

# TECHNOLOGY DEPT COMMUNITY/TECH COMMUNITY CHECK-IN 2017-10



**WIKIMEDIA**  
FOUNDATION

All content is © 2017 Wikimedia Foundation and is available under [CC-BY-SA 4.0](#) unless noted otherwise

# TECH GOAL QUADRANT

**A** Foundation level goals

**C** Features that we build to improve our technology offering

**B** Features we build for others

**D** Modernization, renewal and tech debt goals

The goals in each team pack are annotated using this scheme  
illustrate the broad trends in our priorities

# Programs: Planned Work



**WIKIMEDIA**  
FOUNDATION

## Departmental programs in this deck:

- Program 4: Technical community building
- Program 10. Public cloud services & support
- Program 5. Scoring Platform (ORES)
- Program 7. Smart tools for better data
- Program 9. Growing Wikipedia across languages via recommendations
- Program 12: Grow contributor diversity
- Program 11. Improving citations across Wikimedia projects

## Cross-departmental programs in this deck:

- Community Health (Segment 3. Research)
- Structured Data on Commons (Segment 4. Programs)

# Department Programs



WIKIMEDIA  
FOUNDATION

# Program 4: Technical Community Building



Wikimedia's software products and platforms have a diverse collection of technical and non-technical communities that have not always been well recognized for their contributions and supported in their work.

We will address this shortcoming by providing better documentation, facilitating community building, and establishing better pathways for communication between these communities and the Foundation.



Wikimedia Hackathon badge stickers By Wikimedia Österreich

## Outcome 1: Becoming a technical contributor to the Wikimedia movement is easier than it has been

### Objective 1: Collaborate with community on improving documentation

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Form a documentation Special Interest Group focused on WMCS products (A, B, C)</b>	LAST QUARTER	
	Technical Collaboration	<ul style="list-style-type: none"> <li>Create <a href="#">Special Interest Group framework</a> (non-goal work)</li> </ul>
	NEXT QUARTER	
	Technical Collaboration	<ul style="list-style-type: none"> <li>Form a WMCS Documentation Special Interest Group</li> <li>WMCS Doc SIG holds at least 1 meeting</li> </ul>

**STATUS: OBJECTIVE ON TRACK**



2017-10

**Cloud Services**4.1 [\[LINK\]](#)

## **Outcome 1: Becoming a technical contributor to the Wikimedia movement is easier than it has been**

### **Objective 2: Create tutorial content for common issues**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Plan contract documentation work (B, C)</b>	LAST QUARTER	The plan was to prepare materials needed to hire an outside contractor. We are pivoting on this and will instead be working with an existing Foundation staff member who has technical writing experience and an interest in returning to that field.

**STATUS: OBJECTIVE NOT DONE (Changed plan)**

2017-10

**Cloud Services**4.1 [\[LINK\]](#)

## **Outcome 1: Becoming a technical contributor to the Wikimedia movement is easier than it has been**

### **Objective 2: Create tutorial content for common issues**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Create one example of a "good" tutorial (A, B, C)</b>	NEXT QUARTER	
		<ul style="list-style-type: none"> <li>Improve <a href="#">My first Flask OAuth tool tutorial</a> to "good" status</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

2017-10

**Research**4.2 [\[LINK\]](#)

## Outcome 2: Adoption of Wikimedia technology can be measured

**Objective 1: Design a set of formal KPIs to measure tech audience growth & diversity**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Conduct interviews with teams building platforms or services for external audiences to characterize these core audiences and how adoption of these platforms or services can be measured (A,B,C)</b>	NEXT QUARTER	
	<ul style="list-style-type: none"> <li>Tech Management</li> </ul>	<ul style="list-style-type: none"> <li>Publish a report identifying key external audiences, adoption measures in current use, recommendations for possible KPIs (per team and/or across teams)</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

