and Caitlin Kelleher. 2020. Exploring programmers' api learning processes: Collecting web resources as external memory. In *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, 1–10.

[30] Graham Glanville and Brett A Becker. 2016. Reflective Learning Journals in Computer Science: The Student Experience. (2016).

[31] Carl Christopher Haynes-Magyar and Nathaniel James Haynes-Magyar. 2022. Codespec: A Computer Programming Practice Environment. *Proceedings of the 2022 ACM Conference on International Computing Education Research - Volume 2* (August 2022), 32–34. https://doi.org/10.1145/3501709.3544278

[32] Donghan Hu and Sang Won Lee. 2022. Scrapbook: Screenshot-Based Bookmarks for Effective Digital Resource Curation across Applications. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. ACM, 1–13.

[33] Donghan Hu and Sang Won Lee. 2024. Exploring the Effectiveness of Time-lapse Screen Recording for Self-Reflection in Work Context. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 1–14.

[34] Matthew Jörke, Yasaman S Sefidgar, Talie Massachi, Jina Suh, and Gonzalo Ramos. 2023. Pearl: A Technology Probe for Machine-Assisted Reflection on Personal Data. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. 902–918.

[35] Christina Kelly, Bongshin Lee, and Lauren Wilcox. 2017. Self-tracking for Mental Wellness: Understanding Expert Perspectives and Student Experiences. *CHI Conference on Human Factors in Computing Systems* (Feb. 2017). https://www.microsoft.com/en-us/research/publication/self-tracking-mental-wellness-understanding-expert-perspectives-student-experiences/

[36] Kimia Kiani, Parmit K. Chilana, Andrea Bunt, Tovi Grossman, and George Fitzmaurice. 2020. "I Would Just Ask Someone": Learning Feature-Rich Design Software in the Modern Workplace. In *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 1–10. https://doi.org/10.1109/VL/HCC50065.2020.9127288

[37] Young-Ho Kim, Jae Ho Jeon, Eun Kyoung Choe, Bongshin Lee, Kwonhyun Kim, and Jinwook Seo. 2016. TimeAware: Leveraging Framing Effects to Enhance Personal Productivity. *CHI Conference on Human Factors in Computing Systems* (2016). https://www.microsoft.com/en-us/research/publication/timeaware-leveraging-framing-effects-enhance-personal-productivity/

[38] Young-Ho Kim, Jae Ho Jeon, Bongshin Lee, Eun Kyoung Choe, and Jinwook Seo. 2017. OmniTrack: A Flexible Self-Tracking Approach Leveraging Semi-Automated Tracking. *Interactive, Mobile, Wearable and Ubiquitous Technologies* (jul 2017). https://www.microsoft.com/en-us/research/publication/omnitrack-flexible-self-tracking-approach-leveraging-semi-automated-tracking/

[39] Young-Ho Kim, Bongshin Lee, Arjun Srinivasan, and Eun Kyoung Choe. 2021. Data@ Hand: Fostering Visual Exploration of Personal Data on Smartphones Leveraging Speech and Touch Interaction. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17. https://doi.org/10.1145/3411764.3445421

[40] Malcolm S. Knowles. 1975. Self-directed learning: a guide for learners and teachers. *ERIC* (1975).

[41] Brian Krisler and Richard Alterman. 2008. Training towards mastery: overcoming the active user paradox. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*. ACM, 239–248.

[42] Harsh Kumar, Ruiwei Xiao, Benjamin Lawson, Ilya Musabirov, Jiakai Shi, Xinyuan Wang, Huayin Luo, Joseph Jay Williams, Anna N Rafferty, John Stamper, et al. 2024. Supporting Self-Reflection at Scale with Large Language Models: Insights from Randomized Field Experiments in Classrooms. In *Proceedings of the Eleventh ACM Conference on Learning@ Scale*. 86–97.

[43] Youn kyung Lim, Justin Donaldson, Heekyoung Jung, Breanne Kunz, David Royer, Shruti Ramalingam, Sindhia Thirumaran, and Erik Stolterman. 2008. Emotional Experience and Interaction Design. *Affect and Emotion in Human-Computer Interaction: From Theory to Applications* (2008), 116–129.

[44] Thomas D LaToza, Maryam Arab, Dastyni Loksa, and Amy J Ko. 2020. Explicit Programming Strategies. *Empirical Software Engineering* 25, 4 (2020), 2416–2449.

[45] LeetCode. 2024. LeetCode: Online Coding Platform. https://leetcode.com Accessed: September 12, 2024.

[46] Dastyni Loksa, Amy J. Ko, Will Jernigan, Alannah Oleson, Christopher J. Mendez, and Margaret M. Burnett. 2016. Programming, Problem Solving, and Self-Awareness: Effects of Explicit Guidance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 1449–1461.

[47] Dastyni Loksa, Lauren Margulieux, Brett A. Becker, Michelle Craig, Paul Denny, Raymond Pettit, and James Prather. 2022. Metacognition and Self-Regulation in Programming Education: Theories and Exemplars of Use. *ACM Transactions on Computing Education* (Feb. 2022), 3487050. https://doi.org/10.1145/3487050

[48] Yuhan Luo, Young-Ho Kim, Bongshin Lee, Naeemul Hassan, and Eun Kyoung Choe. 2021. FoodScrap: Promoting Rich Data Capture and Reflective Food Journaling Through Speech Input. *DIS '21* (May 2021). https://doi.org/10.1145/3461778.3462074

[49] Yuhan Luo, Bongshin Lee, Donghee Yvette Wohn, Amanda L Rebar, David E Conroy, and Eun Kyoung Choe. 2018. Time for break: Understanding information workers' sedentary behavior through a break prompting system. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 1–14.

