



# Metrics standardization

Dario Taraborelli • Aaron Halfaker

Wikimedia Research and Data showcase

March 2014

# Wikimetrics

Welcome to the Wikimedia Foundation's Wikimetrics homepage. This API allows you to

select a set of users, also known as a "cohort" (for example, **all users who signed up via the Thank You campaign**)

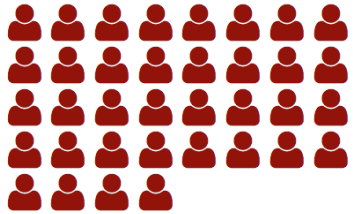
select a metric to be computed for each of these users (for example, how many [bytes they've added](#)) with optional parameters (for example, a **time range**) and retrieve the response in JSON or CSV format.

You can also compute a single, aggregate value for the cohort (like the **mean** revert rate) .

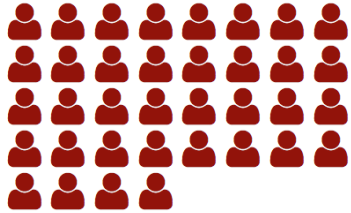
[Learn More](#)[Analyze](#)

[Login](#) to see your cohorts, metrics, and reports

summer 2013



**cohort-level metrics**



**cohort-level metrics**



**project-level metrics**

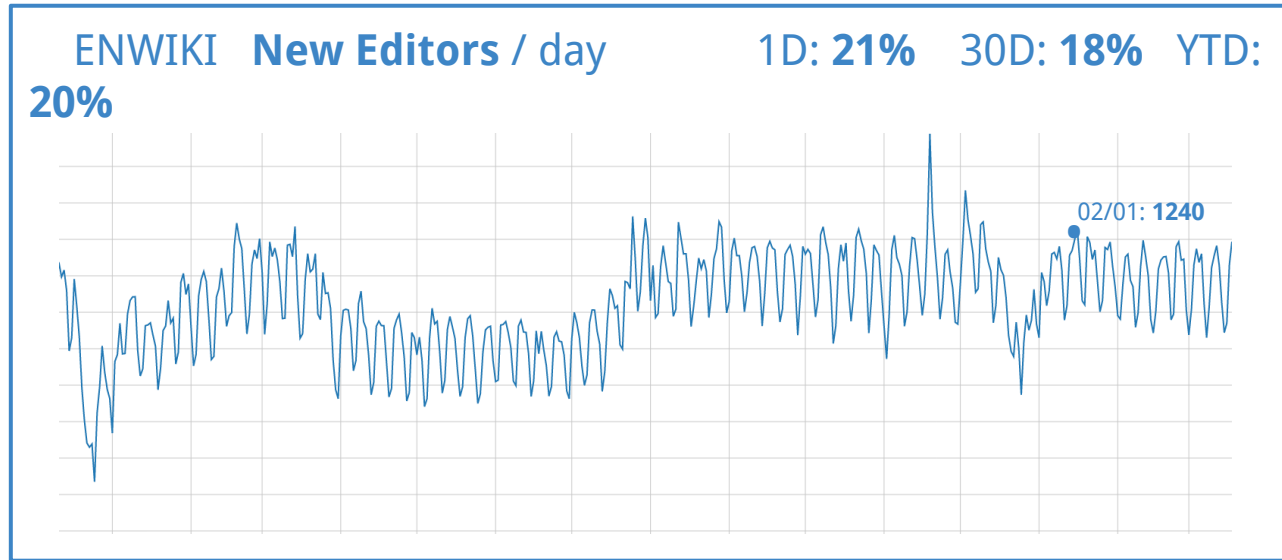


**project-level metrics**



project-level metrics

# Editor engagement vital signs



key performance indicators for user engagement, community and content growth  
aggregated daily / weekly / monthly  
for every single Wikimedia project

[https://www.mediawiki.org/wiki/Analytics/Epics/Editor\\_Engagement\\_Vital\\_Signs](https://www.mediawiki.org/wiki/Analytics/Epics/Editor_Engagement_Vital_Signs)

summer 2014

# Key metrics

## New users

**Newly registered users**

**New editors**

**Productive new editors**

**Surviving new editors**

...

