

# Technology Group

Research and Data, Design Research,  
Analytics Engineering,  
Performance & Availability  
FY Q1: July - September 2016/17

# Quarterly review

## **Research and Data**

### FY Q1: July - September 2016/17

Approximate team size during this quarter:  
5.5 FTE, 2 research fellows, 14 collaborators

# Q1 - Research and Data

Objective: Broaden ORES usage



Objective	Measure of success	Status
<b><i>EXPERIMENT</i></b> <b>Revscore in production</b>	ORES extension deployed to 6 wikis ( <a href="#">T140002</a> )	<b>completed</b> on 8 wikis! (wikidata, fawiki, enwiki, ptwiki, trwiki, nlwiki, plwiki, ruwiki)
<i>Team members involved: 1</i> <i>Collaborators: 2</i>	Release article score dataset for use in ElasticSearch ( <a href="#">T135684</a> )	<b>completed</b> current & historical (enwiki, frwiki, ruwiki)
	Write comprehensive story about ORES ( <a href="#">T140429</a> )	<b>completed</b> announcements & followup discussions

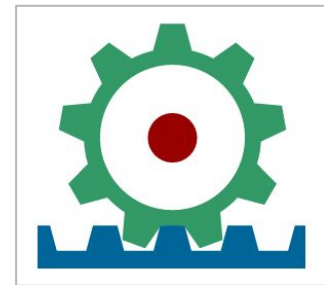
ORES reached production level as a service in Q4. In Q1 we focused on broadening ORES adoption to a larger number of wikis to meet demand for scores. We released article quality datasets for Discovery and the research community broadly.

**Acknowledgments.** Amir Sarabadani, Sabyasachi Ruj

<https://meta.wikimedia.org/wiki/ORES>

### ORES/revision scoring

Hosted a dedicated session at the Product+Tech management onsite to determine **resourcing and long-term maintenance** of the platform.



### Other achievements

**ORES capacity increased** by a factor of 5 ([T143105](#), [T141603](#))

**Substantial performance improvements** for common scoring patterns ([T139408](#))

**New tools** using ORES ([WikiEd drafts](#), [1000 random articles](#), [POPULARLOWQUALITY](#))

Monthly article quality **dataset released** (DOI:[10.6084/m9.figshare.3859800](#))

Explorations into **new signal sources** ([PCFG](#), [HashingVectorization](#))

# Q1 - Research and Data

Objective: Discussion modeling



Objective	Measure of success	Status
<b>FOCUS</b> <b>Discussion Modeling</b>	Design and evaluate attack and aggressiveness models on article talk comments ( <a href="#">T139703</a> )	<b>completed</b>
<i>Team members involved: 1</i> <i>Collaborators: 2</i>	Release notebooks; write up and present results ( <a href="#">T139704</a> )	<b>completed</b>

The first major outputs of the Detox project came to fruition in this quarter.

We'll be continuing work in Q2 using the model to study the impact of harassment and personal attacks on retention.

**Acknowledgments.** Nithum Thain, Lucas Dixon (Jigsaw); Patrick Earley (CE)



<https://meta.wikimedia.org/wiki/Research:Detox>

# Q1 - Research and Data

## Discussion modeling

### Outreach

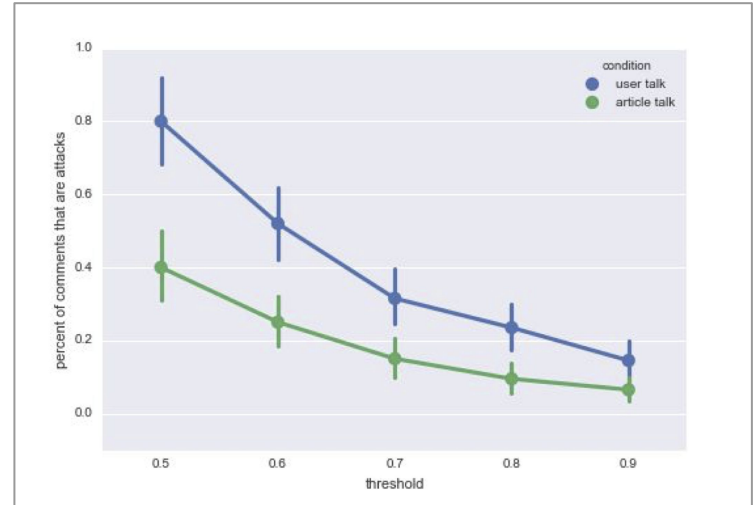
Presentation at *July Research Showcase*

Presentation at *July Monthly Metrics*

Interview and resulting article with *Wired*

### Personal Attack Data Visualization

Prototypes for an interactive application visualizing live and historical personal attacks.



% of comments classified as personal attacks

# Q1 - Research and Data

Objective: Open Notebooks Infrastructure



Objective	Measure of success	Status
<b><i>EXPERIMENT</i></b> <b>Open Notebooks Infrastructure</b>  (T140430)  <i>Team members involved: 0</i> <i>Collaborators: 1</i>	Define, monitor, and ensure PAWS availability of not less than 0.5% less than Labs' NFS availability over a rolling 30da window	completed
	Release a set of notebooks showcasing analysis that can be performed on PAWS by 2nd month of Q1	deferred to Q3
	Publish an announcement and call-to-action targeted at research community in 3rd month of Q1	deferred to Q3

**Acknowledgments.** Yuvi Panda

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





# Q1 - Research and Data

Objective: Productize the article recommendation API



Objective	Measure of success	Status
<b>FOCUS</b> <b>Productize the article recommendation API</b> <i>Team members involved: 3</i> <i>Collaborators: 1</i>	Stable version of production-ready article recommendation API ( <a href="#">T140431</a> )	<b>missed</b>

**Learning.** We made significant progress towards productization of the service but couldn't complete this work in Q1. The set of requirements for bringing research-driven services to productization needs to be sharpened.

