

Table of Contents

Introduction.	3
What Is a Digital Learning Platform?	5
What Is CLICKSTREAM and What Sets It Apart from Other Digital Learning Platforms?	5
Opportunities for Instructional Practice: Providing Data to Inform Instructional Decisions	10
Opportunities for Educational Research	12
Scaling CLICKSTREAM as a DLP for the Promotion and Study of Learning.	13
What Does CLICKSTREAM Require of Instructors?	14
How Does Ongoing Development and Incorporation of AI Lighten the Load on Instructors?	14
Building a Community around the Promotion and Study of Learning on CLICKSTREAM	15
References	17

Introduction

Those who propose theories about learning set out to observe individuals and the ways they learn and richly describe the way learning unfolds (Greene, 2022). Theories of learning are effectively stories that position the learner as the protagonist and document the learner's journey as they pursue their learning goals. In order to fully frame this endeavor as a hero's quest, theorists must further describe the origin story of the learner, the experiences that formed them, the qualities those experiences imbued in them (Ocumpaugh et al., 2024), and the challenges that learners are likely to encounter as a result of experiences in their past. Educators, researchers, and developers are keen on describing the *learning journey* (Wilder et al., 2018) and produce detailed accounts of individuals' learning experiences using methods like journey maps (deLaRosa, 2025). These methods rely upon rich, idiosyncratic accounts from learners who share details about the key moments that were essential to their learning experiences and influential to their outcomes. Administrators and educators care deeply about their learners, wish to know their stories, and endeavor to provide each learner with the aid they need along their quest. They are also hard-pressed to do so, as each student is the protagonist of their own story, but every course is the setting for hundreds of learners' stories, and telling all these stories in real time so that aid can be provided is a grand challenge.

Much can be learned from individual journeys. They can provide qualitative insights about the individual experience and serve as powerful examples of theorized phenomena. However, qualitative accounts are not meant to provide opportunities for generalization or empirical testing of educational theories. Inferentially testing theoretical assumptions requires quantitative research approaches that sample learners from populations, collect measures that approximate key elements of the learning experience, and require inferencing about how these data reflect key attributes of the learner, their actions, the context in which the learning takes place, and how these interact to produce learning outcomes. Just as each learner's journey is idiosyncratic, so is the selection of samples, instruments, and learning contexts that are documented in empirical, quantitative research. This makes the task of synthesis a challenge. A reasonable way to describe the synthesis of the empirical research on learning is that researchers conduct individual investigations to uncover the stories of individual learners and samples, and that efforts to reconcile across studies are attempts to get their stories straight.

Digital learning platforms (DLPs) serve as a potential resource for telling rich stories about the learning process. When DLPs can capture details about the learner, their learning process, and the context in which it unfolds, learners' stories can be told in ways that hold constant or systematically vary features of the learner, the context, or the process so that empirical tests of an educational theory can be reconciled. Achieving this requires establishing data platforms that can collect data that are essential for describing learners, process, and context in detail, along with tools to enrich those data with the setting and plot established by an institution, course, and learning task. The characters include the learner and learning resources they encounter along the way. When a DLP can host the raw data that establish these story elements, the elements can be enriched to form the story and shared with instructors who can follow stories and adaptively support the learners whose journeys they describe. Thereafter, multiple theoretical lenses can be applied to tell the story that unfolds through a distinct narrator's viewpoint as learners engage in the world their instructors have built and journey toward the conclusions they ultimately reach.

In this white paper, we introduce *CLICKSTREAM* (Cyberinfrastructure for Learning Insights & Curricular Knowledge using Student Trace-data for Research in Educational Analytics & Measurement) as a digital learning platform where these kinds of stories can be told. We document the construction that instructors undertake when they design courses, objectives, assessments, and assets that they provide to learners as resources. We document how co-designing or co-documenting courses with the instructors who build them can produce a rich schema that describes the features of courses that need to be understood to appreciate a learner's activities as they pursue their learning goals. We first show how these enriched data can be shared with the instructors and learners who produce it. We demonstrate how the data can inform data-driven instructional adaptations as educators respond to course-wide patterns of engagement, as well as student-specific reports of engagement, learning, and likelihood of success. We do so to document how course-specific data modeling can be used to improve educational practice. Then we widen our lens to show how the same platform provides a base on which data can be further enriched with theoretical tagging to power scaled analyses of complex assumptions from educational theories and refine the ways they describe learning (Gibson, 2017). We highlight how theory-aligned feature engineering of *CLICKSTREAM* data can be applied within a single course to focus deeply on complex learning processes in a single context to test explicit theoretical assumptions, as well as how the shared, fixed data schema can be applied across contexts to investigate how learning works for different types of individual learners and under differing instructional and institutional conditions. Both of these are necessary approaches to leverage the potential of big data for education to inform the refinement of learning theories (Lodge et al., 2017). Finally, we consider the scalability of the *CLICKSTREAM* approach and document the intensity of the tagging process and how incorporating artificial intelligence can lighten the burden on educators who wish to contribute their course data and learn from those data so they can provide adaptive instruction to learners during the semester and improved course designs in the next.

What Is a Digital Learning Platform?

When students interact with technologies while learning, every click, tap, keystroke, or other interaction leaves a trace on the technology's server log. Those interactions can be instrumented to capture the behavior of the learner who used the technology, the asset they used, the action they took with it, the context in which the event took place, and the time at which it occurred. These raw features of event data have long been considered key information about user activity and have been established as essential data required for learning technologies to be interoperable (e.g., caliper standards; [1EdTech, 2025](#)). When these data are collected from any learning technology, educators, developers, and researchers can begin to build descriptions of learners, assets, course contexts, and timestamps so that they can be understood in the context of the learning task and as representative of learning events (Bernacki, 2018). Because educators endeavor to provide active learning experiences for students (Chi & Wylie, 2014; Eddy & Hogan, 2014), they often draw upon multiple technologies to provide active, constructive, and interactive learning opportunities, which they host in a common space like a learning management system (LMS), where they use tools for integrating technologies. When learners' engagement in the LMS and the tools it hosts can be observed via digital learning platforms, educators can collect and model those data in order to observe the learner's engagement and act on those observations during instruction, while researchers can study the learning process.