2017-10

**MediaWiki Platform**4.3 [\[LINK\]](#)

## Outcome 3: Vendors and 3rd parties are more confident in Wikimedia software

**Objective 1: Establish a canonical point of contact for 3rd party concerns**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Publish contact information and meet with third party users and developers electronically and in person</b>	NEXT QUARTER	
		<ul style="list-style-type: none"> <li>Established and published channels of communication</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

## Outcome 3: Vendors and 3rd parties are more confident in Wikimedia software

### Objective 2: Clarify the Foundation's short- and long-term 3rd party commitments

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Develop third party strategy and supporting artifacts with input from third party users and developers</b>	NEXT QUARTER	<ul style="list-style-type: none"> <li>• Strategy for MediaWiki for 3rd party usage</li> <li>• Catalog of Wikimedia software projects frequently used by third parties</li> <li>• Published multi-tiered, third-party support level system for Wikimedia software projects frequently used by third parties</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

2017-10

Research

4.4 [\[LINK\]](#)**Outcome 4: Collaboration with research in industry and academics is supported****Objective 1: Organize and host the Wiki Research Workshop**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Prepare for the annual Wiki Workshop</b>  <b>(B)</b>	<p><b>NEXT QUARTER</b></p> <ul style="list-style-type: none"> <li>• Workshop co-organizers (Stanford, EPFL)</li> <li>• Admin</li> </ul>	<ul style="list-style-type: none"> <li>• Write and submit a workshop proposal for The Web Conference 2018 (<a href="#">task T174741</a>)</li> <li>• If proposal accepted, we will write a call for papers for the workshop (<a href="#">task T174743</a>)</li> <li>• Invite reviewers to review workshop submission (<a href="#">task T174744</a>). The accepted submissions will present their research as part of the workshop poster session in April 2018 in Lyon.</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

## Outcome 4: Collaboration with research in industry and academics is supported

### Objective 2: Maintain the current capacity for formal research collaborations

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<p><b>Create formal collaborations that can expand our capacity towards meeting the annual plan objectives</b></p> <p><b>(B,C)</b></p>	<p>LAST QUARTER</p>	<ul style="list-style-type: none"> <li>● 6 new formal collaborations created in Q1:               <ul style="list-style-type: none"> <li>○ ETH, TSE (TP12)</li> <li>○ UMich (SDC)</li> <li>○ Stanford, Cornell (Community Health)</li> <li>○ Macalester (TP7)</li> </ul> </li> <li>● Centralized tracking of formal collaborations, NDAs and server access with Legal/Ops.</li> <li>● Hosted 5 presentations at our <a href="#">monthly research showcase</a> in Q1.</li> </ul>
	<ul style="list-style-type: none"> <li>● Legal</li> <li>● TechOps</li> </ul>	

**STATUS: OBJECTIVE ON TRACK**

2017-10

CTO Team

4.5 [\[LINK\]](#)

## Outcome 5: Organize Wikimedia Developer Summit

**Objective 1: Developer Summit web page published four months before the event includes dates and location, main theme, call for participation, call for scholarship requests, and calendar with deadlines. A good representation of non-WMF stakeholders related to the main theme are invited**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Planning the Dev Summit 2018 event (B, C)</b>	<b>LAST QUARTER</b>	
	<ul style="list-style-type: none"> <li>Technical Collaboration team</li> </ul>	Published event location, dates and process for deciding, theme and attendees
	<b>NEXT QUARTER</b>	
	<ul style="list-style-type: none"> <li>Developer Relations</li> <li>Program Committee (PC) members</li> </ul>	<ul style="list-style-type: none"> <li>PC evaluates submitted position statements; and selects invitees</li> <li>Invitations are sent and responses processed</li> <li>PC identifies themes and defines sessions</li> <li>All necessary event planning takes place</li> </ul>