**Acknowledgments.** Ori Livneh provided substantial support and mentoring for this work.

# Q1 - Research and Data

## Recommendation API

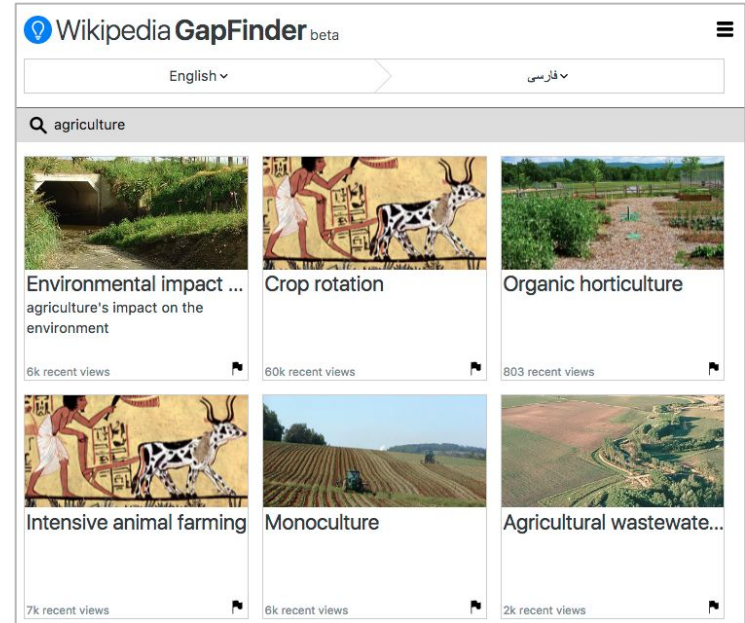
Redefined productization requirements ([T148129](#)):

- request two dedicated VMs ([T148125](#))
- request a service IP and load-balance the two backends ([T148126](#))
- complete security review ([T148133](#))
- deploy using scap3 instead of git ([T148128](#))

Input validation ([T143390](#))

## GapFinder

The application now uses fully responsive design and provides consistent experience across browsers.



[recommend.wmflabs.org](https://recommend.wmflabs.org)

# Q1 - Research and Data

Objective: Research discoverability



Objective	Measure of success	Status
<b><i>STRETCH</i></b> <b>Research Landing Page</b> <i>Team members involved: 1</i>	Complete information architecture, audience analysis and draft of preliminary contents for a landing page for Wikimedia Research	<b>missed</b>

**Learning.** This was a stretch goal for the quarter. Despite making substantial progress, we weren't able to complete it in Q1. Input received during the quarter from multiple stakeholders about findability or research-related initiatives confirms this is still a high priority.

# Q1 - Research and Data

Other successes and misses 

## Other achievements in Q1

### WikiCite

Full **report** and quarterly **newsletter** published:  
[m:WikiCite/Newsletter](https://m:WikiCite/Newsletter)

### outreach

Closing keynote at *VIVO 2016*  
2 invited talks:

*NIH Data Science Lecture series*  
COASP 2016

September Monthly Metrics presentation

### Acknowledgments:

Jonathan Dugan, Anna Filippova, Daniel Mietchen, Cameron Neylon, Lydia Pintscher

{ } wikicite



# Q1 - Research and Data

Category	Workflow	Comments	Type
	NDA's / MOUs	4 new MOUs for <u>research collaborations</u> : <i>ISI Foundation, GESIS, TU Dresden</i>  2 endorsement letters for grant proposals (pending funding decision): <i>University of Washington, University of Pittsburgh</i>	M
	Showcases and talks	Hosted 3 <u>research showcases</u>	M
	Papers submitted	3 paper submissions (CHI 2017, CSCW 2017)	M

- [Research & Data team page](#)
  - Describing goals, processes and projects.
- [Goals for Q2 FY17](#)
  - What we are planning to do in the coming quarter
- [FY17 priorities \(Annual Plan\)](#)
  - Top priorities for the fiscal year
- [Phabricator workboard](#)
  - What we are currently doing (see also our dedicated [project boards](#))

## Q1 collaborators (14)

Lucas Dixon, Jonathan Dugan, Patrick Earley, Anna Filippova, Jure Leskovec, Ori Livneh, Daniel Mietchen, Cameron Neylon, Yuvi Panda, Lydia Pintscher, Sabyasachi Ruj, Amir Sarabadani, Nithum Thain, Robert West.

