Webel: Web Browsing Event Logger

Elijah Cantu[1,*], Gyunam Park[2]

[1]Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, United States
[2]Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, the Netherlands

Abstract

Web browsing behavior offers a rich yet underexplored source of data for process mining (PM) research. However, existing web analytics tools primarily rely on server-side logs or unstructured browsing history, which lack detailed and organized event logging needed for effective process analysis. To address this gap, we introduce Webel, a Chromium-based browser extension tailored for PM researchers. Webel captures structured web browsing events by automatically grouping related activities into coherent process traces while tracking tab relationships and navigation patterns. It supports multiple trace modeling strategies, including parent-child tab relationships, per-tab isolation, and session-wide grouping, providing flexible analytical perspectives. Webel also enables seamless export of logs in the XES format for immediate use in PM tools and offers optional Google Drive integration for cloud-based log synchronization.

Keywords

Process Mining, Web Browsing Behavior, User Journey, Web Event Logging, Browser Extension

Metadata description	Value
Tool name	Webel
Current version	1.0
Legal code license	MIT
Languages, tools and services used	Chrome Extensions API (Manifest V3)
Supported operating environment	Chromium-based browsers
Download/Demo URL	webels.org/install
Documentation URL	webels.org
Source code repository	github.com/elijahcantu/webel
Screencast video	webels.org/demo

1. Introduction

PM research has traditionally focused on business processes derived from enterprise systems, with limited exploration of user-level web browsing behavior as a data source. Although Web server logs provide information on user interactions with specific websites [1], they do not capture the entire user journey between multiple sites. Existing browser history tools and web analytics platforms collect extensive browsing data but lack the semantic structure required for

ICPM Doctoral Consortium and Demo Track 2025, October 20-24, 2025, Montevideo, Uruguay
*Corresponding author.
caelijah@umich.edu (E. Cantu); g.park@tue.nl (G. Park)
https://elijahcantu.com (E. Cantu); https://www.gyunam.com (G. Park)
0009-0008-4333-6248 (E. Cantu); 0000-0001-9394-6513 (G. Park)

PM analysis. Specifically, they do not organize events into meaningful traces with clear process boundaries [2].

This gap between raw browsing data and PM-ready event logs presents a significant barrier to researchers analyzing web-based user behavior. Current tools lack the ability to create logical web browsing processes [3], often requiring extensive data preprocessing that can introduce artificial trace boundaries that may not accurately reflect navigation behavior.

Webel addresses these challenges by providing a browser extension that enables the automatic grouping of web browsing events into coherent, semantically meaningful process instances.

2. Tool Innovations and Features

Webel is a Chromium-based extension specifically designed for PM applications. Unlike conventional web analytics tools that log events as flat, unstructured sequences, Webel implements multiple trace modeling strategies to capture the semantic structure of user browsing behavior. It preserves relationships between tabs and navigation events and exports the resulting event logs directly in XES format, facilitating seamless integration with PM tools.

2.1. Trace Modeling Strategies

Webel supports three trace methods to group web browsing behavior into process instances. Each offers a different analytical perspective:

- **Parent-Child Tab Traces:** This method groups browsing activities based on tab relationships. When a user opens a new tab page, it starts a new trace as the parent. If a link from this tab opens in another tab, the new tab inherits the parent's trace ID. This way, it captures the user's main tasks along with any subtasks branching off.
- **Per-Tab Traces:** Each browser tab is treated as a separate trace, regardless of how it was opened. This allows for independent analysis of navigation behavior per tab and is particularly useful when examining parallel exploration patterns.
- **Session-Based Traces:** All activity during a single browser session is grouped into a single trace. This models coarse-grained user behavior across a session, useful for longitudinal task analysis or macro-level patterns.

2.2. Trace Export and Formats

Each trace method is logged simultaneously, allowing researchers to export and compare all three variants in XES format using the Webel popup or the Google Drive sync. These logs are fully compatible with PM tools such as *ProM*, *Disco*, *Celonis*, and *pm4py*, enabling immediate analysis without preprocessing.