## Community

**Editors**

**Active editors**

**Very active editors**

**IP editors**

**Bots**

**Page creators**

...

## Content

**Edits**

**Bot edits**

**Uploads**

**Pages**

...

## Curation

**Page deletions**

**Reverts**

...



# Rationale

## **Relevant**

Measure quantities that describe important phenomena

## **Replicable**

Make research easily reproducible and verifiable

## **Transparent**

Provide formal specifications, remove ambiguity

## **Consistent**

Replace proprietary, ad-hoc metric definitions; compare apples to apples

## **Robust**

Make metrics replicable via multiple data sources at any point in time

## **Granular**

Computable at different time scales

# Anatomy of a metric

## 1. specification

### Research:Productive new editor

**Productive new editor** is a [standardized user class](#) used to measure the number of first-time editors in a wiki project over time who make [productive contributions](#). It's used as a proxy for [editor productivity](#), and to a lesser extent, [editor activation](#). A "productive new editor" is a [new editor](#) who saves revisions to [content namespace](#) pages that are not [reverted](#).

#### Contents [hide]

##### 1 Discussion

- 1.1 Excluding edits to deleted content
- 1.2 The  $n$  productive edits threshold
- 1.3 The  $t$  time cutoff
- 1.4 Time to revert cutoff
- 1.5 Limitations

##### 2 Analysis

- 2.1 German Wikipedia
- 2.2 English Wikipedia
- 2.3 Spanish Wikipedia
- 2.4 French Wikipedia
- 2.5 Polish Wikipedia
- 2.6 Portuguese Wikipedia
- 2.7 Factor comparison of  $n$  and  $t$

##### 3 Usage

##### 4 References

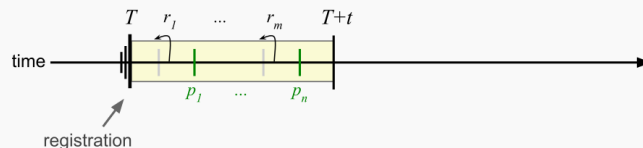
### Discussion [edit]

#### Excluding edits to deleted content [edit]

Spammers and other non-productive new editors tend to create articles that are non-

#### Metrics standardization

#### Productive new editor



#### Specification

A **productive new editor**( $n, t$ ) is a [new editor](#) who completes at least  $n$  [productive edit](#)(s) within  $t$  time since [registration](#) ( $T$ ).

#### WMF Standard

- $n = 1$  [productive edit](#)
- $t = 1$  [day](#)

#### Measures

[Editor productivity](#)

#### Aliases

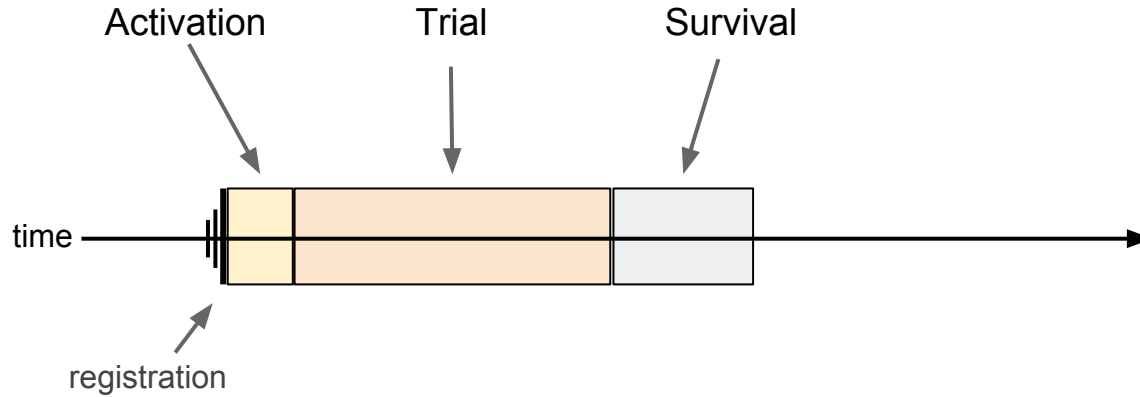
[Productive newcomer](#)