What Is CLICKSTREAM and What Sets It Apart from Other Digital Learning Platforms?

CLICKSTREAM is a cloud-hosted service designed to collect, enrich, and analyze students' timestamped digital interaction data from multiple learning platforms with instructional and theory-aligned tagging schema (Figure 1). Those schemas are the result of a collaborative development process that centers the instructor's voice through co-design and the analysis of key instructional documents (e.g., syllabi and course schedules). This process ensures that every resource and its role within an authentic instructional design is understood and established in the data.

Figure 1.

Conceptual diagram of educational technology applications as sources of data operated upon by CLICKSTREAM (cloud icon), including the intake of interaction data, instrumentation and enrichment with metadata from course and institutional sources, and tagging schema from instructors and educational theorists. Those enriched and interpreted data can tell the stories of individuals and groups to educators, who can responsively adapt instruction, or researchers, who can use data to test and refine assumptions from educational theory.



CLICKSTREAM builds on 1EdTech’s Caliper-formatted data (i.e., key fields capturing learner, learning object, context, time, and more) by adding details about the task and the way the event reflects engagement in it. Unlike prior attempts to scale interoperable data access for closed systems of users (e.g., UNIZEN, which serves only institutions in one higher education network), CLICKSTREAM is built for broad adoption by instructors and adaptability by researchers who can further enrich the tagging infrastructure to more fully describe learning through multiple theoretical lenses.

Illustrative Example: Instructional Design and Application to Learner Engagement in CLICKSTREAM

Consider the example case of an introductory physics course at a large public American university. The instructor who designed the introductory physics class provided the details captured in the following table (Figure 2), which documents the kinds of learning resources students can encounter in the class and the intended uses of those resources as assets to enact learning processes.

Figure 2.

Example assets reflecting digital resources that encourage learning processes in an introductory physics course.

Instructional Intent	Definition	Examples
Course Administration	Documents and tools that provide structural organization and course policies.	📄✅🌐📅 Syllabus
Course Navigation	Viewing of course site pages and external links.	👤💡🌐📄 Mastering Link
Instructional Content Materials	Core instructional materials that provide foundational knowledge on course topics.	👤✅🌐📄 Assigned Readings 👤💡🌐📄 Zoom Recordings
Instructional Content Scaffolds	Resources designed to guide students in understanding and organizing new information.	📄✅🌐📄 Guided Reading Questions 📄🌟📄📄 Class Outlines
Active & Accountable Learning	Activities intended to encourage engagement where students apply and explore concepts interactively.	📄👤🌐📄 Guided Reading Questions 📄👤🌐📄 iClicker Questions
Collaboration & Peer Learning	Opportunities for students to engage in peer-supported learning experiences.	📄👤🌐📄 Small-group activities
Retrieval Practice & Mastery Learning	Strategies that enhance retention and recall of key concepts through structured practice.	📄🌟📄📄 Lecture Slides 📄🌟📄📄 Previous Semester Exam Forms
Reflection & Metacognition	Resources and activities that encourage students to reflect on their learning process.	📄👤🌐📄 Learning Journal Entries
Formative Assessment	Low-stakes assessments providing students with feedback on their understanding.	📄👤🌐📄 Prior Knowledge Assessment 📄👤🌐📄 Online Practice Exams
Summative Assessment	High-stakes evaluations used to measure students’ mastery of the course material.	📄👤🌐📄 Midterm Exam
Content Specific Supplemental Instruction	Supplemental instruction of course specific content provided by the learning center, teaching-assistants, peer-mentors, etc.	👤💡🌐📄 End of Week Review Session
General Campus Support & Resources	University-based services to develop students’ academic writing, and research.	👤💡🌐📄 Learning Center Link

Instructors offer these resources as fixed assets and recommend how students should use them at specific moments during the course. For example, a chapter outline can first serve as a preview before a lecture, an organizer for taking notes during the lecture, or a review resource when preparing for an exam. When usage data is logged in CLICKSTREAM, the time at which a learner chooses to use the resource is recorded.

This resource + time combination establishes the temporal relationship of access of assets to the timing of the course assessments that will produce a record of their learning and achievement. To capture all these temporal and contextual details requires a more thorough mapping of the course’s instructional design. This can be observed in Figures 3 and 4, which document the same resources in greater detail, expanding on the iconography in Figure 2 to more extensively expose the instructional design dimensions of the same physics course that adheres to an active learning design.

Figure 3.

An introductory physics course’s learning resources as they are designed to be used before, during, or after classes in a typical learning cycle, and resources that function in non-cyclical ways.