# Quarterly review

## **Design Research**

### FY Q1: July - September 2016/17

Approximate team size during this quarter: 4 FTE and then 3FTE

# Q1 - Design Research

Objective: Evaluative Design Research



Objective	Measure of success	Status
<p><b>FOCUS</b> Product Research</p> <p><i>Team members involved: 2</i></p>	<p>Deliver design requirements, heuristic evaluations, and user study findings for key products across Reading, Editing, and Discovery.</p>	<p><b>5 research projects studies completed</b> (see next page for details)</p>



## Completed research projects

### Reading Web

- [Reading Web open / closed sections](#)
- [Hovercards usability](#)

### iOS app

- [Random feature experience testing](#)

### Android

- [Android workflows](#)

### Discovery

- [Discovery portal and language drop down assessment](#)

# Q1 - Design Research

## Objective: Generative Design Research



Objective	Measure of success	Status
<p><i>FOCUS</i></p> <p>Generative Research</p> <p><i>Team members involved: 1</i></p> <p><i>Collaborators involved: Reboot teams</i></p>	<p>New Readers</p> <ul style="list-style-type: none"><li>Analyze and share results of Mexico, Nigeria and India contextual inquiries <a href="#">T132799</a></li><li>Concept generation, evaluation and development <a href="#">T129201</a> <a href="#">T132800</a> <a href="#">T132801</a></li></ul>	<ul style="list-style-type: none"><li>Integrated Mexico findings with Nigeria and India findings.</li><li>Cross team prioritization of findings to solve for.</li><li>Cross team concept generation and evaluation for <b>finding #20</b> “People are increasingly getting accessing information online, and consuming and sharing it offline.”</li></ul>

### Learning:

Cross team collaboration provides quick access to diverse perspectives and various forms of expertise during concept development and evaluation. This enables bringing up issues and ideas early in the process, and bringing broader awareness of parallel and informing work around the same topic.

# Q1 - Design Research

## New Readers Research completed and shared:

- **2 day workshop** with Reboot July 13, 14
- **New Readers findings** from India and Nigeria delivered in third week of July
- **August metrics** high level presentation of findings from India and Nigeria
- **Findings from all 3 countries** presented to staff and community September 28, 2016 (led by Communications).
- **Documentation** of New Readers research on **Meta** (documentation is ongoing, check it out for updates!)



# Q1 - Design Research

Objective: Generative Design Research



Objective	Measure of success	Status
<p><i>FOCUS</i></p> <p>Generative Research</p> <p><i>Team members involved: 1</i></p> <p><i>External collaborators: 4</i></p>	<ul style="list-style-type: none"><li>● Define design requirements and use cases for <a href="#">Edit Review Improvements (ERI) T137987</a></li><li>● Evaluate impact, sustainability of community driven new user support (Wikipedia Adventure, <a href="#">Teahouse</a>) <a href="#">T132809</a></li></ul>	<ul style="list-style-type: none"><li>● <b>Completed 9 research interviews</b> with Wikipedians who work with new editors @ the Teahouse, AfC, PageCuration, FeedbackDashboard</li><li>● <b>Provided design and evaluation guidance</b> for ERI feed project (w/ Collaboration team)</li><li>● <b>Completed impact evaluation</b> of Wikipedia Adventure (w/ collaborators at Northwestern, UW)</li><li>● <b>Completed sustainability evaluation</b> of Wikipedia Teahouse (w/ collaborator at Carnegie Mellon)</li></ul>

## Learning:

- Speaking with target users early about their workflows, motivations, and challenges can help validate key concepts, challenge assumptions before development resources are committed.
- Evaluating the success of previous WMF onboarding interventions can inform product strategy and help us understand new editor experience and community health.

# Q1 - Design Research

Objective: Research Capacity Building



Objective	Measure of success	Status
<p><i>STRENGTHEN</i></p> <p>Research capacity building</p> <p><i>Team members involved: 4</i></p>	<ul style="list-style-type: none"><li>• Implement user testing tools to support moderated and unmoderated user testing on desktop and mobile (apps and web)</li><li>• Experiment with recruiting research participants through Wikimedia/Wikipedia social media channels</li></ul>	<p><b>Tooling:</b></p> <ul style="list-style-type: none"><li>• Support product UX researchers to debug and document UserZoom processes</li></ul> <p><b>Recruiting:</b></p> <ul style="list-style-type: none"><li>• <b>Continued recruiting via Social media</b> in collaboration with Jeff Elder in Communications.</li><li>• Samantha created a <u>plan</u> to implement a participant outreach campaign</li><li>• Recruiting for 3 research projects.</li><li>• Samantha is on medical leave (August 24 - present)</li><li>• To address her absence, the DR team all pitched in to support each other with recruiting needs.</li></ul>

# Q1 - Design Research

## Began using UserZoom:

- Initial training with team
- Created [page on Office wiki](#) describing use of UserZoom and FAQ
- Used in hover cards usability study
- Used participants on the UZ panel
- Addressed some issues with UZ

# Q1 - Design Research

## Objective: Research Data Mapping



Objective	Measure of success	Status
<p><i>EXPERIMENT</i></p> <p>Research data mapping (stretch)</p> <p><i>Team members involved: 3</i></p>	<ul style="list-style-type: none"><li>● Document the sources of user data we manage, how we store it, and what potentially personally identifiable information it contains.</li><li>● Ensure all archived user study data is retained and shared in compliance with WMF data retention policy and privacy policy.</li><li>● Draft data access guidelines to inform our future data collection and dissemination practices.</li></ul>	<p><b>Completed:</b></p> <ul style="list-style-type: none"><li>● <b>Summarize data collected by UserZoom</b></li><li>● <b>Created a plan for UZ data destruction</b> in compliance with policy.</li><li>● Began prototyping process for releasing <b>New Readers research corpus via OA policy.</b></li><li>● <b>Assessment of all DR surveys</b>, where the data is, and what can be removed, destroyed, retained.</li><li>● Trello workboards clean up</li></ul>

# Q1 - Design Research

Category	Workflow	Comments	Type
	Participant recruiting	Recruiting for three evaluative research projects. Samantha is on medical leave (August 24 - present). To address her absence, the DR team all pitched in to support each other with recruiting needs.	M