#### Related metrics

[Newly registered user](#) • [New editor](#) • [Productive edit](#)

# Anatomy of a metric

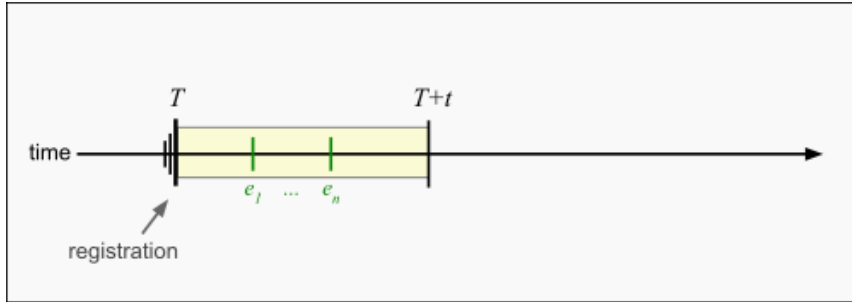
## 2. visualizations



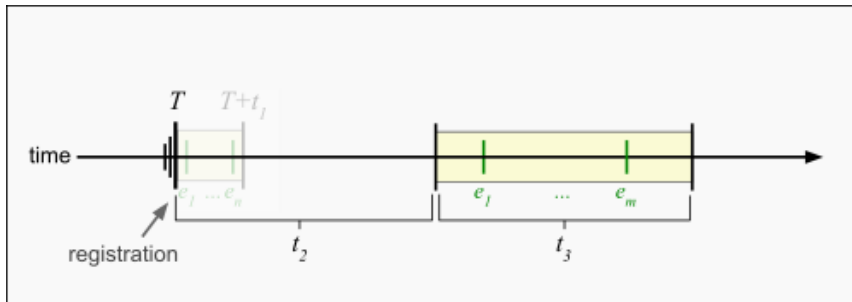
# Anatomy of a metric

## 2. visualizations

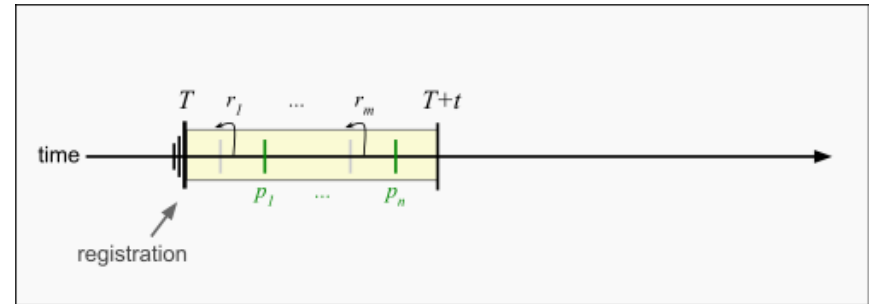
New editor



Surviving new editor



Productive new editor



# Anatomy of a metric

## 3. discussion

### Discussion

---

#### Excluding edits to deleted content

Spammers and other non-productive new editors tend to create articles that are non-productive and those articles tend to be deleted rather than the edits to the articles being reverted (and therefore excluding them from the [productive edit](#) criteria). Edits to articles that are deleted by the end of a new editor's first week since registration are *not* included in counts of productive edits.

#### The $n$ productive edits threshold

Like choosing an  $n$  for any metric based on counts (e.g. [new editor](#) and [active editor](#)), choosing a threshold is somewhat arbitrary. Choosing a higher threshold will result in a smaller proportion of [newly registered users](#) being considered productive.