	 Non-Cyclical	 Before Class	 During Class	 After Class
Course Administration	Schedule Syllabus Seating charts			
Course Navigation	Book/E-book (OpenStax) Gradescope (Link to platform) Piazza forum			
Instructional Content Materials	Terms	Lesson Readings		
Instructional Content Scaffolds		Additional Readings	Lecture Slides	
Active & Accountable Learning			In-Class Activities (Poll Everywhere) Complementary Material	Homeworks (ExperTa)
Collaboration & Peer Learning			Studio Worksheets	Studio Submission
Retrieval Practice & Mastery Learning				Studio Worksheet Keys Post-Lecture Slides
Reflection & Metacognition		Practice Exam Keys		Midterms Exams Form Midterms Exams Form Keys
Formative Assessment		Warm-up assignments Practice Exams		
Summative Assessment			Midterms Exams Final Exam	
Content Specific Supplemental Instruction			Q&A End of Week Sessions	
General Campus Support & Resources				

Cycle Types

- Non-Cyclical
- Onboarding
- Content Coverage
- Lab/Studio
- Weekly
- Examinations

Active learning designs often include highly structured *learning cycles* (Berro et al., 2025) where each lesson involves some pre-class activity, such as pre-reading, formative assessment, and homework, and preparations including the download of course outlines and notes. Additional activities like formative assessments happen in class, and learners follow up a class session by revisiting newly hosted notes or recordings of session activities that can inform their ongoing learning process and preparations for major exams. This cycle repeats lesson by lesson within each content unit, as does a unit-by-unit content coverage and assessment cycle. Data can be modeled across multiple cycles of lessons, chapters, or content units, and learner engagement can be modeled over the full span of the course, from orientation on the first day of class (FDOC; Figure 4), through instruction and during periods of preparation for summative exams. Additional details can be gleaned from interoperable, enriched data when the timing of a learning event and the action undertaken with a learning resource can be interpreted based on the instructional and temporal mapping of the event’s occurrence to the course objectives, content, and schedule (Figures 4 and 5), and how the event’s timing in relation to the course schedule reflects its alignment to the anticipated use of the learning resource (Figure 6).

Figure 4.

Additional metadata about the opportunities learning resources provide for learning activity, along with alignment to a cycle and to an assessment.

	Act	Exp	Cyc	Pha	Common Name (Instructor's nomenclature)
FDOC	↓	✓	●	🕒	Syllabus
	↑	📄	●	🕒	Syllabus Quiz
	👁️	★	●	🕒	Pearson Account Creation Instructions
	↑	📄	●	▶️	Prior Knowledge Assessment
Before Class	↓	✓	●	🕒	Guided Reading Questions
	👁️	✓	●	🕒	Assigned Readings
	👁️	💡	●	🕒	Mastering Link
	↑	📄	●	🕒	Mastering Homework Assignments
	↑	📄	●	🕒	Guided Reading Questions
In-Class	↓	★	●	▶️	Class Outlines
	↑	📄	●	▶️	iClicker Questions
After Class	↓	★	●	🔄	Lecture Slides
	👁️	💡	●	🔄	Zoom Recordings
Add. Cycles	↓	✓	●	🕒	Lab Instructions
	↑	📄	●	▶️	Lab Worksheet Assignments
	👁️	💡	●	🔄	End of Week Review Session
Exam Preparation	↓	★	●	🕒	Exam Study Guides
	↑	📄	●	🕒	Online Practice Exams
	↓	★	●	🕒	Previous Semester Exam Forms
	↓	★	●	🕒	Previous Semester Exam Keys
Admin. & Feedback	↑	📄	●	▶️	Midterm Exam
	↓	💡	●	🔄	Current Semester Exam Keys

Cycle Types

- Non-Cyclical
- Onboarding
- Content Coverage
- Lab/Studio
- Weekly
- Examinations

Cycle Phases

- 🕒 Non-Cyclical
- 🕒 Before Class
- ▶️ During Class
- 🔄 After Class

Potential Actions

- ↓ Download
- ↑ Submit/Create
- 👁️ View/Navigate

Level of Expectation

- 📄 Required (Graded)
- ✓ Expected
- ★ Recommended
- 💡 Optional

Figure 5.

Elaborated mapping of learning resources to course design schema by the cycle(s) on which they can be observed, stakes of the activity, and actions the resources enable.

Cycle Phases

- 🕒 Non-Cyclical: Resources available throughout the course, not tied to phases.
- 🕒 Before Class: Preparation activities completed prior to scheduled class sessions.
- ▶️ During Class: Engagement in instructional activities during class time.
- 🔄 After Class: Review and reinforcement activities following class sessions.

Cycle Types

- Non-Cyclical: Resources and activities that are not related to a specific class meeting or topic.
- Onboarding (FDOC): Introductory activities to familiarize students with course expectations.
- Content Coverage: Instructional sessions focused on delivering and exploring course material.
- Lab/Studio: Hands-on, practical sessions for applying concepts in real-world scenarios.
- Weekly: Recurring sessions/assignments that maintain course rhythm and reinforce learning.
- Examinations: High-stakes, summative assessments.

Potential Actions

- ↓ Download (e.g., .doc, .ppt, .pdf)
- ↑ Submit/Create (e.g., assignments, quizzes, discussion posts)
- 👁️ View/Navigate (e.g., course pages, discussion pages, external links)

Level of Expectation

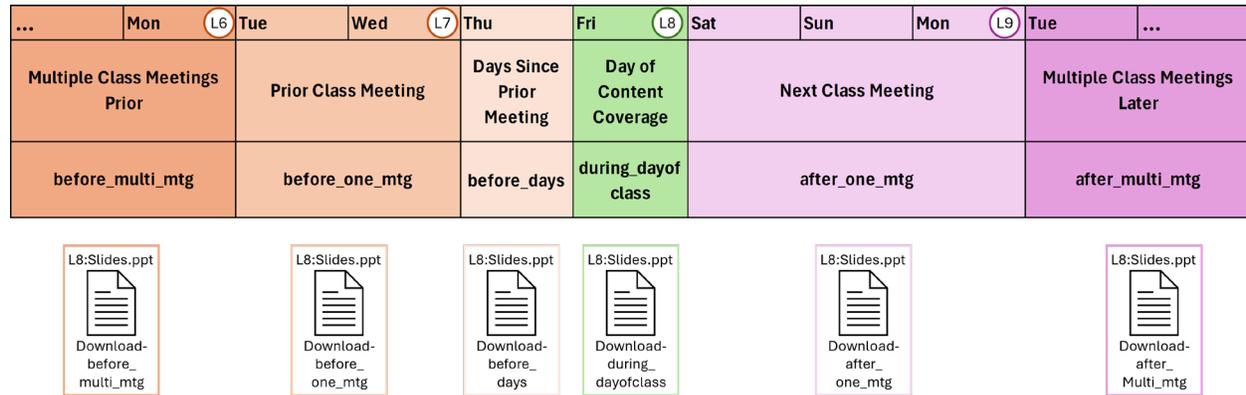
- 📄 Required (Graded): Required and assessed for credit.
- ✓ Expected: Mandatory but not graded (e.g., pre-class readings, participation activities).
- ★ Recommended: Strongly encouraged but not required.
- 💡 Optional: Extra/enrichment material for deeper learning or review.

