

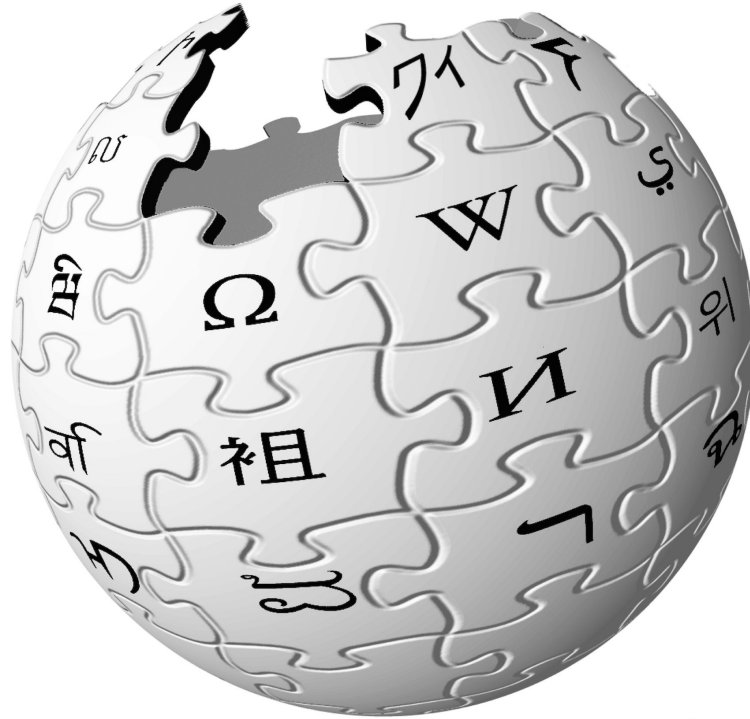


Analytics Engineering



WIKIMEDIA

F O U N D A T I O N



WIKIPEDIA
La enciclopedia libre

We build analytics infrastructure

analytics@lists.wikimedia.org

The Analytics Team sees as its primary responsibility making **Wikimedia related data available** for querying and analysis to both **WMF** and the **different Wiki communities** and stakeholders. We develop infrastructure so all our users, both within the Foundation as within the different communities, can access data in a self-service fashion that is ***consistent with the values of the movement.***

**We do not handle data
requests (for the most
part)**

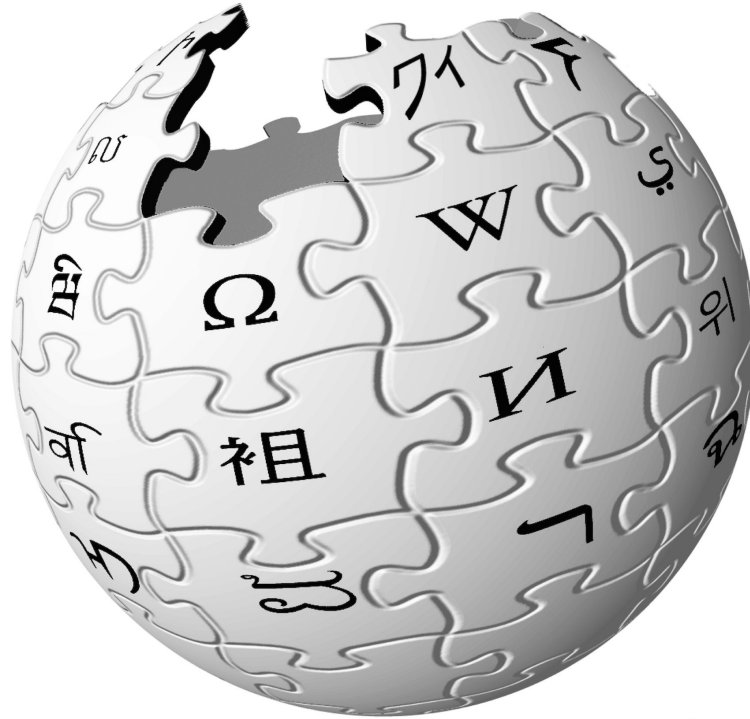
We try for (all) data to
be public by default.

The more accessible the data is, the more impact it can have.

But we are not there
Yet

Public Data

**Data that is Useful for the
world at large.**



WIKIPEDIA
La enciclopedia libre

6th largest website [Alexa]

Wikipedia reaches **hundreds of millions of unique devices** every month and, as such, are a good barometer of **browser popularity**.

The most popular browser

?