**STATUS: OBJECTIVE ON TRACK**



CHECK IN

TEAM/DEPT

PROGRAM

WIKIMEDIA  
FOUNDATION  
ANNUAL PLAN

2017-10

CTO Team

4.5 [\[LINK\]](#)

## Outcome 5: Organize Wikimedia Developer Summit

Objective 2: Discussions start before the event with involvement of relevant stakeholders

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Preparations for the Dev Summit 2018 event (B, C)</b>	NEXT QUARTER	
	<ul style="list-style-type: none"><li>• Developer Relations</li><li>• Program Committee members</li></ul>	For each theme or session (TBD), an online discussion is initiated

**STATUS: OBJECTIVE ON TRACK**



# Program 10: Public Cloud Services And Support

We will offer services and support that empower developers to create robust technical solutions for the problems of the Wikimedia movement.



2017-10

**Cloud Services**10.1 [\[LINK\]](#)

## Outcome 1: PaaS is easy to use

### Objective 2: Migrate workflows to Striker

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Manage shared tool accounts via Striker (C)</b>	LAST QUARTER	
		<ul style="list-style-type: none"> <li>• Migrate Tool account creation from Wikitech to Striker</li> <li>• Add support for documenting purpose and OSI approved software license of new and existing Tools</li> </ul>
<b>Measure community opinion on ease of use (A, B)</b>	NEXT QUARTER	
	Research	<ul style="list-style-type: none"> <li>• Run annual <a href="#">Toolforge developer survey</a></li> <li>• Publish survey results</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

2017-10

**Cloud Services**10.1 [\[LINK\]](#)

## Outcome 2: Rebranding

### Objective 1: Complete core rebranding

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Perform Cloud Services rebranding (C)</b>	<b>LAST QUARTER</b>	
		<ul style="list-style-type: none"> <li>● Rename Phabricator projects</li> <li>● Rename mailing lists</li> <li>● Rebrand wiki content</li> <li>● Publish <a href="#">blog post</a></li> </ul>
	<b>NEXT QUARTER</b>	
		<ul style="list-style-type: none"> <li>● Rebrand wiki content</li> <li>● Rebrand remaining software products</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

## Outcome 3: Outreach

### Objective 1: Promote services and products

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
Attend Wikimania to support and promote Cloud Services products with the Wikimedia communities (B)	LAST QUARTER	
		<ul style="list-style-type: none"> <li>• <a href="#">"Introduction to Wikimedia Cloud Services"</a> presentation</li> <li>• <a href="#">"What is Cloud Services?"</a> poster</li> <li>• <a href="#">"FLOSS best practices for bots and tools"</a> poster</li> </ul>
Promote Toolforge Tools and maintainers	NEXT QUARTER	
	Technical Collaboration	<ul style="list-style-type: none"> <li>• Interview Tool maintainers</li> <li>• Publish blog post(s)</li> </ul>

**STATUS: OBJECTIVE COMPLETE**

2017-10

**Cloud Services**10.1 [\[LINK\]](#)**Outcome 4: First line tech support****Objective 1: Provide first line technical support resources**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Hire first line technical support contractor (B)</b>	<b>LAST QUARTER</b>	
		<ul style="list-style-type: none"> <li>• Define duties for support contract position</li> <li>• <del>Advertise and fill position</del></li> </ul>
	<b>NEXT QUARTER</b>	
		<ul style="list-style-type: none"> <li>• Advertise and fill position</li> </ul>

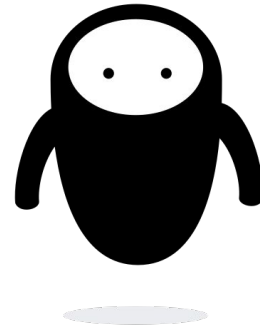
**STATUS: OBJECTIVE BEHIND SCHEDULE**

# Program 5: Scoring Platform (ORES)





We will help increase the efficiency of production activities on the wikis with machine prediction services and we will build accountability mechanisms to mitigate the effects of prediction errors and bias