#### The $t$ time cutoff

There are a few ways that the timespan for identifying productive edits can be drawn. The two most common ways are based on time bounds and events. A time-bounded approach is based on the use of some  $t$  cutoff to limit observations to a certain amount of time after a user registered their account. An event-based approach will use some event as the starting point to count user contributions. Another candidate time-span includes edits that a newcomer performed in their first [edit session](#). Since *productive new editor* qualifies the activity of a [new editor](#) we set  $t = 1$  day, which effectively makes the class of productive new editors a proper subset of new editors. We analyze the effect of choosing a different value for  $t$  below.

#### Time to revert cutoff

It is theoretically possible that a revision could be reverted years after it was originally saved, observations taken at any time would truncate any future reverts (see [en:Censoring\\_\(statistics\)](#), specifically "right censoring"). In order to minimize this issue and compare editors' contributions fairly, a revert is only counted if it occurred within 48 hours of the original edit. Recent research suggests that, at least for the English Wikipedia, nearly all reverts take place within 48 hours<sup>[1]</sup>.

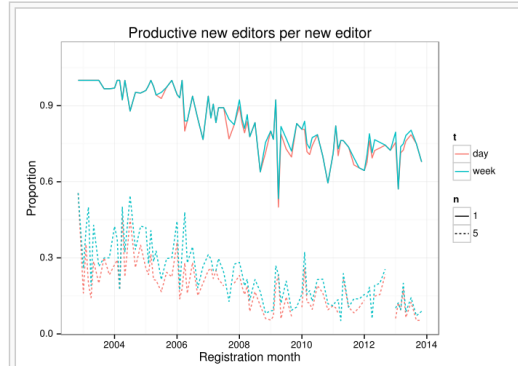
#### Limitations

- This metric represents productivity as a binary attribute of a user, it does not measure *how productive* a new editor is. New editors who make many productive edits and contribute substantial amounts of content will look identical (under this metric) to new editors who fix a few typos.
- The most clever vandalism/vandals may go unnoticed for more than 48 hours.

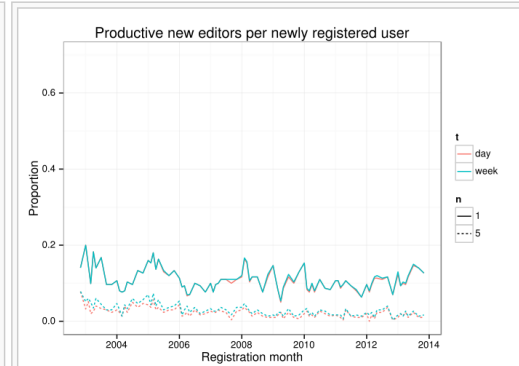
# Anatomy of a metric

## 4. sensitivity analysis

Spanish Wikipedia [\[edit\]](#)

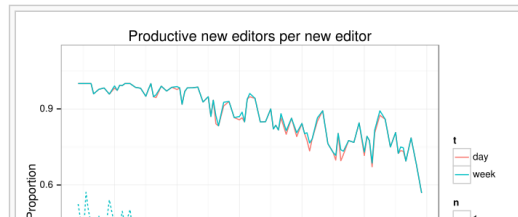


**Productive new editors per new editor.** The proportion of productive new editors is plotted by registration month for two values of  $t$ .



