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**

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EXPERIMENT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revscoreing in production</td>
<td>ORES extension deployed to 6 wikis (T140002)</td>
<td>completed on 8 wikis! (wikidata, fawiki, enwiki, ptwiki, trwiki, nlwiki, plwiki, ruwiki)</td>
</tr>
<tr>
<td>Team members involved: 1 Collaborators: 2</td>
<td>Release article score dataset for use in ElasticSearch (T135684)</td>
<td>completed current &amp; historical (enwiki, frwiki, ruwiki)</td>
</tr>
<tr>
<td></td>
<td>Write comprehensive story about ORES (T140429)</td>
<td>completed announcements &amp; followup discussions</td>
</tr>
</tbody>
</table>

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](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)

https://meta.wikimedia.org/wiki/ORES
Q1 - Research and Data

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FOCUS</strong> Discussion Modeling</td>
<td>Design and evaluate attack and aggressiveness models on article talk comments</td>
<td>completed</td>
</tr>
<tr>
<td></td>
<td>(T139703)</td>
<td></td>
</tr>
<tr>
<td><strong>Team members involved:</strong></td>
<td>Release notebooks; write up and present results (T139704)</td>
<td>completed</td>
</tr>
<tr>
<td>1 Collaborators: 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

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

% of comments classified as personal attacks
## Objective: Open Notebooks Infrastructure

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

**Acknowledgments.** Yuvi Panda

[https://wikitech.wikimedia.org/wiki/PAWS](https://wikitech.wikimedia.org/wiki/PAWS)
Q1 - Research and Data

Open Notebooks Infrastructure

Revisited timeline

Due to competing priorities, Yuvi’s bandwidth to work on the project in Q1 was limited. We decided that additional end-user documentation work is needed in preparation for the official launch. Finally, the availability of new hardware and fully configured databases will require additional time. As a result of these factors, we decided with Ops to defer the official launch of the platform to Q3 (January-March 2017).

Outreach

In Q1 we continued our outreach efforts to identify early adopters. We worked with professors at University of Washington and University of Colorado Boulder to use PAWS with students in their classes.
**Q1 - Research and Data**

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FOCUS</strong></td>
<td>Stable version of production-ready article recommendation API (T140431)</td>
<td>missed</td>
</tr>
</tbody>
</table>

**Objective: Productize the article recommendation API**

*Team members involved: 3*  
*Collaborators: 1*

**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
## Q1 - Research and Data

**Objective: Research discoverability**

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>STRETCH</strong></td>
<td>Complete information architecture, audience analysis and draft of preliminary contents for a landing page for Wikimedia Research</td>
<td>missed</td>
</tr>
</tbody>
</table>

**Research Landing Page**

*Team members involved: 1*

---

**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 achievements in Q1

WikiCite

Full report and quarterly newsletter published: 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
<table>
<thead>
<tr>
<th>Category</th>
<th>Workflow</th>
<th>Comments</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDAs / MOUs</td>
<td></td>
<td>4 new MOUs for research collaborations: ISI Foundation, GESIS, TU Dresden</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 endorsement letters for grant proposals (pending funding decision):</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>University of Washington, University of Pittsburgh</td>
<td></td>
</tr>
<tr>
<td>Showcases and talks</td>
<td></td>
<td>Hosted 3 research showcases</td>
<td>M</td>
</tr>
<tr>
<td>Papers submitted</td>
<td></td>
<td>3 paper submissions (CHI 2017, CSCW 2017)</td>
<td>M</td>
</tr>
</tbody>
</table>

Type: new, reactive, maintenance
Q1 - Research and Data

- **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)**
Quarterly review
Design Research
FY Q1: July - September 2016/17

Approximate team size during this quarter: 4 FTE and then 3FTE
### Objective: Evaluative Design Research

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FOCUS</strong> Product Research</td>
<td>Deliver design requirements, heuristic evaluations, and user study findings for key products across Reading, Editing, and Discovery.</td>
<td>5 research projects studies completed (see next page for details)</td>
</tr>
<tr>
<td>Team members involved: 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Q1 - Design Research

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
### Objective: Generative Design Research

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
</table>
| **FOCUS** Generative Research | New Readers  
  - Analyze and share results of Mexico, Nigeria and India contextual inquiries [T132799](#)  
  - Concept generation, evaluation and development [T129201](#) [T132800](#) [T132801](#) |  
  - Integrated Mexico findings with Nigeria and India findings.  
  - Cross team prioritization of findings to solve for.  
  - Cross team concept generation and evaluation for **finding #20** “People are increasingly getting accessing information online, and consuming and sharing it offline.” |

#### 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

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FOCUS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generative Research</td>
<td>● Define design requirements and use cases for <em>Edit Review Improvements</em> (ERI) <a href="#">T137987</a></td>
<td><strong>Completed 9 research interviews</strong> with Wikipedians who work with new editors @ the Teahouse, AfC, PageCuration, FeedbackDashboard</td>
</tr>
<tr>
<td></td>
<td>● Evaluate impact, sustainability of community driven new user support (Wikipedia Adventure, Teahouse) <a href="#">T132809</a></td>
<td><strong>Provided design and evaluation guidance</strong> for ERI feed project (w/ Collaboration team)</td>
</tr>
<tr>
<td>Team members involved: 1</td>
<td></td>
<td><strong>Completed impact evaluation</strong> of Wikipedia Adventure (w/ collaborators at Northwestern, UW)</td>
</tr>
<tr>
<td>External collaborators: 4</td>
<td></td>
<td><strong>Completed sustainability evaluation</strong> of Wikipedia Teahouse (w/ collaborator at Carnegie Mellon)</td>
</tr>
</tbody>
</table>