**SCORING  
PLATFORM**  
TEAM

Team [\[ edit \]](#)

100%

20 → 50%

80%

15%



Aaron Halfaker  
Principal Research Scientist  
Team Lead



Amir Sarabadani  
Software Engineer  
(Wikimedia Deutschland /  
volunteer)



Adam Wight  
Software  
Engineer  
(WMF)



Sumit  
Volunteer

## Former

- とある白い猫 (IEG)
- Arthur Tilley (IEG)
- He7d3r (IEG/Volunteer)
- Yuvipanda (Volunteer)

## Outcome 1: Tool developers and Product teams can innovate tools that use machine prediction to make wiki-work more efficient.

Objective 1: Expand vandalism & good-faith detection models to more wikis (focus on Emerging Communities)

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Deploy thresholds selection system (T162217)</b>  <b>Deploy advanced support for Albanian and Romanian Wikipedias and basic support for Greek &amp; Tamil Wikipedia</b>  <b>(C)</b>	LAST QUARTER	
	<ul style="list-style-type: none"> <li>Operations</li> </ul>	Thresholds selection system <b>(due for deployment in early Q2)</b>  Support for target wikis (Albanian, Romanian, Greek, and Tamil) deployed successfully.
	NEXT QUARTER	
		Deploy full edit quality support to Swedish and Spanish Wikipedia as well as Spanish Wikibooks  Train/test a draft topic modeling AI (experimenting with mitigating new page curation issues)

## Outcome 2: Volunteers are empowered to track trends in prediction bias and other failures of AI in the wiki.

**Objective 1: Develop best practices for using community input to improve/correct predictions**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Design schema for meta ORES (T153152)</b>  <b>Report on implementations of meta ORES (T166053)</b>  <b>(C)</b>	LAST QUARTER	
	<ul style="list-style-type: none"> <li>Community Engagement</li> </ul>	<p><a href="#">Designed</a> schema for meta ORES</p> <p><a href="#">Reported</a> on implementations of meta ORES</p>
	NEXT QUARTER	
	<ul style="list-style-type: none"> <li>Community Engagement</li> </ul>	<p>Engage in a community lead discussion about the design and use of JADE</p> <p>Deploy an MVP of the JADE API in Cloud VPS (state database, event generation, structured UI support)</p>

# Notable accomplishments

- 4 papers accepted at ACM GROUP, ACM OpenSym, ACM CSCW, ACL EMNLP
- Lots of blog activity
  - <https://blog.wikimedia.org/2017/07/19/scoring-platform-team/>
  - <https://blog.wikimedia.org/2017/08/30/wikipedia-bot-pocalypse/>
  - <https://blog.wikimedia.org/2017/09/29/ores-distributed-processing/>
  - [https://phabricator.wikimedia.org/phame/post/view/68/more\\_better\\_model\\_information\\_and\\_the\\_reshold\\_optimizations/](https://phabricator.wikimedia.org/phame/post/view/68/more_better_model_information_and_the_reshold_optimizations/)
  - [https://phabricator.wikimedia.org/phame/post/view/69/wikilabels\\_incident\\_reversed\\_diffs/](https://phabricator.wikimedia.org/phame/post/view/69/wikilabels_incident_reversed_diffs/)
- Public talks
  - Wikimania'17 (Keilana Effect, How can I get ORES for my Wiki?)
  - OpenSym'17 (Keilana Effect)
  - Peer production Workshop (Backlogs and ORES)
  - Quora (Wikipedia's socio-technical dynamics)



# Program 7: Smart Tools for Better Data

We will maintain and increase public access to past, present and real time data for Wikimedia projects. We will provide the infrastructure to measure the impact and reach of projects and features for editors, communities and WMF.





...measure the impact and reach of projects and features for editors, **communities** and WMF.