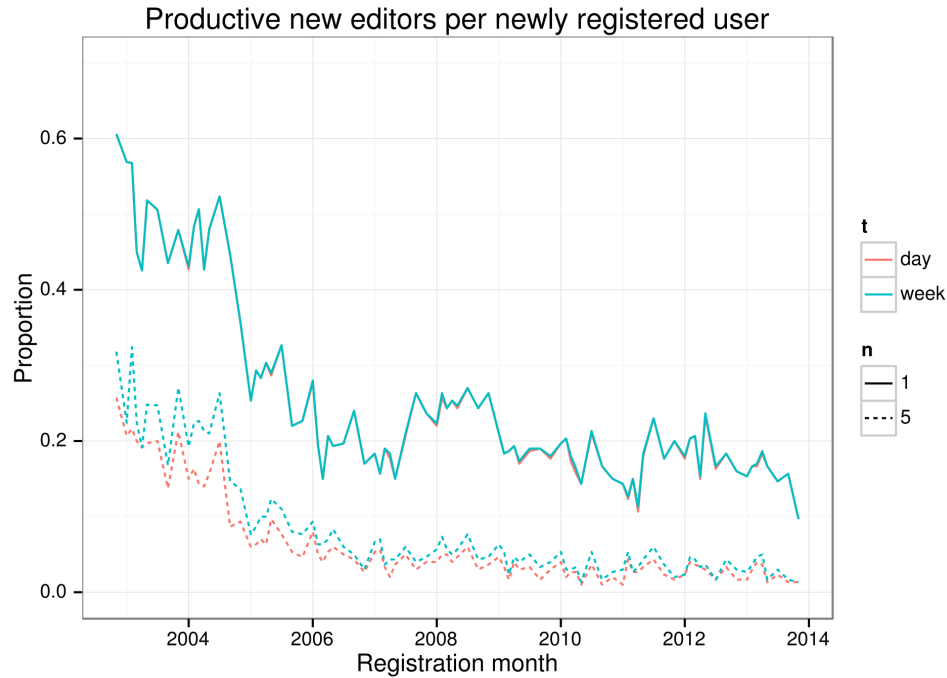
**Productive new editors per newly registered user.** The proportion of productive new editors is plotted by registration month for two values of  $t$ .

French Wikipedia [\[edit\]](#)



# Sensitivity analysis

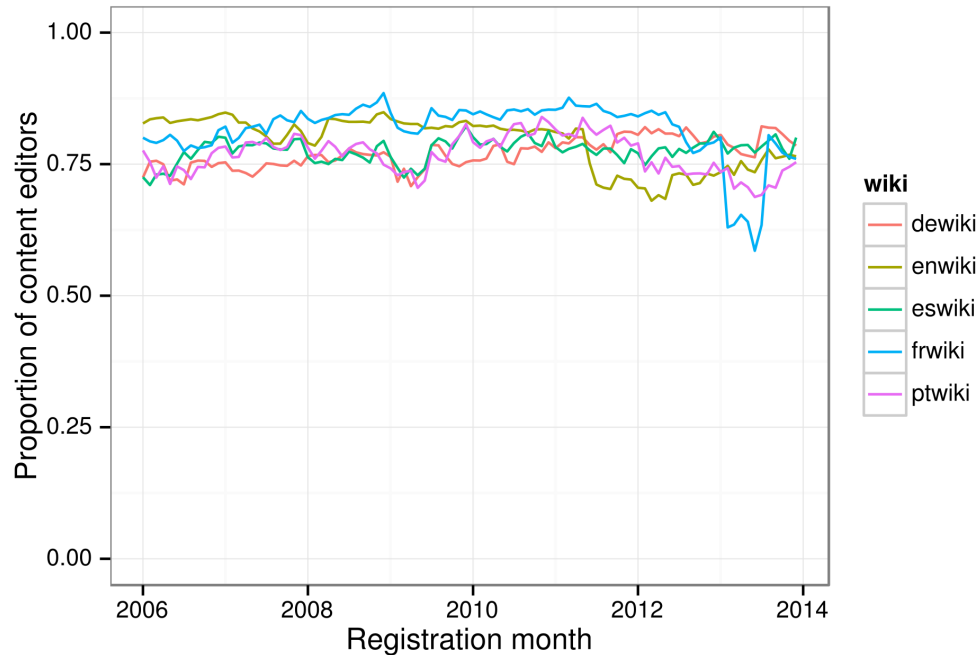
Does *new editor productivity* vary when we measure it over the first day or the first week?