The most
popular browser
in April 2017

**Was Chrome 56
with 25% market
share**

<https://analytics.wikimedia.org/dashboards/browsers/#all-sites-by-browser>

User Agent Breakdowns Wikimedia Foundation

All Sites by OS ▾ **All Sites by Browser ▾** All Sites by OS and Browser ▾ Mobile Site by OS ▾ Mobile Site by Browser ▾ Desktop Site by OS ▾

Browser Family Timeseries **Browser Family and Major Hierarchical View** Browser Family and Major Tabular View

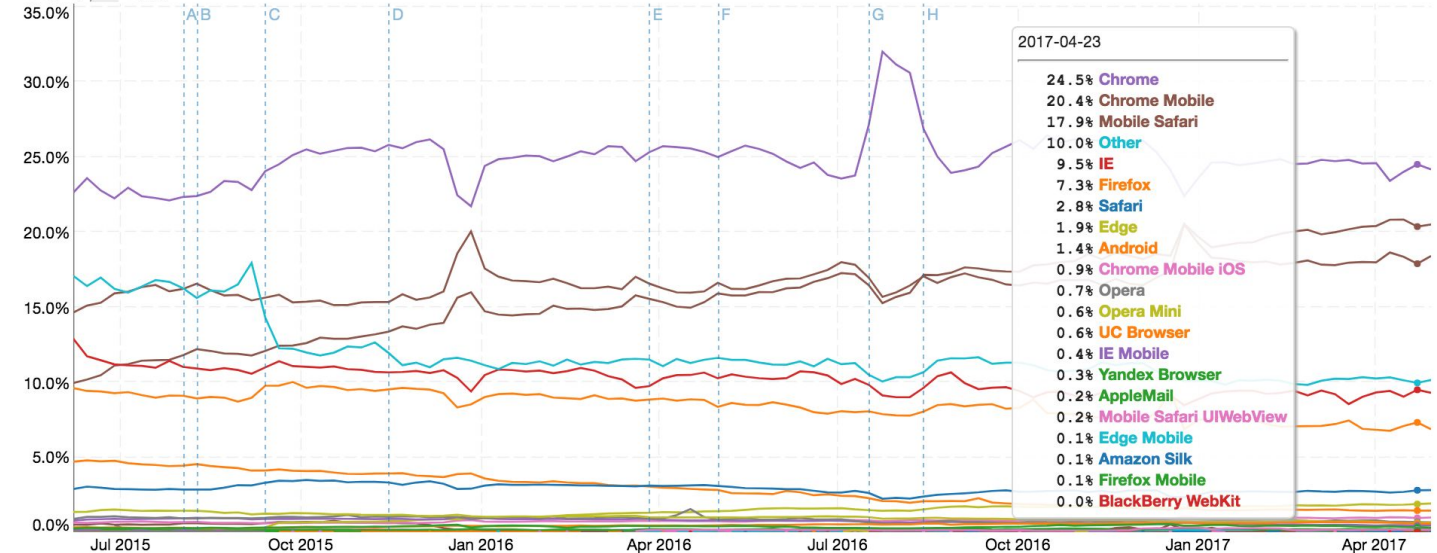
Browser Family Timeseries

Showing data from 2015-06-01 to today

filter by date

- Amazon Silk
- Android
- AOL
- AppleMail
- bingbot
- BingPreview
- BlackBerry WebKit
- Camino
- CFNetwork
- Chrome
- Chrome Frame

Averaging: 1 day(s)