2.3. Google Drive Integration

Webel supports optional Google Drive synchronization. Users can enter a Google Drive folder ID or URL they can access. The logs are automatically synced to that folder using OAuth 2.0

and the Google Drive API. This feature facilitates cross-device analysis and shared research workflows while retaining user control over storage and sharing.

2.4. Privacy and Control

All data remains local unless explicitly exported or synced to Google Drive. The extension popup lets users toggle logging on or off, trigger XES exports, or sync to Google Drive at any time. If logging is disabled, reminder notifications and badge indicators prompt users to re-enable it after they start a new browser session or clear their history.

3. Case Study 1: Course Resource Navigation

To evaluate Webel's capabilities, we conducted a case study focused on how a user accessed course-related resources on the Canvas platform, a widely used learning management system in higher education, at the University of Michigan. In this scenario, the user visited the Canvas page for each enrolled course and then attempted to reach all external resources linked from that course page.

Although this browsing pattern is common among students, it involves a complex sequence of navigation steps that make it well-suited for testing Webel's trace logic. We recorded the user's browser sessions during this process and analyzed the resulting logs using Disco [1] after applying basic URL-based regular expressions to label each resource.

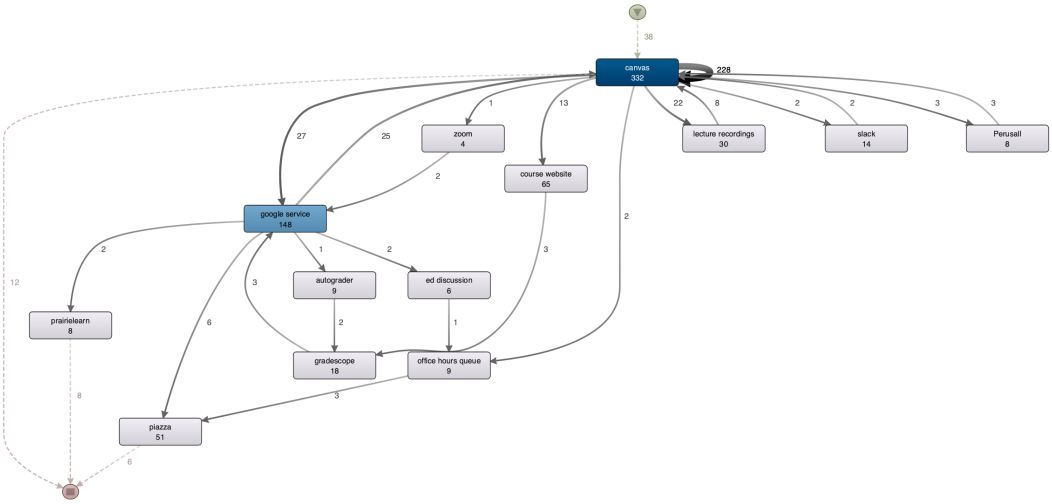
Figure 1a–c present process models generated from the same set of browsing sessions, each visualized using a different trace modeling strategy. The parent-child trace map captures how external resource visits are organized under each course as distinct sub-tasks, revealing the hierarchical structure of student navigation and how main activities generate subtasks. The per-tab model provides insight into how the user structured their navigation across browser tabs, showing clear separation of activity per course and highlighting parallel exploration behavior. The session trace offers a comprehensive temporal view of the entire session, reflecting the user's sequential progress through a course and its associated resources, and capturing habitual return points and overall workflow patterns.

The generated models reveal several important insights. Canvas emerges as the central hub across all perspectives, with substantial self-loop activity (228 returns), indicating frequent returns as a primary "home base." A dual-hub pattern also appears, with Google Services (148 visits) serving as a secondary integration hub, reflecting the role of third-party tools in educational workflows.