[https://meta.wikimedia.org/wiki/Research:Productive\\_new\\_editor](https://meta.wikimedia.org/wiki/Research:Productive_new_editor)

# Sensitivity analysis

*Does it really matter to limit **new editor activation** to main namespace edits only?*





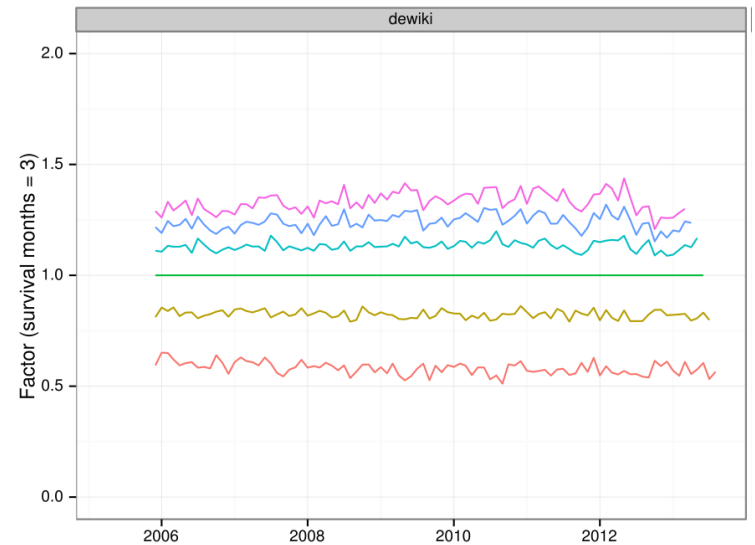
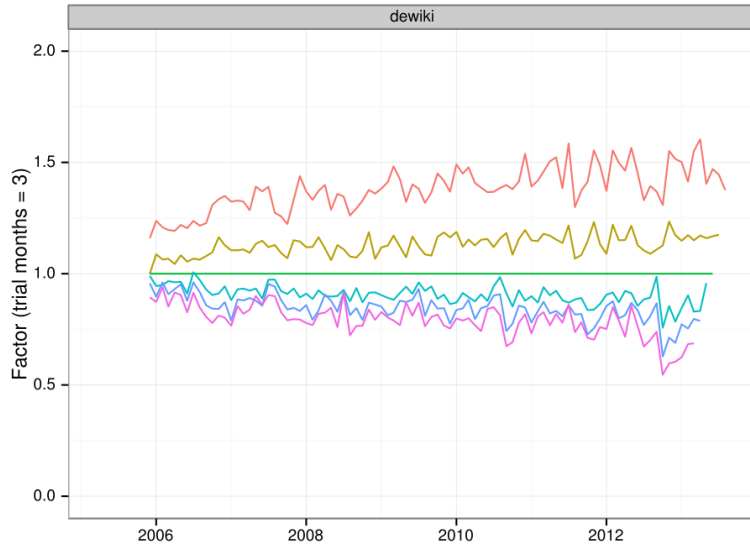
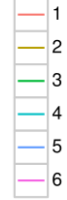
# Sensitivity analysis

*Does the length of the trial and survival period affect the measurement of **new editor survival**?*

factor(trial\_months)



factor(survival\_months)



# Why does this matter at all?

## 1. Data exploration

*“Newly registered users on German and Dutch Wikipedia have a higher activation rate than newbies on English Wikipedia”*

*“Spanish Wikipedia adds every day twice as many new editors than German Wikipedia, despite having only half its new user activation rate”*

## 2. Natural experiments

*“A change in abuse filter rules on the Italian Wikipedia significantly increased new editor survival”*

# Evaluation

Metric specification



Sensitivity analysis



Parameter recommendation



Release



Evaluation

# Evaluation



# Questions?

[dario@wikimedia.org](mailto:dario@wikimedia.org)

[ahalfaker@wikimedia.org](mailto:ahalfaker@wikimedia.org)

## Read more

[https://meta.wikimedia.org/wiki/Research:Metrics\\_standardization](https://meta.wikimedia.org/wiki/Research:Metrics_standardization)

## Image credits

W.Wood (1839) Index Entomologicus

<http://dx.doi.org/10.5962/bhl.title.12503>