Issues

IE7 making a
comeback... up
more than 1%
last year

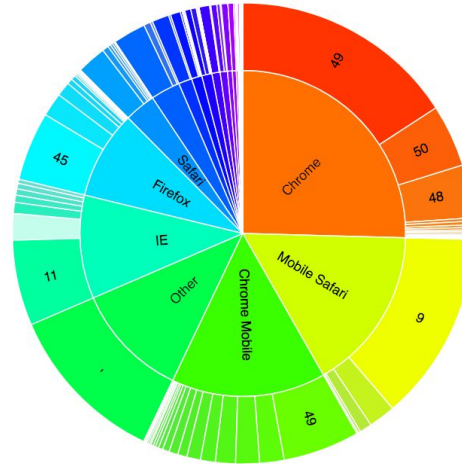
2016-03-06 - 2016-05-11



IE

7

1.8%



2017-03-06 - 2017-05-11

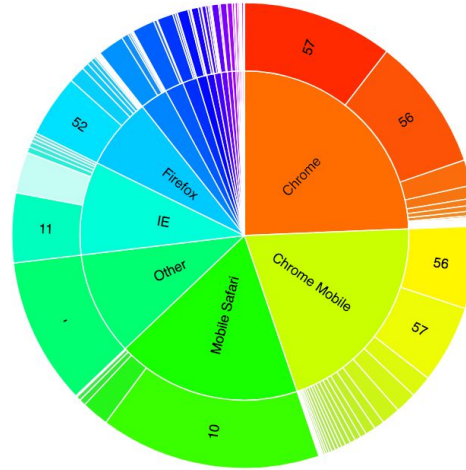


IE 7 2.8%

FROM 2017-03-06 TO 2017-05-11

Apply Cancel

Mar 2017							May 2017						
Su	Mo	Tu	We	Th	Fr	Sa	Su	Mo	Tu	We	Th	Fr	Sa
26	27	28	1	2	3	4	30	1	2	3	4	5	6
5	6	7	8	9	10	11	7	8	9	10	11	12	13
12	13	14	15	16	17	18	14	15	16	17	18	19	20
19	20	21	22	23	24	25	21	22	23	24	25	26	27
26	27	28	29	30	31	1	28	29	30	31	1	2	3
2	3	4	5	6	7	8	4	5	6	7	8	9	10



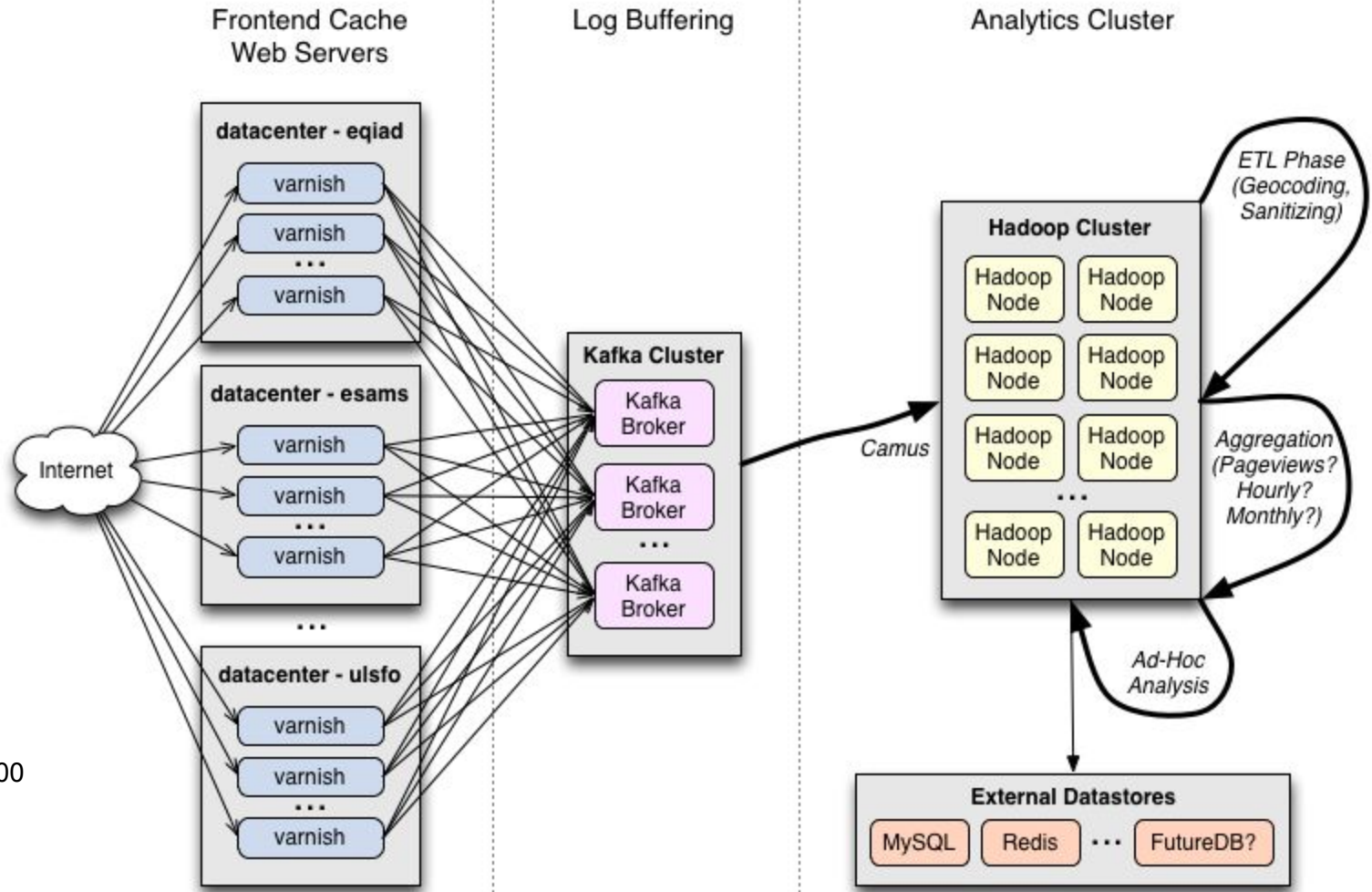
Bots ...