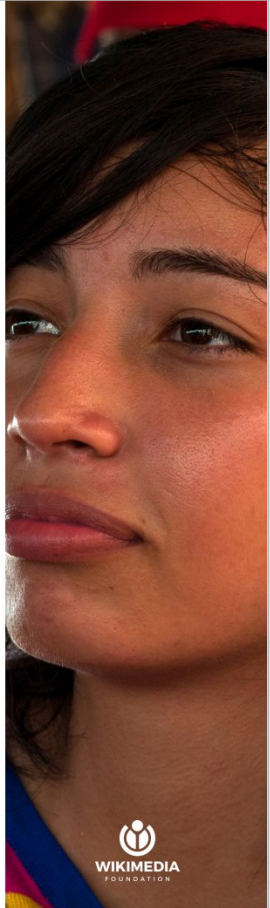


# Q1 - Design Research

## Other achievements in Q4

### Personas:

- Created persona from Mexico research >
- Personas being used in concept development, evaluation and prototyping for New Readers
- Met with Reading UX team and handed off iterating the pragmatic Reading personas (Michelle and Sandra)
- Worked with Zack and Blanca in Communications on a consistent template for use with all of the personas.






### Ximena

**New Reader Mexico, 20**  
Puebla, Mexico  
Student about to graduate from Universidad de las Américas Puebla  
Fluent in Spanish and speaks limited English (moderately conversational)

**TECHNOLOGY COMFORT LEVEL**

Not at all	Comfortable	Very	Extremely
Always needs assistance	Sending emails, browsing internet using software	Downloading & setting up own software	Advanced programming, building own software

**DEVICES USED**


 Android (refurbished)     Tablet (refurbished)     Vio Laptop

**GOALS**

- Get into grad school in a place relevant to biochem
- Establish her career
- Improve English competence
- Find content in Spanish
- Fast frictionless translation

**CHALLENGES**

- Limited English competence and limited relevant content in Spanish
- Constant access at home (wifi), less stable, slow outside of home (school, transit)



[https://commons.wikimedia.org/wiki/File:Maracabo\\_typical\\_girl.jpg](https://commons.wikimedia.org/wiki/File:Maracabo_typical_girl.jpg)

- **Design Research team page**
  - Describing processes and projects.
- **Goals for Q2 FY17**
  - Plans for next quarter
- **Phabricator workboard**
  - What we are currently doing

# Quarterly review

## **Analytics Engineering**

### FY Q1: July - September 2016/17

Approximate team size during this quarter 5.5 FTE (2 devops)  
and 1 PT

Key performance indicator: Velocity

Quarter: 719	April: 230	May: 288	June: 211
--------------	------------	----------	-----------

# Q1 - Analytics Engineering

Objective: Operational Excellence

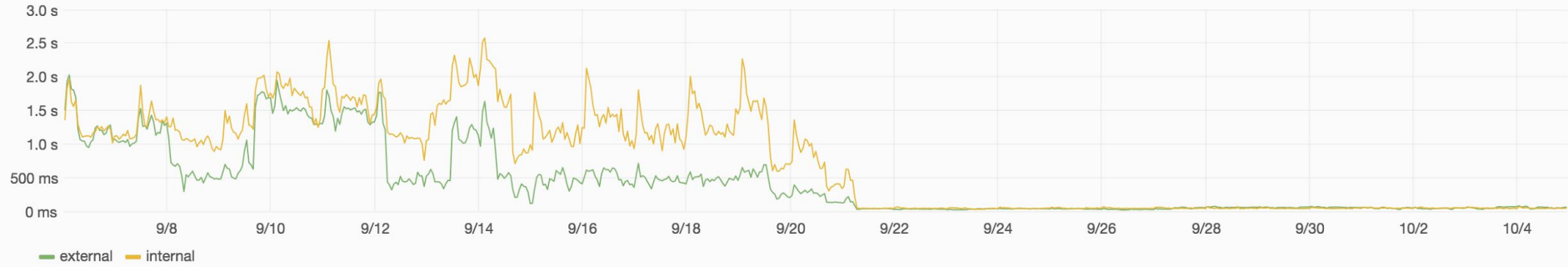


Objective	Measure of success	Status
Better response times pageview API	Pageview API can sustain with much lower latencies a higher number of fresh requests	Done

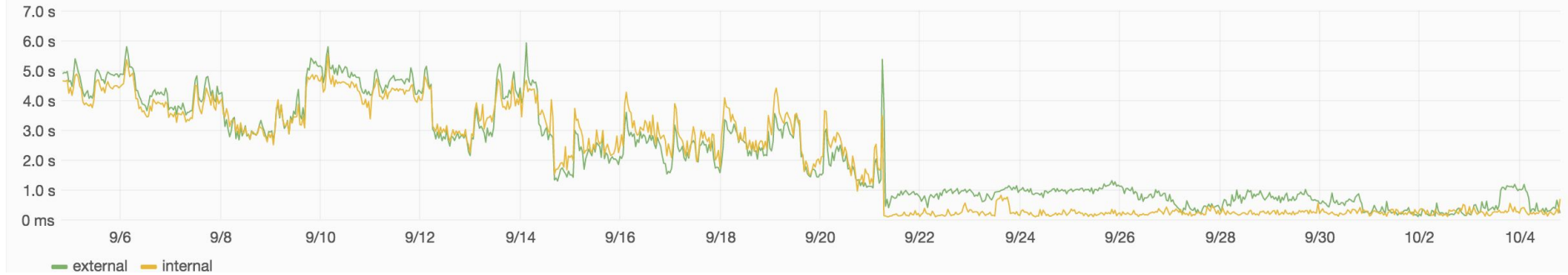
## Learning:

Carry-on goal from last quarter. **Scaling took three months longer than anticipated**, the bulk of the time went into having to load 6TB + of data into a new set of nodes. Anticipating needs when it comes to hardware and software for Cassandra will have avoided much of this work.