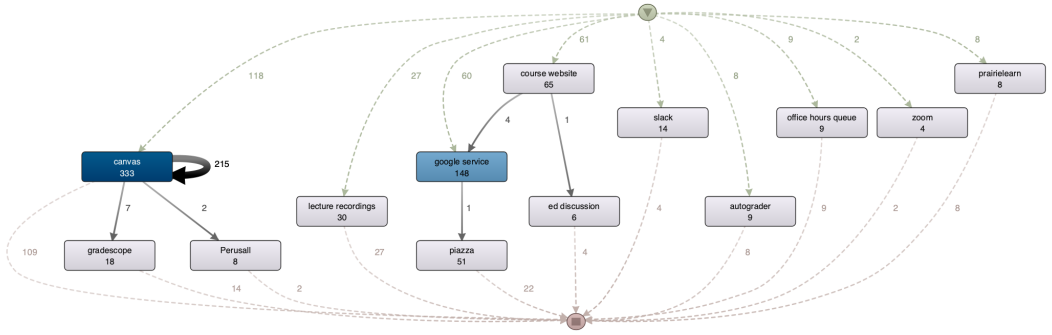
Resource usage shows clear functional distinctions: lecture recordings (30 visits) and Slack (14) are heavily used, while assessment platforms like Gradescope (18) and discussion forums like Piazza (51) see moderate activity. Synchronous tools such as Zoom (4) and office hours queues (8–9) have lower but consistent usage.

Taken together, the three trace perspectives provide complementary views: parent-child traces reveal task hierarchies and branching behavior, per-tab traces emphasize parallel exploration, and session traces illustrate sequential workflows and macro-level patterns. This case study

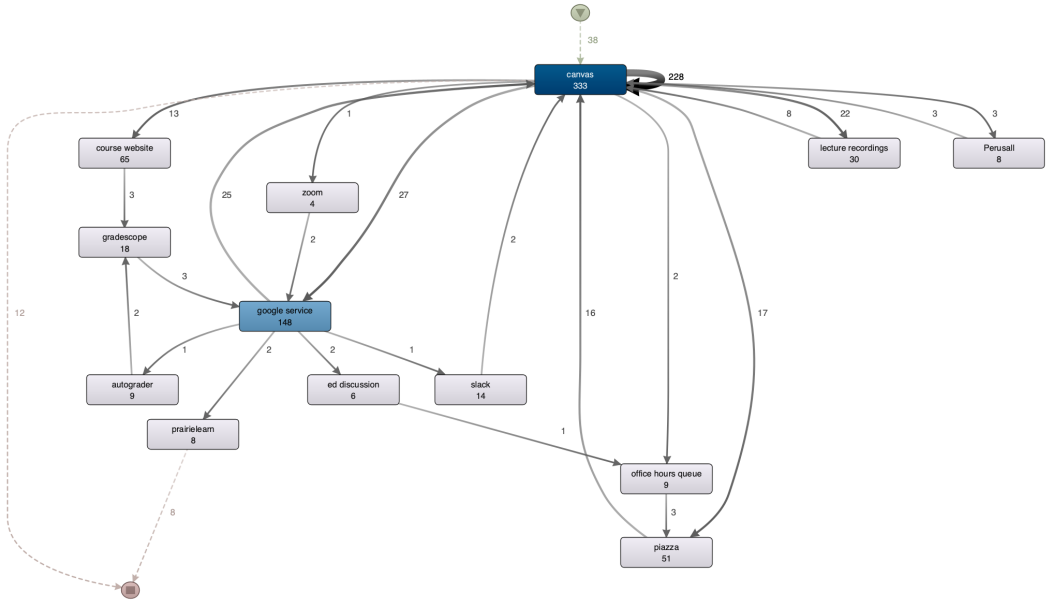
[1] https://fluxicon.com/disco/



(a) Parent-child tab process map



(b) Per-tab process map



(c) Per-session process map

Figure 1: Comparing navigation flows using three trace modeling strategies

highlights Webel's flexibility in supporting multiple trace perspectives out of the box, allowing researchers to choose the most appropriate level of abstraction for their analysis without requiring manual segmentation or complex preprocessing.