**Data useful to WMF,
Researchers and
Community**

Pageviews

We process about
200,000 HTTP requests / second at peak

At peak we process about 200.000 requests per second



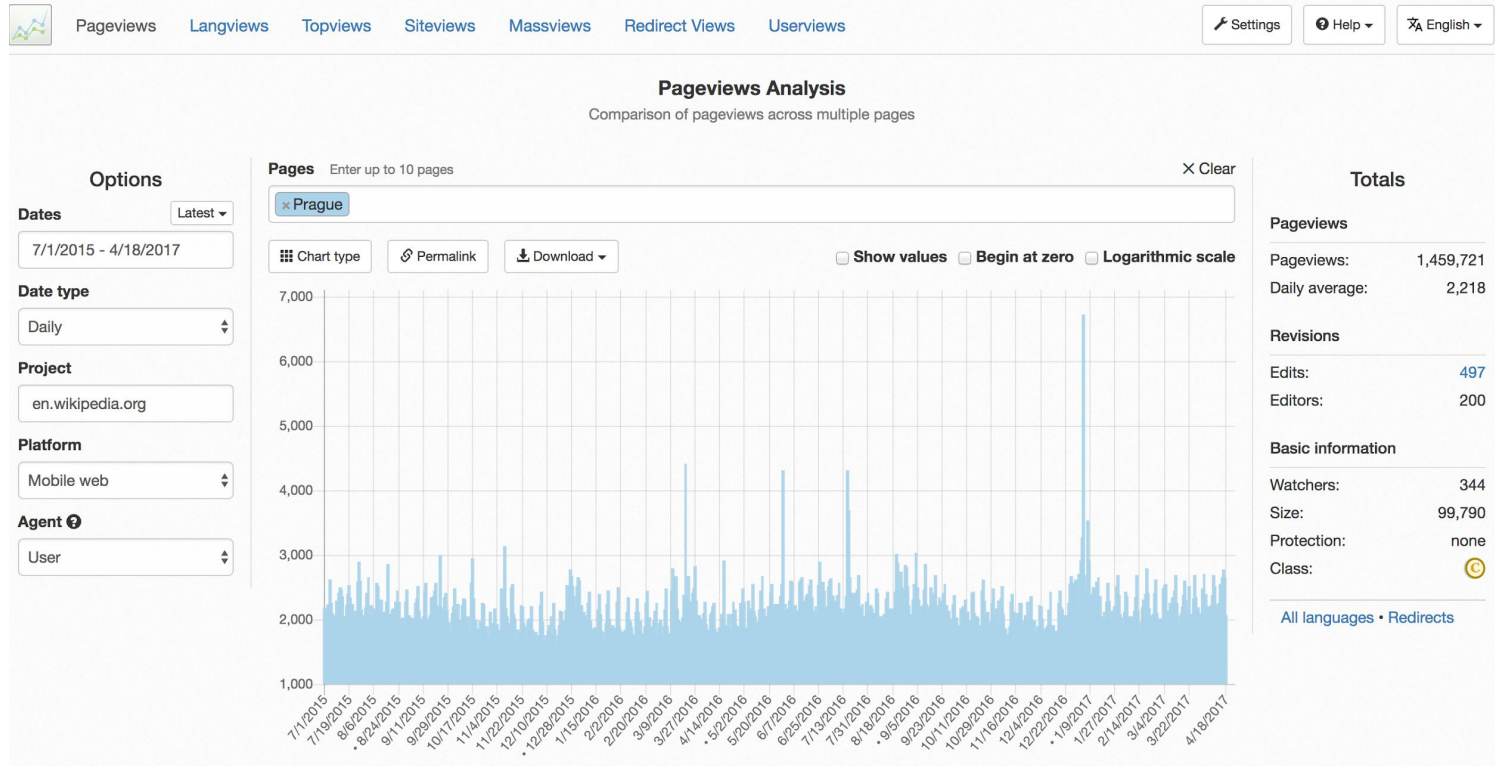
Pageview API

<https://wikitech.wikimedia.org/wiki/Analytics/PageviewAPI>

Get a pageview count time series of en.wikipedia's article
Albert Einstein for the month of October 2015:

[http://wikimedia.org/api/rest_v1/metrics/pageviews/per-article/
en.wikipedia/all-access/all-agents/Albert_Einstein/daily/20151
00100/2015103100](http://wikimedia.org/api/rest_v1/metrics/pageviews/per-article/en.wikipedia/all-access/all-agents/Albert_Einstein/daily/2015100100/2015103100)

