

Wikimedia Foundation

March 7, 2013

http://bit.ly/ELWorkshop

Program

1.30-3.00. EventLogging tutorial

- Intro and rationale (*Dario*)
- From an idea to a data model: Designing a schema (*Maryana*)
- Instrumenting code, collecting and validating EventLogging data (Ori)
- Accessing, QA'ing and analyzing EventLogging data (Dario)

3.00-3.30 Coffee break

Coffee, cookies and snacks

3.30-5.00 Hacking session

Hands-on session to learn how to use EventLogging

Service announcements

IRC channel

#wikimedia-e3 @ freenode.net

Streaming/recording

You're on the air (1.30-3pm)

http://www.youtube.com/watch?v=ZBuTjl0lh9Y

Mailing list

Sign up at http://bit.ly/EL-News

Introduction and rationale

What we do

Experiments

The following pages hold the *methodology*, *experimental design and results* of each experiment. For technical information and UX design documentation please view our E3 hub on MediaWiki.org. For a short *overview of the results* of completed experiments check out this page.

Active

- Research: Onboarding new Wikipedians
- Research: Account creation UX, including campaigns to drive registration

Planned

- Research:Section edit modification
- Research:Improve your edit
- Research:Wikimania unconference experiment (suppress cleanup templates)
- Research:Left-hand navigation

Complete

- Research: Donor engagement
- Research: Community portal redesign
- Research:Post-edit feedback
- Template A/B testing
- Research: Necromancy
- Research: Timestamp position modification
- Research: Editor milestones

Editor Engagement Project

Editor Engagement Experiments

project leads

Munaf Assaf

Wikimedia Foundation

Ryan Faulkner

Wikimedia Foundation

Matt Flaschen

Wikimedia Foundation

Ori Livneh

Wikimedia Foundation

S Page

Wikimedia Foundation

Dario Taraborelli

Wikimedia Foundation

Steven Walling

Wikimedia Foundation

start

2012-April

W Editor engagement initiatives

Typical problems

How many new users complete a funnel?

How many users fail to complete an edit?

Is the conversion rate of a new experimental feature significantly higher than a control?

What is the breakdown by user segment in the adoption of a feature?

open source usable scalable transparent

Web analytics

for rapid prototyping and A/B testing

High-quality, granular data to measure how users interact with features and the UI.

client-side events:

clicks
impressions
bucketing events
errors

server-side events: transactions

Third-party solutions



Legacy solutions



Legacy solutions

Extension:ClickTracking

This feature is used to track users' navigation around the wiki by logging each click they make. It has a couple of API modules.

ClickTracking was developed specifically to be run as part of the Beta rollout (of the Vector skin and Enhanced editing toolbar in 2010) on Wikimedia wikis; it may or may not be useful to third-party users. The extension can of course be run on non-WMF wikis, but documentation is very limited.

ClickTracking logs the user's edit counts using functionality and a database table implemented by Extension:UserDailyContribs, so you need to install that as well.

This extension has been migrated from Extension: UsabilityInitiative. During this migration this extension was ported to make use of ResourceLoader functionality, making it only compatible with MediaWiki 1.17+.

If you are upgrading from the UsabilityInitiative (MediaWiki 1.16 compatible) version, you may need to pay attention to the change in naming for configuration variables.

See README for an example.

1. scalability

```
* action=clicktracking *
```

Track user clicks on JavaScript items

This module requires read rights

Parameters:

eventid - string of eventID

This parameter is required

namespacenumber - the namespace number being edited

This parameter is required

token - unique edit ID for this edit session

This parameter is required

redirect to (only used for links that go off the page)

additional - additional info for the event, like state information

2. flexibility

Log format

Clicktracking data is logged using the following, tab-separated format:

enwiki ext.articleFeedbackv5@0-option1-impression-bottom 20120113013257 0 w4dtyfWSw714KsqmMAAkdC7isy8CmHvP9 0 0 0 Pokémon 274278

Log format specification

- 1. database name (e.g. enwiki or ptwikisource)
- 2. event name (as documented above)
- 3. timestamp (in YYYYMMDDhhmmss format)
- 4. user category: 1 for logged in users, NULL for anons
- 5. user token: an anonymous token replacing user names or IP addresses
- 6. namespace
- 7. lifetime edit count (logged in only)
- 8. 6-month edit count (logged in only)
- 9. 3-month edit count (logged in only)
- 10. last month edit count (logged in only)
- additional data (page title, rev_id)

2. flexibility

event names

ext.articleFeedbackv5@4-option3-cta_learn_more-button_click-overlay

additional data

e90RMgpyyOjKArMiyFUadnI1LgqkPXSU|http://en.wikipedia.org/wiki/Main_Page|Foo

3. data modeling

No mechanism for data validation

3. data modeling

Limited support for data documentation and provenance

4. data-centric collaboration

Product managers
Data analysts
Engineers
UX designers

EventLogging: design principles

Support getting desired metrics, not simply all available metrics

Provide data provenance/ enable long-term use of data

Support data-centric collaboration: using a wiki for data modeling

Support versioning/permit data sets to be compared

Support transparency/community scrutiny

Do not capture every action on site, but rich, well-documented data that fit the question

Enforce lightweight client and server-side validation

Creates a contract between data analyst, product manager, engineer around data definitions

Simplify access to data, QA and data analysis

Next

- From an idea to a data model: Designing a schema (*Maryana*)
- Instrumenting code, collecting and validating EventLogging data (Ori)
- Accessing, QA'ing and analyzing EventLogging data (Dario)