Per article latency (mean) breakdown



Per article latency (p99) breakdown



# Q1 - Analytics Engineering

Objective: Better tools to access data



Objective	Measure of success	Status
Productionize Druid Pageview Pipeline and UI (pivot) on Druid	<a href="http://pivot.wikimedia.org">http://pivot.wikimedia.org</a> working on top of pageview data	Not Done (*)

(\*) It was completed done couple weeks into Q2.

- DIMENSIONS**
- Time
  - Access Method
  - Agent Type
  - Continent
  - Country
  - Country Code
  - Language Variant
  - Project
  - Referer Class
  - Ua Browser Family
  - Ua Browser Major
  - Ua Os Family
  - Ua Os Major
  - Ua Os Minor
  - Ua Wmf App Version
- MEASURES**
- View Count

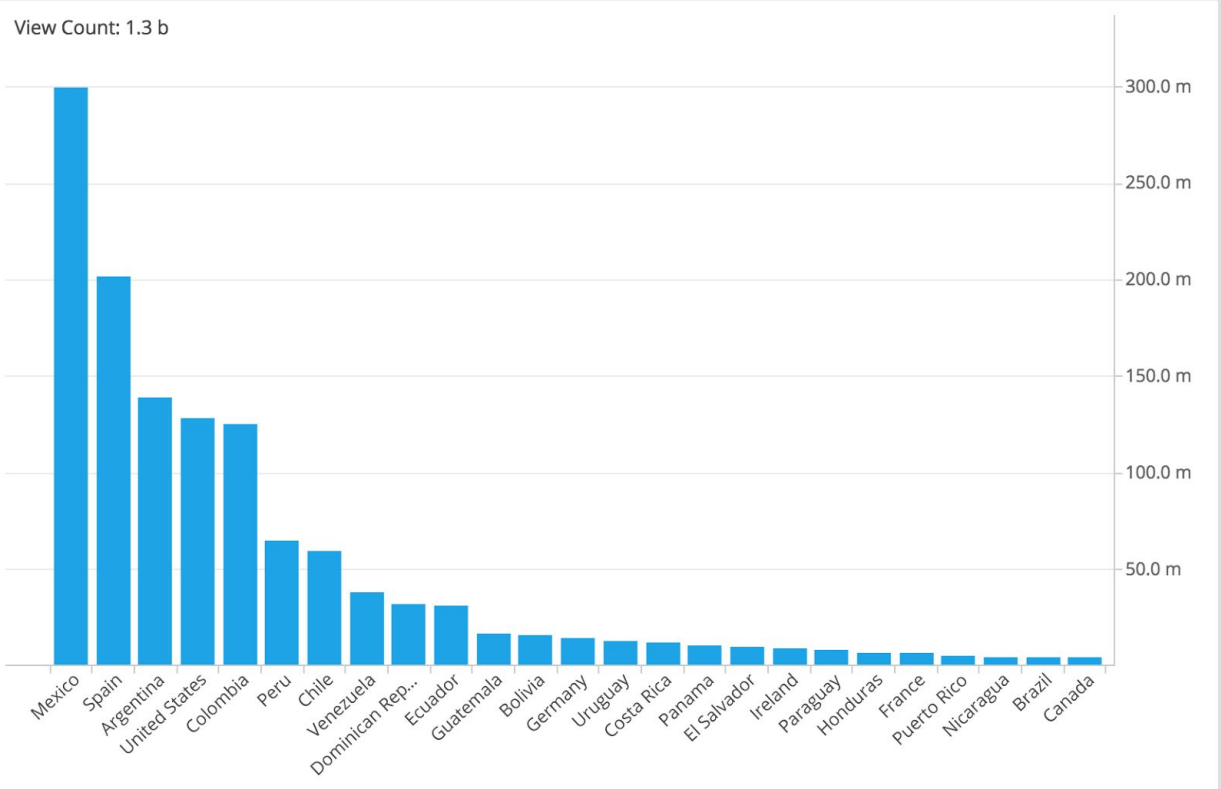
**FILTER** Aug 31 - Sep 29 Project: es.wikipedia