<http://tools.wmflabs.org/pageviews/>





Pageviews

Langviews

Topviews

Siteviews

Massviews

Redirect Views

Userviews

Settings

Help

English

Siteviews Analysis

Comparison of total pageviews or unique devices across multiple projects

Options

Dates

Latest 20

4/20/2017 - 5/10/2017

Date type

Daily

Metric

Pageviews

Platform

All

Agent

User

Query for:

Individual projects

All projects

Projects Enter up to 10 projects

X Clear

x tr.wikipedia.org

Chart type

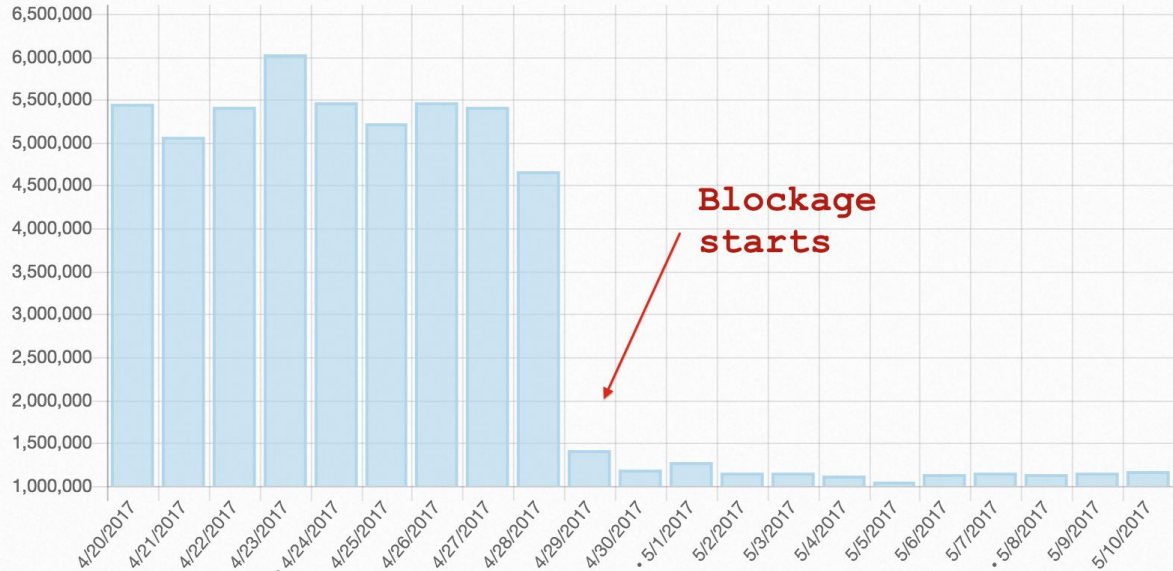
Permalink

Download

Show values

Begin at zero

Logarithmic scale



Totals

Pageviews

Pageviews: 62,315,035

Daily average: 2,967,383

Statistics (all time)

Pages: 1,485,694

Articles: 292,224

Edits: 19,616,279

Images: 29,957

Users: 963,885

Active users: 2,760

Admins: 28

[Most viewed pages](#)