Stakes: Classifier or Continuous Variable?

- **None:** Not graded
- **Low:** Participation, binary result
- **Medium:** Accuracy, small %
- **High:** Accuracy, large %

Figure 6.

Alignment of the timing of a slide deck for a lesson (L8 Slides.ppt) and whether the day on which it is accessed is multiple days, before, on the day of, or after the season when the lesson is covered in the course. This temporal coding of resources to the schedule enables interpretation of asset access events and inferences about the resource’s affordances for learning when accessed early, on time, or later than the target date.



These visualizations capture the underlying intake of interoperable data that is the typical product of learning technologies and raw currency of DLPs, as well as the tagging infrastructure provided by CLICKSTREAM that enriches educators’ and educational theorists’ ability to describe learning (i.e., the second aim of educational theories; Greene, 2022) as a detailed story (i.e., a record of learning events, Bernacki, 2018) given meaning by a course context that provides setting and plot. This allows educators and theorists to understand learners’ quests and model their learning journeys to better know how to empathize with and support individual learners (Tracey & Baaki, 2022). These data also enable observation of the many divergent pathways individuals take over one or more courses along their STEM degree pathway (Rosenzweig et al., 2021) that comprise the story of individuals and groups of learners. Educators and educational theorists can use this information to describe learning journeys through the narrative lens of their choosing, by aligning it to a theory of instruction or learning (Bernacki, 2025).

We next consider the implications of CLICKSTREAM for educational practice and the development of educational theory. We begin by highlighting opportunities for data storytelling in higher education and the ways that instructors can derive insights about learning processes in their instructional environments and how they could adapt their instructional design and delivery to improve the learner experience. These include a long view of CLICKSTREAM data that inform design-loop adaptation of courses through iterative improvement of course designs (Aleven et al., 2017), and dashboards of student preparedness to learn that can improve instructor responsiveness within a learning cycle by altering lecture plans in light of students’ preparation for the lesson (i.e., step-loop adaptivity; Aleven et al., 2017). We demonstrate how this kind of mapping affords opportunities to observe learning events and apply schema to new data that accrue in a timely enough fashion that they can inform instructors’ decision-making.

Opportunities for Instructional Practice: Providing Data to Inform Instructional Decisions

Through instructor tagging and temporal mapping of interactions to course timelines and instructional design tags for each resource, CLICKSTREAM enables instructors to see what students do and when they do it, using familiar terminology they or their peers have provided. This enables instructors to make inferences about the learning process and ask questions of the data, with an eye toward improving instructional decisions during classes, refining tools to identify and support struggling students, and producing summaries of learners' behaviors and performances, all of which can inform course redesign. All of these are pressing concerns given students' struggle to manage the demands of active learning (Shekhar et al., 2020) and succeed (Freeman et al., 2014).

Opportunity #1: Telling an Instructor the Semester-Long Story about Students' Learning

An example of the way tagged data in CLICKSTREAM can be provided back to instructors appears in Figure 7. This chart provides a semester-long view of learners' engagement with resources and the trends in the timeliness of student activity against the instructors' expectations. When this dashboard is visualized to include a semester-long corpus of data, an instructor can make observations about the way learning unfolded during a course and consider ways to adapt their instructional designs from one semester to the next (i.e., design-loop adaptivity, Alevin et al., 2017). For instance, where student-user behavior is delayed compared to the recommended timing of activity (e.g., when students download slides just in time for, during, or right after lecture on the day of class; middle magenta box), an instructor can provide pre-training about the importance of using advanced organizers. This is a key adaptive learning design approach (Plass & Pawar, 2020) that leverages multimedia learning theory (Pre-Training Principle, Mayer & Pilegard, 2005).

Opportunity #2: Appraising Students' Preparedness to Learn, in Time for Instructors to Adapt

When the same kind of visualization is filtered down to reflect a single lesson, and raw data are flowed into CLICKSTREAM and enriched in real time, an instructor can use this same view to observe the percentage of students who have completed the pre-class activities assigned in a learning cycle (e.g., read the chapter, completed reading comprehension activities and homework, and downloaded the course outline and slides). When the majority of learners are fully prepared for a lesson, instructors can simply conduct business as usual and engage in the active, constructive, and interactive elements of their lesson that are typical of active learning designs. On other days, where calendars are disrupted, workloads have spiked, or other events might threaten students' opportunity to prepare for a session, an instructor may find that far fewer students appear to have completed the work necessary to form and pose curiosity questions and questions to extend students thinking or productively engage in collaborative activities. The dashboard could signal this to an instructor in the hours before a lesson, and they could quickly pivot to the less engaging but more adaptive didactic lecturing that students may need to acquire basic content knowledge for that lesson.

Figure 7.

CLICKSTREAM dashboard prototype visualizing learner engagement timing.