### 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

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
</table>
| **STRENGTHEN** Research capacity building | - Implement user testing tools to support moderated and unmoderated user testing on desktop and mobile (apps and web)  
- Experiment with recruiting research participants through Wikimedia/Wikipedia social media channels | **Tooling:**  
- Support product UX researchers to debug and document UserZoom processes  
**Recruiting:**  
- **Continued recruiting via Social media** in collaboration with Jeff Elder in Communications.  
- Samantha created a plan to implement a participant outreach campaign  
- Recruiting for 3 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. |

Team members involved: 4

### Learning:
Consistent recruiting with social media both contributes to growing the database of participants as well as functions well for recruiting for specific research projects.
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

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EXPERIMENT</strong> Research data mapping (stretch)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| *Team members involved: 3* | ● Document the sources of user data we manage, how we store it, and what potentially personally identifiable information it contains.  
● Ensure all archived user study data is retained and shared in compliance with WMF data retention policy and privacy policy.  
● Draft data access guidelines to inform our future data collection and dissemination practices. | **Completed:**  
● Summarize data collected by UserZoom  
● Created a plan for UZ data destruction in compliance with policy.  
● Began prototyping process for releasing New Readers research corpus via OA policy.  
● Assessment of all DR surveys, where the data is, and what can be removed, destroyed, retained.  
● Trello workboards clean up |
## Core workflows and metrics

<table>
<thead>
<tr>
<th>Category</th>
<th>Workflow</th>
<th>Comments</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Participant recruiting</td>
<td>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.</td>
<td>M</td>
</tr>
</tbody>
</table>

**Type:** new, reactive, maintenance
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.
Q1 - Design Research

- **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

<table>
<thead>
<tr>
<th>Quarter: 719</th>
<th>April: 230</th>
<th>May: 288</th>
<th>June: 211</th>
</tr>
</thead>
</table>

All content of these slides is (c) Wikimedia Foundation and available under a CC BY-SA 3.0 license, unless noted otherwise.
**Q1 - Analytics Engineering**

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Better response times pageview API</td>
<td>Pageview API can sustain with much lower latencies a higher number of fresh requests</td>
<td>Done</td>
</tr>
</tbody>
</table>

**Objective: Operational Excellence**

**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.
## Q1 - Analytics Engineering

### Objective: Better tools to access data

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

(*) It was completed done couple weeks into Q2.
View Count: 2.5 b
## Q1 - Analytics Engineering

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

**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

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Event Stream POC</td>
<td>Event Stream Endpoint operational. POC that makes arbitrary JSON events available for public consumption from MediaWiki changes to publish data to fulfil schemas task</td>
<td>Done.</td>
</tr>
</tbody>
</table>

**Learning:**

Many takers for this project among Product teams but infrastructure is not productionised yet.
Q1 - Analytics Engineering

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
Visits Over Time

155 Unique visitors
Quarterly review

PERFORMANCE TEAM

FY Q1: July - September 2016/17
Daily median save timing in milliseconds, 06/01 - 10/19
Daily median first paint time, 07/01 - 09/30
## Q1 - Performance

### Objective: Thumbor

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

### 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.
### Objective: Performance Inspector

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>On-wiki performance inspector tool</td>
<td>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.</td>
<td>Not done Feature itself is ready, but still hasn't cleared the process for becoming a beta feature.</td>
</tr>
</tbody>
</table>

**Learning:**

Developers without substantial experience as community members can't skimp on product, community, and design research support.
## Q1 - Performance

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimise critical rendering path.</td>
<td>Make content load more quickly by:</td>
<td>Not done.</td>
</tr>
<tr>
<td></td>
<td>- Inlining above-the-fold CSS  [T124966]</td>
<td>Loading of user and site CSS deduplicated.</td>
</tr>
<tr>
<td></td>
<td>- Making module execution via <code>mw.loader.work()</code> asynchronous  [T142129]</td>
<td>Reusable style modules introduced, allowing</td>
</tr>
<tr>
<td></td>
<td>- Preventing user and site CSS from loading twice  [T108590]</td>
<td>OOUI to be used throughout MediaWiki core</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and extensions. Async execution of stored</td>
</tr>
<tr>
<td></td>
<td></td>
<td>modules went out but had to be reverted.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Revised patch going out next week. Critical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CSS still not inlined; was blocked on async</td>
</tr>
<tr>
<td></td>
<td></td>
<td>module execution.</td>
</tr>
</tbody>
</table>
## Q1 - Performance

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measure of success</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cut latency and improve resilience of core MediaWiki stack</strong></td>
<td>Serve MediaWiki requests from more than one datacenter. Make MediaWiki properly detect master replication lag when: - Master database is unreachable; - Replication is delayed by network conditions; - Replicating from another slave. <em>(T111266)</em> Use MASTER_GTID_WAIT() to ensure reads from slaves that happen after a write to a master are consistent. <em>(T135027)</em></td>
<td><strong>Ongoing.</strong> Quarterly subgoals complete.</td>
</tr>
</tbody>
</table>