2017-10

**Analytics**7.1 [\[LINK\]](#)**Outcome 1: Foundation staff and community have better tools to access data.****Objective 1: Wikistats 2.0 redesign.**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<p><b>(carry on) Initial deployment of wikistats 2.0 UI.</b></p> <p><b>Start development of a new backend and API on top of the Data Lake Edit data in Hadoop.</b></p> <p><b>(A, B)</b></p>	LAST QUARTER	
	TechOps, DBA	We deployed the new wikistats UI with access to Pageview Data (not announced yet) but run into privacy/architecture issues with edit metrics backend.
	NEXT QUARTER	
	TechOps, Research	Deployment of api on top of Public Druid Cluster. <a href="#">Task T176223</a>  <b>Wikistats 2.0 alpha release with edit and pageview data</b>

**STATUS: OBJECTIVE DELAYED**

<https://stats.wikimedia.org/v2/#/bm.wikipedia.org>

## Monthly Overview

Wikipedia - Bambara

### Reading

#### Total Page Views

**350.6K**

September **↑ 10.81 % month over month**



**2.8M** **↓ -8.56 % year over year**

Year Total (2016)

#### Unique Devices

**6.7K**

September **↓ -7.40 % month over month**



**8.1K** **↓ -17.76 % year over year**

Year Average (2016)

#### Top Viewed Articles

Most viewed articles for October

**6.3k** Nye fólk

**1.9k** Spécial:Connexion

**1.6k** Spécial:Journal

**1.6k** Discussion:Nye fólk

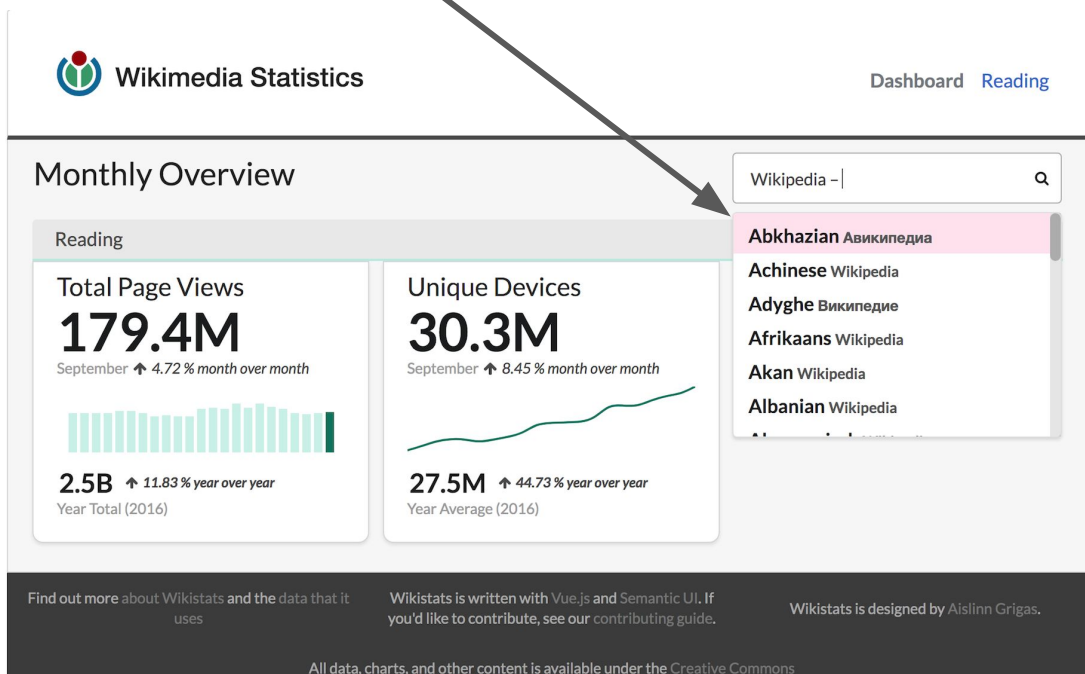
Find out more about [Wikistats](#) and the data that it uses

Wikistats is written with [Vue.js](#) and [Semantic UI](#). If you'd like to contribute, see our [contributing guide](#).