Timing of Student Interactions

Lecture Meeting Materials	Meetings Before Class		Days Before Class	Day of Class	Meetings After Class	
	Multiple	One			One	Multiple
Before Meeting						
Instructional Content Materials	Timely Engagement				Delayed Engagement	
Lesson Reading <i>download</i>	13.19%	19.44%	38.54%	15.28%	5.56%	7.99%
Additional Reading <i>download</i>	1.15%	5.56%	10.15%	11.49%	16.67%	54.98%
Summative Assessment						
Warm Up Assignment <i>accessed</i>	10.92%	22.13%	45.30%	21.17%	0.30%	0.17%
<i>download</i>	2.94%	4.71%	57.06%	33.82%	0.00%	1.47%
During Meeting						
Instructional Content Scaffolds	Advanced Engagement					
Pre-Lecture Slides <i>download</i>	0.00%	0.07%	6.31%	62.49%	14.62%	16.52%
After Meeting						
Active & Accountable Learning						
Homework <i>accessed</i>	0.52%	0.39%	0.80%	3.25%	8.30%	86.74%
Retrieval Practice & Mastery Learning						
Post-Lecture Slides <i>download</i>	0.00%	0.00%	0.00%	8.38%	23.67%	67.95%
Ponderables <i>download</i>	0.00%	2.19%	4.37%	7.65%	13.11%	72.68%
Reflection & Metacognition						
Ponderables Solutions <i>download</i>	0.00%	2.07%	6.21%	6.21%	14.48%	71.03%

Instructional Intentions

Opportunity #3: Telling Individuals' Stories, Making Predictions, and Providing Support

When these data are provided to instructors, they can apply adaptive step-loop responses to ensure the quality of instruction on the fly or design-loop versioning as they refine their instructional practice from term to term. When the data are supplied to artificial intelligence tools such as prediction algorithms, instructional support teams can provide further adaptive support for individual learners. The learning analytics literature is replete with examples of prediction algorithms that can be used to predict student achievement with small windows of data intake (Arizmendi, Bernacki, Rakovic et al., 2023). These prediction algorithms can be more accurate when raw data are enriched to reflect instructional design and learning processes (Bernacki et al., 2020a). Predictions can be made within days when context-responsive temporal coding is applied (Plumley et al., 2024). Additional literature documenting the real-time application of models confirms that designs that include DLP infrastructure, learning analytics to predict learning outcomes, and theory-aligned intervention materials to address students' predicted struggle can improve undergraduates' course success (Bernacki et al., 2020a; Cogliano et al., 2022). CLICKSTREAM has the capacity to facilitate data collection and enrichment, as well as the reapplication of prediction models to additional cohorts of students (Plumley, et al., 2024). This provides opportunities for task-loop adaptivity (Aleven et al., 2017) where cohorts of students whose early learning event data match those in prior semesters who obtained poor outcomes. Those classified as likely to benefit from support can be branched into a group that receives digital learning supports like learning skill training interventions that have been shown to promote more productive learning behavior in a course (Bernacki et al., 2020b), improved course achievement (Bernacki et al. 2021, 2023), and sustained benefits to achievement and retention in future semesters (Bernacki et al., 2025).

Opportunities for Educational Research

DLPs like CLICKSTREAM can not only provide real-time collection and enrichment of data to educators and learners to improve instruction, but they also provide opportunities for educational researchers to adopt narrative lenses to craft better stories describing learning. The second goal of theory is to richly describe the learning process, and theories of education require precise testing and refinement to tell these stories well (Greene, 2022). The enrichment of raw data that CLICKSTREAM provides offers an opportunity to educational theorists to pick up where instructors left off and add further description to the annotation of learning resources and their use as they reflect the complex learning processes described in theories of motivation and learning. In a recent special issue of the *Journal of Educational Psychology* on the power of the learning analytics opportunities provided by DLPs, Bernacki (2025) highlights the ways that educational theories describe learner engagement in task settings, and the assumptions about their relationship to learning and performance proposed within theories of motivation and learning. CLICKSTREAM's extensible data modeling approach enables educational researchers to apply their own libraries of metadata to learning events to make them into theory-aligned learning events. This can take place when events' attributes match the prescribed assumptions about a learner, their action, and the temporal context under which they engage with an object. As the CLICKSTREAM user community grows and more data become available for such modeling, the platform holds promise as a resource for testing two types of assumptions.

Opportunity #1: Close, Sustained Observation of Individuals in a Single Setting

Theoretical frameworks like self-regulated learning include assumptions about the moment-to-moment event sequences that might reflect a loosely sequenced process of planning and enactment of cognitive strategies and learners' periodic metacognitive monitoring and adaptation as they pursue a self-set or task-determined goal. Those in the learning analytics community have developed action libraries and sequence-mining methods to model such assumptions in laboratory settings. However, complexities in self-regulated learning frameworks' nested series of assumptions require scaled opportunities to observe the few instances when such conditions are satisfied (Ben-Eliyahu & Bernacki, 2015). Modeling these assumptions with power to detect patterns that support hypothesized processes can require hundreds of observations, and CLICKSTREAM can provide pathways to isolation of those learners under those conditions.

Opportunity #2: Modeling the Same Type of Learner across Contexts

Many theories of motivation and learning document interactions where individuals learn best when they engage in specific learning processes, but also when those opportunities to learn are undertaken under specific conditions (e.g., situated expectancy value theory; Eccles & Wigfield, 2024). The data that are ingested and enriched in CLICKSTREAM include person- and context-level variables that afford the opportunity to examine how any individual who possesses one characteristic (e.g., an achievement goal or dimension of their identity) might benefit differently when their instructional context differs in its instructional affordances or climate. Sociocultural and situated models of motivation and learning require data that accommodate these multilevel and multiple group models to test such assumptions. CLICKSTREAM data can be used to procure samples with person- and event-level criteria and contextual heterogeneity to answer questions about what works, for whom, and under what conditions.