Issues

Bots, Bots, Bots

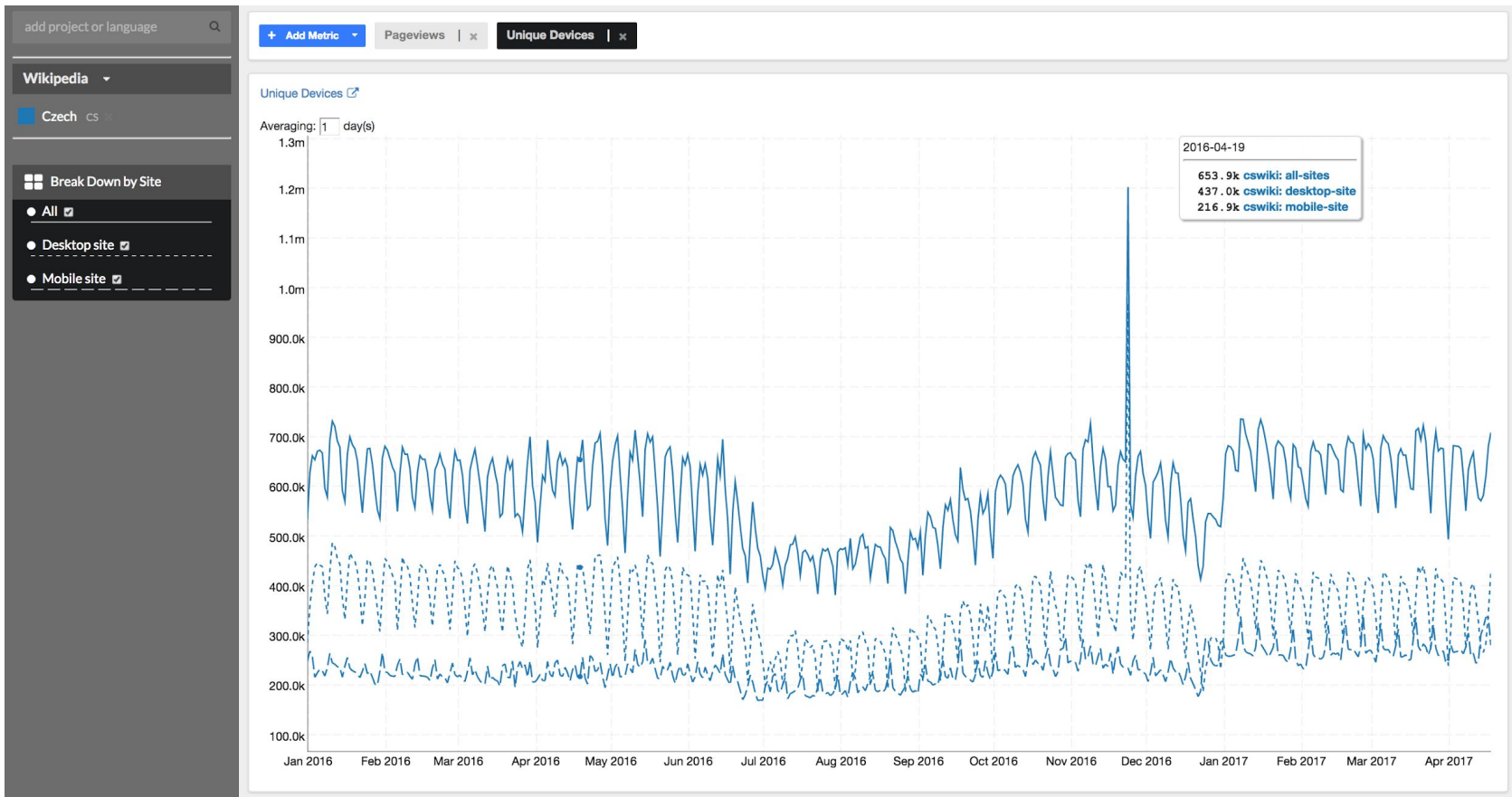
**Data useful to WMF
(mostly)**

Unique Devices

https://wikitech.wikimedia.org/wiki/Analytics/AQS/Unique_Devices

Get the monthly number of unique devices for the mobile version of cs.wikipedia.org for the month of January, February and March 2016:

http://wikimedia.org/api/rest_v1/metrics/unique-devices/cs.wikipedia.org/mobile-site/monthly/20160101/20160301



**Data useful to Community
(mostly)**

Wikistats 2.0

<http://stats.wikimedia.org>

May 2016: The major overhaul of Wikistats reports has entered a new phase.

First phase focused on migrating the traffic analysis reports to our new infrastructure. Those are operational now. The Analytics Team will now proceed to also migrate data collection and reporting about wiki content and contributors. First results are expected later this year.

















More info at [this announcement](#)

You can still tell us which reports you want to see preserved, in this [survey](#).

Wikipedias

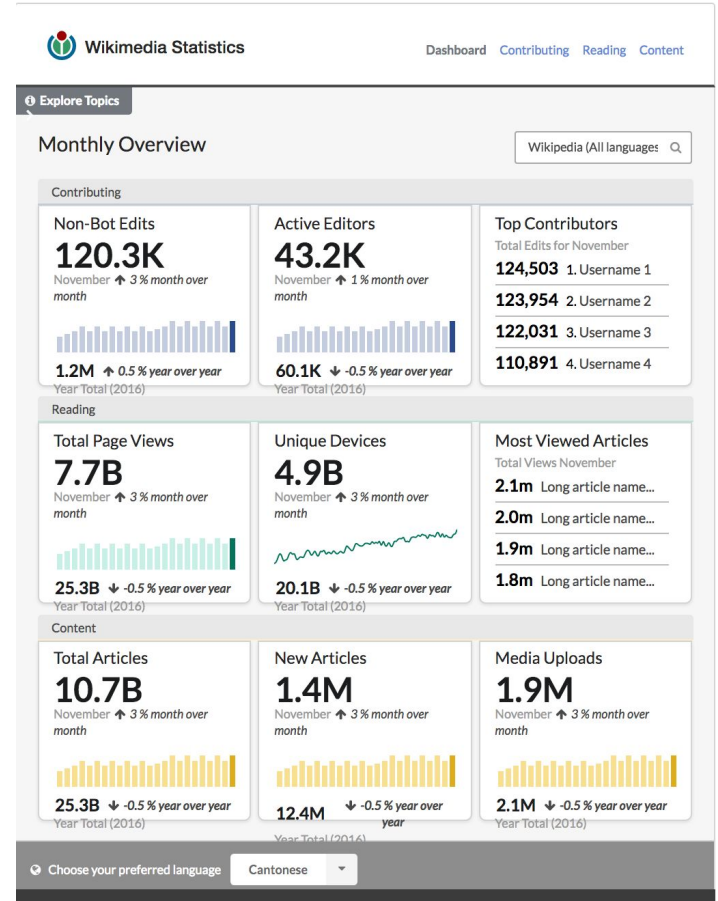
See bottom of page for [language comparisons](#) / [other reports](#) See also the [expanded version of this report](#).