4. Case Study 2: Canvas Files User Journey

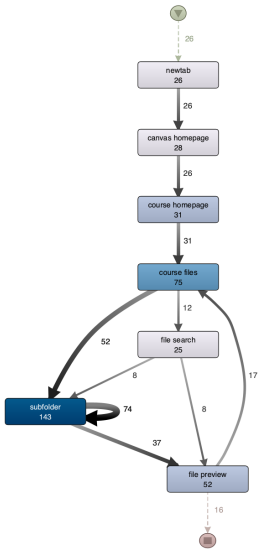


Figure 2: Process Map of a User Navigating Canvas Files

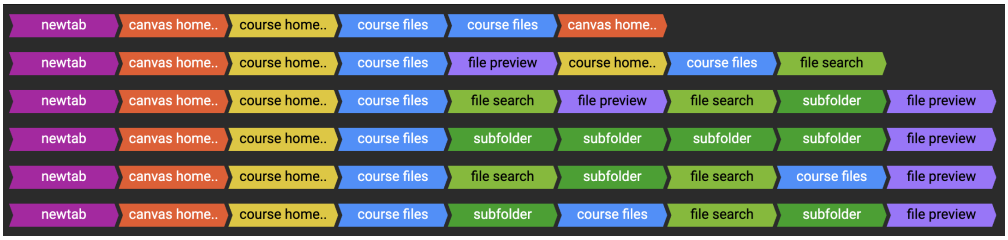


Figure 3: Process Variants of a User Navigating Canvas Files

To demonstrate Webel's applicability to single-tab navigation analysis, we examined a focused scenario within Canvas's file management system. This study was conducted entirely within a single tab and browser session, so we applied the per-tab trace method. Since Canvas does not support subfolder previews, users must repeatedly enter and exit individual folders to explore the file hierarchy. This behavior results in a navigation pattern that is well-suited for analyzing inefficiencies within isolated browsing contexts.

We recorded browsing events during typical file navigation tasks and applied more fine-grained URL-based regular expressions to label each navigation. The resulting event log was analyzed using Disco and Cortado [4]. The process model in Figure 2 reveals frequent loops between *subfolder* and *file preview* activities, quantifying the repetitive overhead imposed by Canvas's interface design. Figure 3 shows multiple process variants, each containing redundant cycles that highlight the absence of efficient file browsing paths. This focused analysis demonstrates Webel's practical value for identifying specific usability bottlenecks. Webel can transform raw browsing data into actionable insights for user experience improvement.

5. Conclusion

The tool demonstrates robust support for logging web browsing events, and its XES export functionality has been verified for compatibility with leading PM platforms. Webel focuses exclusively on Chromium-based browsers, and logging activities outside the browser is out of scope for the current design. Future work includes adding support for Safari.

Multiple trace modeling strategies in Webel address the need for flexible process instance identification in web browsing and allow researchers to select the most appropriate level of abstraction for their research questions without manual preprocessing.

6. Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT and Claude for grammar and spelling checks. After using these services, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] H. Sarirah Husin, S. Ismail, Process mining approach to analyze user navigation behavior of a news website, in: Proceedings of the 4th International Conference on Information Science and Systems, ICISS '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 7–12. doi:10.1145/3459955.3460593.

[2] N. Bhattacharya, J. Gwizdka, Yasbil: Yet another search behaviour (and) interaction logger, in: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 2585–2589. doi:10.1145/3404835.3462800.

[3] GESIS - Leibniz Institute for the Social Sciences, Gesis web tracking, 2025. URL: https://www.gesis.org/en/services/planning-studies-and-collecting-data/collecting/gesis-web-tracking, accessed: [insert your access date].

[4] D. Schuster, S. J. van Zelst, W. M. van der Aalst, Cortado: A dedicated process mining tool for interactive process discovery, SoftwareX 22 (2023) 101373. doi:10.1016/j.softx.2023.101373.