Scaling CLICKSTREAM as a DLP for the Promotion and Study of Learning

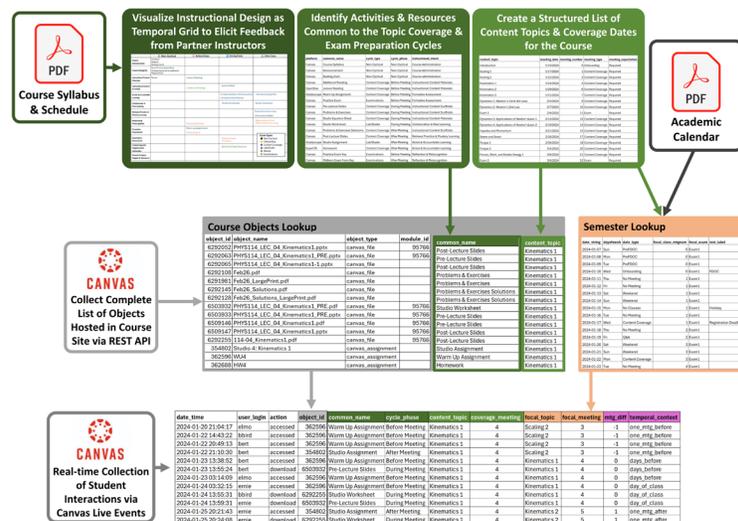
In order to leverage DLPs as critical infrastructure for educational research, the amount of available data and diversity of sources need to be grown, and the quality of instrumentation needs to be held constant as platforms like CLICKSTREAM are scaled to meet the needs of the educational research community. A few key initiatives need to be undertaken to establish CLICKSTREAM as a DLP that can host new sources and kinds of data, tag it efficiently for educator and researcher use, maintain it securely, and connect it with the broader ecosystem of DLPs to maximize the affordances of enriched data to support learning and educational research.

Improving Data Enrichment: Lightening Instructor Load while Maintaining Tagging Validity

During the prototyping of CLICKSTREAM, learning scientists worked closely with the instructors to analyze their course syllabi and schedule to create a course-specific tagging schema centered on the instructor's course design. This step included analysis of weekly topics, assignments, and milestones to ensure instructional alignment of the tagging framework. The tagging schema was then applied to all assets in the Canvas course site (e.g., pages, files, assignments, quizzes, external tools). Each asset was mapped to key elements of the instructional design by linking related assets by resource type, content topic, coverage date, and the expected timing of resource utilization within the instructor-specified learning cycle. A final review with the instructor ensured that each tag accurately reflected the intended purpose of the asset, clarified any ambiguous classifications, and resolved inconsistencies between how the instructor conceptualized a resource and how it was positioned in the tagging framework. The course-tagging schema was then aligned to a generalizable schema that emphasized instructional intent and the before-during-after phase of high-structure course designs. The tagging framework was then mapped to LMS events (e.g., page views, file downloads, assignment submissions) to create enriched student interaction data. This process transformed raw log data into structured event data that reflects not only the underlying course design but also the temporal context of students' engagement (Figure 8).

Figure 8.

CLICKSTREAM data architecture.



What Does CLICKSTREAM Require of Instructors?

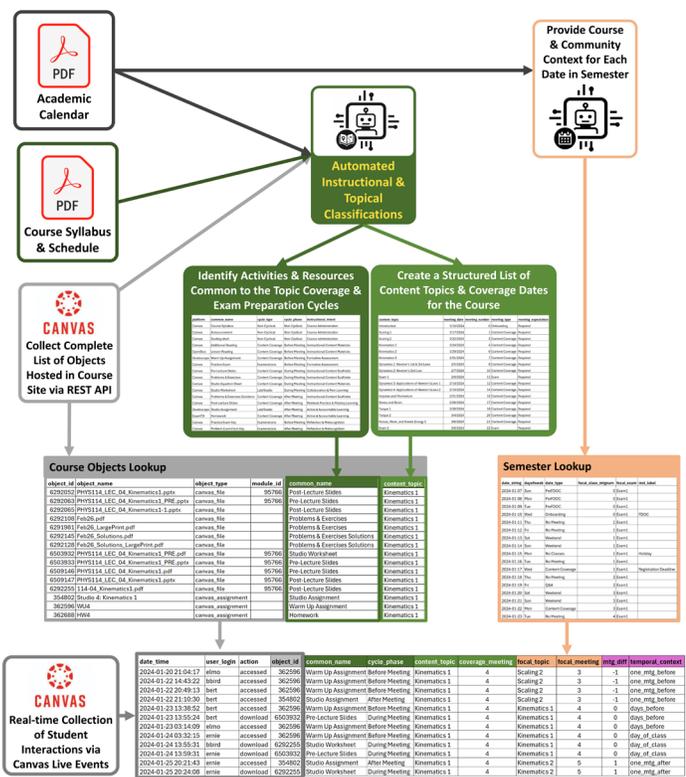
Instructors who wish to use CLICKSTREAM are asked to participate in a collaborative process that ensures their course can be interpreted accurately and meaningfully. This begins by working with educational researchers to create course asset metadata by reviewing the syllabus, clarifying the instructional intent of course materials, and validating the application of the metadata tags to files, pages, assignments, and other resources. Instructors can also support the tagging process by maintaining consistent naming conventions for course resources, which makes classifying assets with resource type, content topic, and instructional intent easier. Overall, such an effort requires just a few hours of instructors' time up front and occasional checking thereafter. Ongoing efforts to improve the tagging workflow continue to lower these time costs.

How Does Ongoing Development and Incorporation of AI Lighten the Load on Instructors?