Mar 2016 !New! Extra columns added. Also, want to bookmark this page with default sort column? Now you can! Add url argument 'sortcol=x' (where x is 4-6|8-17), add D for descending sort. E.g. 'Sitemap.htm?sortcol=13D'

Data Month	Languages				Regions 	Participation			Active editors				Edits		Usage	Content
	Code ⇒ Project Main Page	Language ⇒ Wikipedia article	Speakers in millions (log scale) (?) M=millions k=thousands	Editors per million speakers (5+ edits)		Editors (5+ per million speakers	5+ edits p/month (3m avg)	100+ edits p/month (3m avg)	Admins	Bots	Bot edits	Human edits by unreg. users	Views per hour	Article count		
			⇅													
		Tables Charts	Σ	All languages (238)	AF AS EU NA SA OC CL W											
Dec	Summary	Tables Charts	simple	Simple English	AF AS EU NA OC		1500 M	0.1	98	15	17	31	40%	27%	8,799	122,060
Dec	Summary	Tables Charts	en	English	AF AS EU NA OC		1500 M	20	30171	3307	1276	312	9%	32%	5,971,962	5,321,306
Dec	Summary	Tables Charts	zh	Chinese	AS		1300 M	2	2221	302	83	108	20%	24%	298,263	918,069
Dec	Summary	Tables Charts	hi	Hindi	AS		550 M	0.2	93	15	6	82	47%	10%	6,941	117,586
Dec	Summary	Tables Charts	ar	Arabic	AF AS		530 M	1	664	100	34	34	49%	22%	78,110	455,628
Dec	Summary	Tables Charts	es	Spanish	AF AS EU NA SA		500 M	8	4060	545	71	36	19%	38%	606,243	1,305,779
Dec	Summary	Tables Charts	ms	Malay	AS		300 M	0.3	88	10	16	72	57%	21%	5,985	287,259
Dec	Summary	Tables Charts	pt	Portuguese	AF AS EU SA		290 M	5	1544	201	69	212	26%	36%	176,602	951,463
Dec	Summary	Tables Charts	ru	Russian	AS EU		278 M	12	3257	543	86	84	18%	26%	855,466	1,362,550
Dec	Summary	Tables Charts	id	Indonesian	AS		250 M	2	456	43	29	145	54%	23%	60,242	392,730
Dec	Summary	Tables Charts	bn	Bengali	AS		230 M	0.7	164	26	17	16	61%	9%	1,883	46,693
Dec	Summary	Tables Charts	fr	French	AF AS EU NA OC SA		200 M	22	4388	740	162	107	21%	22%	597,412	1,828,237
Dec	Summary	Tables Charts	de	German	EU		185 M	29	5332	871	198	374	11%	21%	841,233	2,015,094
Dec	Summary	Tables Charts	ja	Japanese	AS		132 M	32	4215	395	48	57	10%	41%	657,755	1,043,689
Dec	Summary	Tables Charts	fa	Persian	AS		107 M	7	787	78	34	166	56%	13%	48,868	518,111

Wikistats exists to motivate our editor community.

*In **Wikistats 2.0** we are not only updating the website interface but we are also providing new access to all our edit data in an analytics-friendly form. This much improves (and fundamentally changes) the way, time and resources it takes to calculate edit metrics, for WMF and community.*



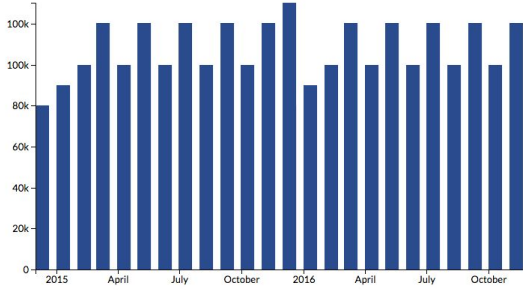
Explore Topics

Wiki
Wikipedia (All lang)

[Add another Wiki](#)

- Metrics
- Non-bot edits
 - Active Editors
 - Top Contributors
- [View more metrics](#)

Non-bot edits Wikipedia



All 2-Year 1-Year 3-Month 1-Month

Total: 2.8m Non-bot edits ↑ 0.5 this year

* Definition of non-bot-edits goes here, pulled from config or maybe dynamically from the wiki page.
More descriptor text will go here assuming that it takes a few sentences to explain a term for a metric.