**SPLIT** Country

Bar Chart

**PINBOARD** View Count

Click or drag dimensions to pin them



DIMENSIONS

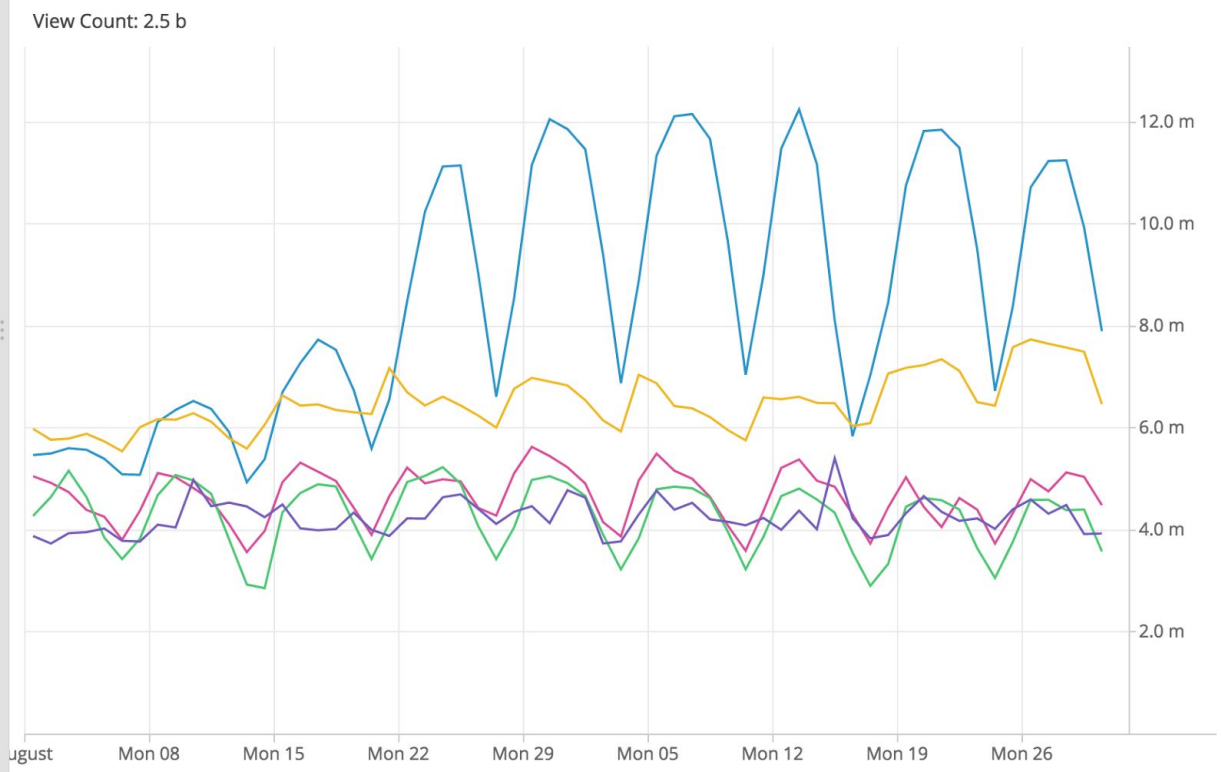
- Time
- Access Method
- Agent Type
- Continent
- Country
- Country Code
- Language Variant
- Project
- Referer Class
- Ua Browser Family
- Ua Browser Major
- Ua Os Family
- Ua Os Major
- Ua Os Minor
- Ua Wmf App Version

MEASURES

- View Count

FILTER Aug 1 - Oct 1 Project: es.wikipedia ×

SPLIT Country × Time (Day) ×



LEGEND View Count ▾

COUNTRY

<span style="color: #0070c0;">■</span> Mexico	523,313.3 k
<span style="color: #ffc000;">■</span> Spain	396,041.9 k
<span style="color: #e91e63;">■</span> Argentina	284,400.0 k
<span style="color: #4caf50;">■</span> Colombia	259,937.9 k
<span style="color: #673ab7;">■</span> United States	258,649.3 k
<span style="color: #bdbdbd;">■</span> Peru	124,355.4 k
<span style="color: #bdbdbd;">■</span> Chile	123,225.9 k
<span style="color: #bdbdbd;">■</span> Venezuela	76,426.4 k
<span style="color: #bdbdbd;">■</span> Ecuador	59,163.0 k
<span style="color: #bdbdbd;">■</span> Dominican Repul	47,415.8 k

[Show selection](#)

PINBOARD View Count ▾

*Click or drag dimensions to pin them*



# Q1 - Analytics Engineering

Objective	Measure of success	Status
<b>Wikistats 2.0 (ongoing goal).</b> Remove dependency of dumps as sources of edit data to be able to replace wikistats <b>edit</b> reports.	<b>Reconstructed edit history for simplewiki and enwiki using a more productionised version of our proof on concept last quarter.</b>	<b>Done.</b>

## Learning:

Reconstructing edit history for enwiki is the most complex problem (in terms of scale and algorithms) that the team has tackled to date.

**A byproduct of this work is the ability to generate dumps from data on hdfs.**

# Q4 - Analytics Engineering

Objective	Measure of success	Status
Public Event Stream POC	Event Stream Endpoint operational. POC that makes arbitrary JSON events available for public consumption from MediaWiki changes to publish data to fulfil schemas task	Done.

## Learning:

Many takers for this project among Product teams but infrastructure is not productionised yet.

# Q1 - Analytics Engineering

Other successes and misses 

## Operational Excellence

Collaborated with traffic team on varnish upgrades

Kafka: Upgrade Kafka on non-analytics cluster

Kafka: Get mirrormaker puppetized

Better deployment to hadoop cluster

## Blogpost

Announcement of browser dashboards:

<https://blog.wikimedia.org/2016/08/19/most-popular-browser/>

## Collaborations


Compiled a cached dataset to be used for cache tuning for java JVM among others

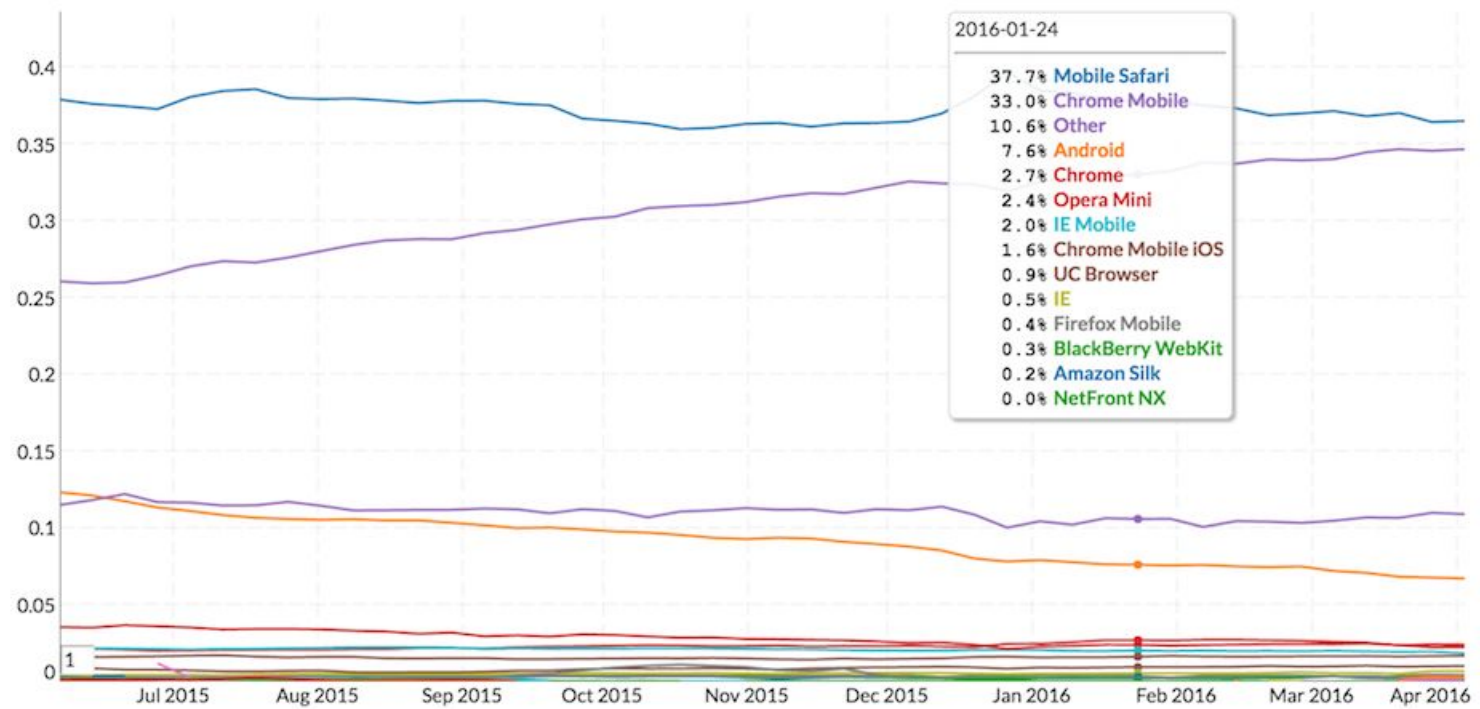
Zika:[https://meta.wikimedia.org/wiki/Research:Quantifying\\_the\\_global\\_attention\\_to\\_public\\_health\\_threats\\_through\\_Wikipedia\\_pageview\\_data](https://meta.wikimedia.org/wiki/Research:Quantifying_the_global_attention_to_public_health_threats_through_Wikipedia_pageview_data)

# 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 ▾**

## Browser Family Timeseries

- filter by date 
- Amazon Silk
  - Android
  - BlackBerry WebKit
  - Chrome
  - Chrome Mobile
  - Chrome Mobile iOS
  - Edge Mobile
  - Firefox Mobile
  - IE
  - IE Mobile
  - Mobile Safari