Early-stage development of AI-driven features for CLICKSTREAM shows promise for automating much of the data enrichment and analytics process that once required extensive manual effort. Large language models (LLMs) have been shown to generate consistent metadata labels and descriptions for course materials by analyzing course syllabi, schedules, and resource naming conventions. The use of LLMs for this process will reduce the need for instructors and learning scientists to create initial versions of course mappings and tagging frameworks. This will allow the humans-in-the-loop of the tagging process to move from a generative to a confirmatory role where corrections made to LLM-created labels will not only produce accurate metadata more quickly but also provide additional data for fine-tuning the classification model going forward (Figure 9).

Figure 9.

AI integration to replicate educators' initial tagging efforts and extend schema to future courses and instructional designs.



Building a Community around the Promotion and Study of Learning on CLICKSTREAM

As we leverage emerging tools like LLMs to build momentum and lower effort barriers to the adoption of CLICKSTREAM, we also need to engage in the more human enterprises of community building around digital learning platforms (a goal of the SEERNet community). In order for a DLP to truly serve as a platform that advances inquiry about and support for learning, we need to overcome challenges to hosting one's own data, accessing data supplied by others, and building collective knowledge and resources to tell the many stories that students' data can describe.

Expanding Adoption: Data Sharing and Governance with Brief, Initial IT Support

Universities protect student data security with carefully crafted policies that align to federal standards for privacy about student admissions and achievement data (protected by FERPA policy), but because data about the learning process and in-course achievement data like exams are not formally protected, institutions also establish internal policies for data governance. When data from a course at a new institution are to be hosted on CLICKSTREAM, data-sharing agreements are established between institutions in alignment with current data governance policy. IT leads who manage the LMS and other learning technologies need to spend only under an hour establishing a data flow to CLICKSTREAM, where hosting and enrichment of those data can begin. Access to data and inclusion of student identifiers are limited to the lowest viable levels to allow data integration from multiple platforms during modeling and enrichment, as well as educator use during instruction. De-identification follows after the semester to further prioritize data privacy, before data are hosted in spaces where open educational research can follow.

Broadening Contribution to the DLP Ecosystem: DLP-to-DLP Integration to Support Multimodal Learning Analytics

Instructors tend to conduct the bulk of their course design on the university LMS, and adopt a variety of digital tools that they integrate to offer the best learning opportunities to their students. CLICKSTREAM has the capacity to ingest and enrich not only LMS data, but data sources from these other technologies via the same learning tools interoperability (LTI) and application programming interface (API) connections that are established to enable students to access technologies and technologies to serve learners. A common example of other essential technologies for undergraduate instruction include digital textbooks, which themselves are being established by OpenStax as a digital learning platform. Researchers at OpenStax and partner institutions trace students' access, navigation, reading, and annotation of assigned texts and tag these events with their own schema describing self-regulated learning (Winne & Hadwin, 1998; Zimmerman, 2000). The co-development of CLICKSTREAM and OpenStax's tagging capacity allows opportunities for further data integration that can more completely describe students' self-regulated learning process, including events in the LMS reflecting the course as a task and students' adherence to planned coverage of lessons, as well as their fine-grained engagement across learning resources and with the texts where they acquire content knowledge. Integration of data across DLPs holds promise for educational theory development, as it broadens the gaze for observation and affords opportunities to richly describe a greater portion of the learning process. Further integration with experimental design DLPs, like those

in SEERNet including Terracotta, and cognate DLPs, like ASSISTments TestBed and Carnegie Learning's UpGrade, provide opportunities for systematic experimentation to test assumptions of educational theories, as well as intervention approaches that can promote desirable and reduce undesirable approaches to learning (e.g., procrastination; Miyake & Kane, 2022).

Building a Data Storytelling Community: Welcoming Educational Theorists to Embrace DLPs

As the ecosystem of DLPs matures into a robust network of data platforms that host, enrich and share data describing learning where it happens, educational researchers who propose and study theories of learning will need to be welcomed as narrators who can tell students' stories and do so from the cognitive and social perspectives their theoretical lenses provide. Documentation of raw and enriched data needs to be offered so that theorists can dig into the work of description, tagging events by learners that occur in tasks and contexts with the insights about when and how these reflect theorized aspects of the learning process. Such theoretical tagging can then be tested and validated (e.g., using multimodal data to confirm inferences about their behaviors made by instructors to corroborate students' own descriptions; Bernacki et al., 2025), applied extensively to datasets, and shared with the research community. When a community-based tagging effort bridges the gap from raw to contextualized to theoretically-grounded descriptions of the learning process, individuals and their learning processes can be seen and more fully known in ways that enable the education community to understand and provide support for each student as they pursue their learning goals.