Choose your preferred language Cantonese

https://www.mediawiki.org/wiki/Wikistats_2.0_Design_Project/RequestforFeedback/Round2

Please Chime in!

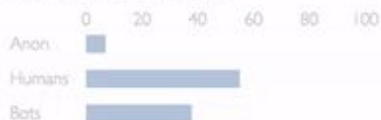
Live Data

EventStreams is a web service that exposes continuous **streams of structured event data**. *Get live updates to Wikimedia projects.*

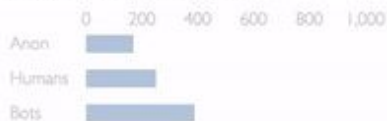


Comparisons

Percentage of Total Edits Made



Average Size Of An Edit (Bytes)



Anonymous Edits

Edits so far:	87	
Additions:	67%	Average Size: 64 bytes
Subtractions:	33%	Average Size: 380 bytes

Registered Human Edits

Edits so far:	686	
Additions:	79%	Average Size: 282 bytes
Subtractions:	21%	Average Size: 133 bytes

Bot Edits

Edits so far:	472	
Additions:	72%	Average Size: 523 bytes
Subtractions:	28%	Average Size: 51 bytes

About

The globe is used to display a real-time random sample of anonymous edits (one every 5 seconds).

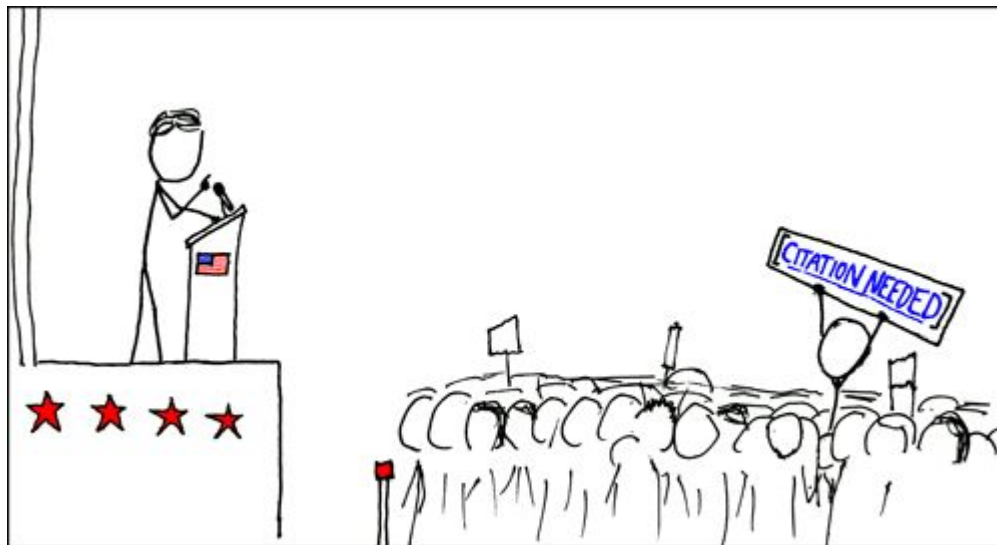
Purple circles represent net additions, red circles represent net subtractions.

The bigger the circle, the larger the edit (circle radius is proportional to the log of the size of the edit in bytes).

Why are only anonymous user edits displayed on the globe? Anonymous users are identified by their IP address, which can be converted to a geolocation. Registered humans and bots are identified by their usernames, which have no geolocation information attached.

Navigate to <http://wikimedia.org> in your browser and open the development console

```
// This is the EventStreams RecentChange stream endpoint
var url = 'https://stream.wikimedia.org/v2/stream/recentchange';
// Use EventSource (available in most browsers, or as an
// npm module: https://www.npmjs.com/package/eventsource)
// to subscribe to the stream.
var recentChangeStream = new EventSource(url);
// Print each event to the console
recentChangeStream.onmessage = function(message) {
    //Parse the message.data string as JSON.
    var event = JSON.parse(message.data);
    console.log(event);
};
```



<https://xkcd.com/285>

Questions?

Most things documented at:

<https://wikitech.wikimedia.org/wiki/Analytics/>