[50] Louise Ann Lyon, Chelsea Clayton, and Emilv Green. 2018. Tinkering in the Wild: What Leads to Success for Female End-User Programmers? *2018 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)* (2018), 125–129.

[51] Janice Malcolm, Phil Hodkinson, and Helen Colley. 2003. *Informality and formality in learning: a report for the Learning and Skills Research Centre*. Learning and Skills Research Centre.

[52] Robert McCartney, Jonas Boustedt, Anna Eckerdal, Kate Sanders, Lynda Thomas, and Carol Zander. 2016. Why computing students learn on their own: Motivation for self-directed learning of computing. *ACM Transactions on Computing Education (TOCE)* 16, 1 (2016), 1–18.

[53] André N. Meyer, Thomas Fritz, and Thomas Zimmermann. 2019. Fitbit for developers: self-monitoring at work. *Rethinking Productivity in Software Engineering* (2019), 261–270.

[54] Andre N. Meyer, Gail C. Murphy, Thomas Zimmermann, and Thomas Fritz. 2017. Design Recommendations for Self-Monitoring in the Workplace: Studies in Software Development. *ACM Human-Computer Interaction* 1, CSCW (2017), 1–24. https://dl.acm.org/doi/10.1145/3134714

[55] OpenAI. 2024. OpenAI API Documentation Overview. https://platform.openai.com/docs/overview. [Accessed 12-09-2024].

[56] Joonyoung Park, Hyunsoo Lee, Sangkeun Park, Kyong-Mee Chung, and Uichin Lee. 2021. Goldentime: Exploring System-driven Timeboxing and Micro-financial Incentives for Self-regulated Phone Use. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 1–17. https://doi.org/10.1145/3410400.3415691

[57] W Jared Parmer. 2022. Manipulative design through gamification. In *The Philosophy of Online Manipulation*. Routledge, 216–234.

[58] James Prather, Brett A. Becker, Michelle Craig, Paul Denny, Dastyni Loksa, and Lauren Margulieux. 2020. What Do We Think We Think We Are Doing? Metacognition and Self-Regulation in Programming. In *Proceedings of the International Computing Education Research Conference*. 2–13.

[59] Kathleen B Rager. 2009. I Feel, Therefore, I Learn: The Role of Emotion in Self-Directed Learning. *New Horizons in Adult Education and Human Resource Development* 23, 2 (2009), 22–33.

[60] Damian J Rivers. 2020. Strava as a discursive field of practice: Technological affordances and mediated cycling motivations. *Discourse, Context & Media* 34 (2020), 100345.

[61] C. Scaffidi, M. Shaw, and B. Myers. 2005. Estimating the Numbers of End Users and End User Programmers. *2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)* (2005), 207–214. https://doi.org/10.1109/VLHCC.2005.34

[62] Donald A. Schön. 2017. *The reflective practitioner: How professionals think in action.* Routledge.

[63] Huma Shafiq, Zahid Ashraf Wani, Iram Mukhtar Mahajan, and Uzma Qadri. 2017. Courses beyond borders: A case study of MOOC platform Coursera. *Library Philosophy and Practice* (2017), 1–15.

[64] Anselm Strauss and Juliet Corbin. 1994. Grounded theory methodology: An overview. (1994).

[65] Eliane Zambon Victorelli and Julio Cesar Dos Reis. 2020. Human-Data Interaction Design Guidelines for Visualization Systems. In *Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems (IHC '20).* ACM, New York, NY, USA, 1–10.

[66] April Y. Wang, Ryan Mitts, Philip J. Guo, and Parmit K. Chilana. 2018. Mismatch of Expectations: How Modern Learning Resources Fail Conversational Programmers. In *CHI Conference on Human Factors in Computing Systems.* 1–13. https://doi.org/10.1145/3173574.3174085

[67] Tianjia Wang, Daniel Vargas Díaz, Chris Brown, and Yan Chen. 2023. Exploring the Role of AI Assistants in Computer Science Education: Methods, Implications, and Instructor Perspectives. In *2023 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC).* IEEE, 92–102.

[68] Philip H. Winne. 2010. Improving Measurements of Self-Regulated Learning. *Educational Psychologist* 45, 4 (Oct. 2010), 267–276. http://www.tandfonline.com/doi/abs/10.1080/00461520.2010.517150

[69] David Wong-Aitken, Diana Cukierman, and Parmit K Chilana. 2022. 'It Depends on Whether or Not I'm Lucky': How Students in an Introductory Programming Course Discover, Select, and Assess the Utility of Web-Based Resources. In *Proceedings of the 27th ACM Conference on Innovation and Technology in Computer Science Education Vol. 1.* 512–518.

[70] Saelyne Yang, Jisu Yim, Aitolkyn Baigutanova, Seoyoung Kim, Minsuk Chang, and Juho Kim. 2022. SoftVideo: Improving the Learning Experience of Software Tutorial Videos with Collective Interaction Data. In *Proceedings of the 27th International Conference on Intelligent User Interfaces.* 646–660.