References

- Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2017). Instruction based on adaptive learning technologies. In R. E. Mayer & P. Alexander (Eds.), *Handbook of Research on Learning and Instruction* (2nd ed., pp. 522-560). New York: Routledge.
- Arizmendi, C. J., Bernacki, M. L., Raković, M., Plumley, R. D., Urban, C. J., Panter, A. T., Greene, J. A. & Gates, K. M. (2022). Predicting student outcomes using digital logs of learning behaviors: Review, current standards, and suggestions for future work. *Behavior Research Methods*, *55*, 3026-3054. doi.org/10.3758/s13428-022-01939-9.
- Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: A focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, *10*(1), 1–13. doi.org/10.1007/s11409-015-9134-6.
- Bernacki, M.L. (2018). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. In D.H. Schunk & J.A. Greene (eds.) *Handbook of Self-Regulated Learning and Performance*. (pp. 370-387). Routledge.
- Bernacki, M. L. (2025). Leveraging learning theory and analytics to produce grounded, innovative, data-driven, equitable improvements to teaching and learning. *Journal of Educational Psychology*, *117*(1), 1.
- Bernacki, M.L., Chavez (nee Dominguez), M. M., & Uesbeck, P. M. (2020a). Predicting STEM achievement with learning management system data: prediction modeling and a test of an early warning system. *Computers & Education*, *158*. doi.org/10.1016/j.compedu.2020.
- Bernacki, M.L., Cogliano, M.C., Kuhlmann, S.L., Utz, J. C., Strong, C., Hilpert, J. C., & Greene, J.A. (2023). Relations between undergraduates' self-regulated learning skill mastery during digital training and biology performance. *Metacognition and Learning*, *18*, 711-747. doi.org/10.1007/s11409-023-09356-9.
- Bernacki, M.L., Vosicka, L. & Utz, J.C. (2020b). Can brief, web-delivered training help STEM undergraduates "learn to learn"? *Journal of Educational Psychology*, *112*(4), 765–781. doi.org/10.1037/edu0000405.
- Bernacki, M.L., Vosicka, L., Utz, J.C. & Warren, C. (2021). Effects of digital learning skill training on the academic performance of undergraduates in science and mathematics. *Journal of Educational Psychology*, *113*(6), 1107–1125. doi.org/10.1037/edu0000485.
- Bernacki, M. L., Yu, L., Kuhlmann, S. L., Plumley, R. D., Greene, J. A., Duke, R. F., Freed, R., Hollander-Blackmon, C., & Hogan, K. A. (2025). Using multimodal learning analytics to validate digital traces of self-regulated learning in a laboratory study and predict performance in undergraduate courses. *Journal of Educational Psychology*, *117*(2), 176–205. doi.org/10.1037/edu0000890.
- Cogliano, M.C., Bernacki, M.L. Hilpert, J.C., & Strong (2022). A self-regulated learning analytics prediction-and-intervention design: detecting and supporting struggling biology students. *Journal of Educational Psychology* *114*(8), 1801–1816. doi.org/10.1037/edu0000745.

- de la Rosa, M., Schmitz, F. M., Tekian, A., & Guttormsen, S. (2025). Enhancing institutional processes and student experience in higher education through Journey Mapping: A systematic review. *Review of Education, 13*(3), e70100.
- Eccles, J. S., & Wigfield, A. (2024). The development, testing, and refinement of Eccles, Wigfield, and colleagues' situated expectancy-value model of achievement performance and choice. *Educational Psychology Review, 36*(2), Article 51. doi.org/10.1007/s10648-024-09888-9.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the national academy of sciences, 111*(23), 8410-8415.
- Gibson, D. (2017). Big Data in Higher Education: Research Methods and Analytics Supporting the Learning Journey. *Technology, Knowledge & Learning, 22*, 237–241. doi.org/10.1007/s10758-017-9331-2.
- Greene, J. A. (2022). What can educational psychology learn from, and contribute to, theory development scholarship? *Educational Psychology Review, 34*(4), 3011-3035.
- Lodge, J. M., Alhadad, S. S., Lewis, M. J., & Gašević, D. (2017). Inferring learning from big data: The importance of a transdisciplinary and multidimensional approach. *Technology, Knowledge and Learning, 22*(3), 385-400.
- Mayer, R. E., & Pilegard, C. (2005). Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. *The Cambridge Handbook of Multimedia Learning, 13*, 316-44.
- Miyake, A., & Kane, M. J. (2022). Toward a holistic approach to reducing academic procrastination with classroom interventions. *Current Directions in Psychological Science, 31*(4), 291-304.
- Plass, J. L., & Pawar, S. (2020). Toward a taxonomy of adaptivity for learning. *Journal of Research on Technology in Education, 52*(3), 275-300
- Plumley, R. D., Bernacki, M. L., Greene, J., Kuhlmann, S., Raković, M., Urban, C. J., Hogan, K. A., Lee, C., Panter, A. T., & Gates, K. M. (2024). Co-designing enduring learning analytics prediction and support tools in undergraduate biology courses. *British Journal of Educational Technology*. doi.org/10.1111/bjet.13472.
- Rosenzweig, E. Q., Hecht, C. A., Priniski, S. J., Canning, E. A., Asher, M. W., Tibbetts, Y., ... & Harackiewicz, J. M. (2021). Inside the STEM pipeline: Changes in students' biomedical career plans across the college years. *Science Advances, 7*(18), eabe0985.
- Shekhar, P., Borrego, M., DeMonbrun, M., Finelli, C., Crockett, C., & Nguyen, K. (2020). Negative student response to active learning in STEM classrooms. *Journal of College Science Teaching, 49*(6), 45–54.
- Tracey, M. W., & Baaki, J. (2022). Empathy and empathic design for meaningful deliverables. *Educational Technology Research and Development, 70*(6), 2091-2116.

- Wilder, J., & Lillvist, A. (2018). Learning journey: A conceptual framework for analyzing children's learning in educational transitions. *European Early Childhood Education Research Journal*, 26(5), 688–700.
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice* (pp. 277–304). Lawrence Erlbaum Associates.
- 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.

1 Center for Learning Analytics
School of Education
University of North Carolina at Chapel Hill

For more information about CLICKSTREAM or to inquire about partnerships,
contact the authors at <https://learninganalytics.unc.edu/>

Suggested Citation

Bernacki, M.L., and Plumley, R.D. (2026, February). *CLICKSTREAM: Essential educational infrastructure to promote student learning and educational theory development*. Digital Promise. <https://doi.org/10.51388/20.500.12265/285>



Attribution-NonCommercial-NoDerivs 4.0 International



Washington, DC:
1001 Connecticut Avenue NW, Suite 935
Washington, DC 20036

Redwood City, CA:
702 Marshall St., Suite 340
Redwood City, CA 94063

Website: <https://digitalpromise.org/>