## Visits Over Time



**Jun 27 - Jul 3, 2016**  
**155 Unique visitors**



ANNOTATIONS

26



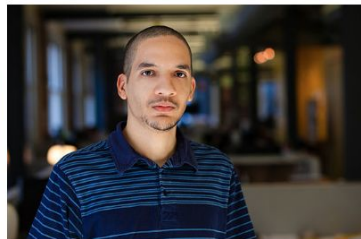
# Quarterly review

## PERFORMANCE TEAM

### FY Q1: July - September 2016/17



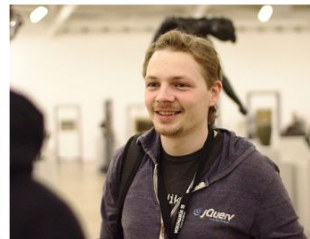
Ori Livneh



Aaron Schulz



Gilles Dubuc



Timo Tijhof

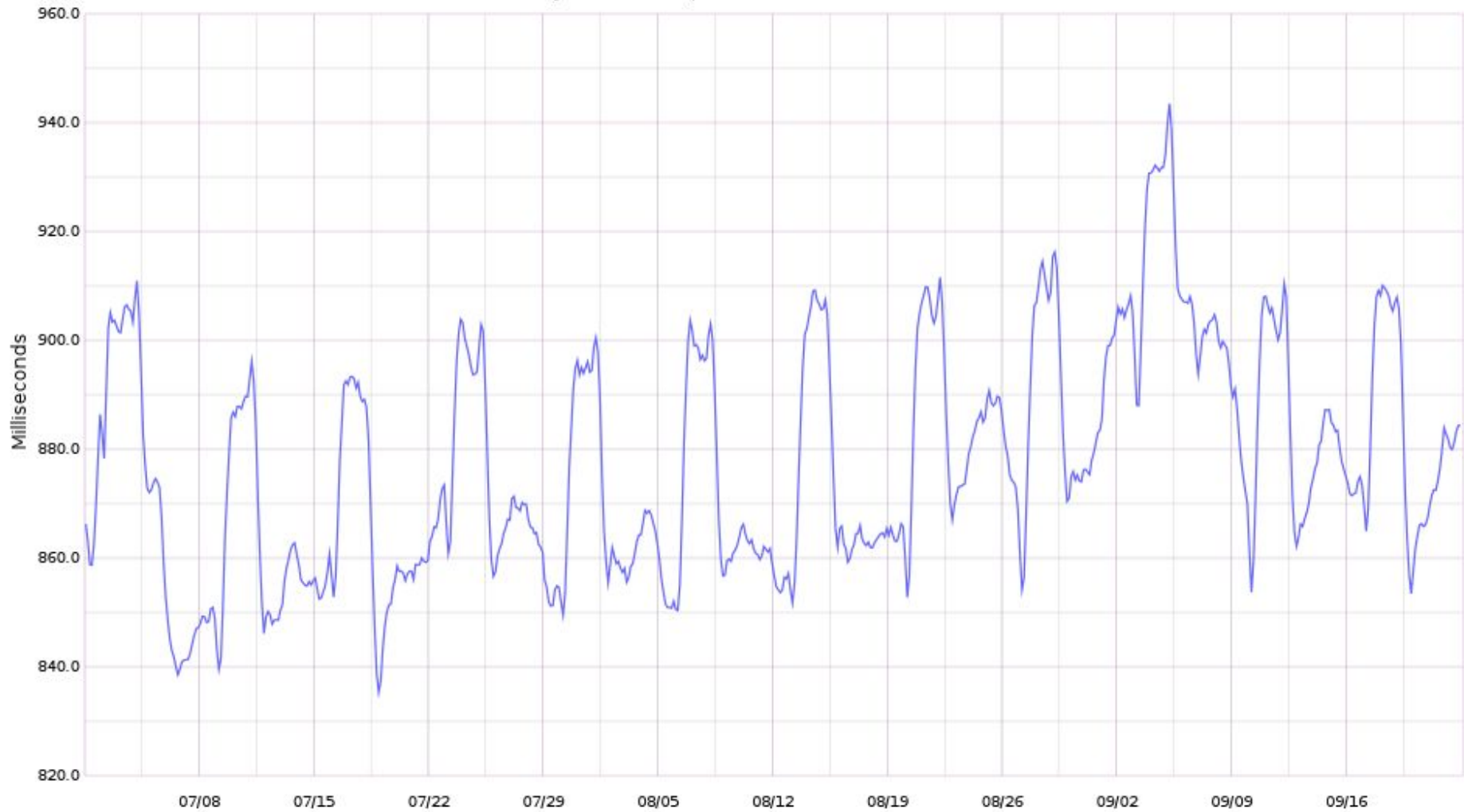


Peter Hedenskog

Daily median save timing in milliseconds, 06/01 - 10/19



Daily median first paint time, 07/01 - 09/30





# Q1 - Performance

Objective: Thumbor



Objective	Measure of success	Status
<b>Deploy Thumbor</b>	Package Thumbor and plugins for production; add instrumentation and logging; deploy to production; shadow-serve all production traffic.	<b>Done</b>  Will begin serving live traffic by the end of Q2.

## Learning:

Focused cross-team partnerships between developers and ops work really well. If you provide opportunities for operational experience to inform the development process, the end result is more resilient and easier to maintain.

# Q1 - Performance

Objective: Performance Inspector



Objective	Measure of success	Status
<b>On-wiki performance inspector tool</b>	Editors can see the impact of an article page's assets (modules, custom styles, total size, images) on page load time and data cost, and use that to make informed decisions about content and presentation.	<b>Not done</b>  Feature itself is ready, but still hasn't cleared the process for becoming a beta feature.

## Learning:

Developers without substantial experience as community members can't skimp on product, community, and design research support.



Objective	Measure of success	Status
<b>Optimise critical rendering path.</b>	<p>Make content load more quickly by:</p> <ul style="list-style-type: none"><li>- Inlining above-the-fold CSS (<a href="#">T124966</a>)</li><li>- Making module execution via <code>mw.loader.work()</code> asynchronous (<a href="#">T142129</a>)</li><li>- Preventing user and site CSS from loading twice (<a href="#">T108590</a>)</li></ul>	<p><b>Not done.</b></p> <p>Loading of user and site CSS deduplicated. Reusable style modules introduced, allowing OOUI to be used throughout MediaWiki core and extensions. Async execution of stored modules went out but had to be reverted. Revised patch going out next week. Critical CSS still not inlined; was blocked on async module execution.</p>

# Q1 - Performance

Objective: Multi-DC



Objective	Measure of success	Status
<b>Cut latency and improve resilience of core MediaWiki stack</b>	<p>Serve MediaWiki requests from more than one datacenter.</p> <p>Make MediaWiki properly detect master replication lag when:</p> <ul style="list-style-type: none"><li>- Master database is unreachable;</li><li>- Replication is delayed by network conditions;</li><li>- Replicating from another slave.</li></ul> <p>(<a href="#">T111266</a>)</p> <p>Use <code>MASTER_GTID_WAIT()</code> to ensure reads from slaves that happen after a write to a master are consistent. (<a href="#">T135027</a>)</p>	<b>Ongoing.</b>  Quarterly subgoals complete.