Wikistats is designed by [Aislinn Grigas](#).

All data, charts, and other content is available under the [Creative Commons CC0](#) dedication.

# Having all wikis on labs replicas will make possible to show data for all wikis in wikistats 2.0. This is a 3 way team effort: DBAs, Cloud Services and Analytics



2017-10

**Cloud Services**7.3 [\[LINK\]](#)

## Outcome 3: Data services

### Objective 2: Complete migration of Wiki Replica access to new high-availability cluster

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Migrate all users to new Wiki Replica cluster (C, D)</b>	<b>LAST QUARTER</b>	
	DBA	<a href="#">Announce availability of cluster</a> (non-goal work)
	<b>NEXT QUARTER</b>	
	DBA	Fix issues identified by user testing of new servers Help users adapt to feature deprecations Decommission old servers

**STATUS: OBJECTIVE ON TRACK**

# MediaWiki databases

How many  
**new editors**  
across all  
wikis?

```
new_editors.old.sql (~/projects/tmp) - GVIM
1 select month,
2   count(*)
3   from (select substr(log_timestamp, 1, 6) month,
4     log_user
5     from logging l
6     inner join
7     (select rev_user,
8      rev_timestamp
9      from revision
10
11     union all
12
13     select ar_user rev_user,
14      ar_timestamp rev_timestamp
15     from archive
16    ) edits
17     on rev_user = log_user
18     and log_type = 'newusers'
19     and log_action = 'create'
20
21   where unix_timestamp(str_to_date(rev_timestamp, 'yyyyMMddHHmmss')) -
22     unix_timestamp(str_to_date(log_timestamp, 'yyyyMMddHHmmss'))
23     <= 86400
24     and log_timestamp between '20160101000000' and '20160201000000'
25
26   group by substr(log_timestamp, 1, 6),
27     log_user
28
29   ) new_editors
30
31 group by month
32 order by month
33 limit 1000
34 ;
~/projects/tmp/new_editors.old.sql" 33L, 1067C written      1,2      All
```

2 joins, 3 subqueries  
3 tables on 800+ wikis  
**days** to run

# MediaWiki databases

Primary ids are **missing**  
for most events!

Instead they have page  
title or user name.

Pages and users are **renamed** all the time!

Data is stored in **different formats**

depending on MediaWiki version  
that created the log record

sometimes in **PHP encoded  
blobs!**

page and user properties might  
**be in different fields** depending  
on age of event!

Durations of user blocks is  
stored in **7 different date  
formats!**

Some revisions are  
historically **orphaned**  
(no existent page IDs!)

Some page **IDs were reused** for  
different pages! (I guess IDs were a  
scarce commodity in 2006?)

# MediaWiki history

How many  
new editors  
across all  
wikis?

```
new_editors.new.sql (~/.projects/tmp) - GVIM
1 select month,
2     count(*)
3     from (select substr(event_user_creation_timestamp, 0, 6) month,
4             event_user_id
5           from history
6           where event_entity = 'revision'
7                 and event_type = 'create'
8                 and event_user_id is not null
9                 and event_user_is_created_by_self
10                and unix_timestamp(event_timestamp, 'yyyyMMddHHmmss') -
11                      unix_timestamp(event_user_creation_timestamp, 'yyyyMMddHHmmss')
12                      <= 86400
13
14          group by substr(event_user_creation_timestamp, 0, 6), event_user_id
15
16         ) new_editors
17
18 group by month
19 order by month
20
```

~/projects/tmp/new\_editors.new.sql" 20L, 639C written 20,1 All

0 joins, 0 subqueries  
1 table  
**minutes** to run

A background collage of various line-art icons representing different fields: science (satellite, telescope, planet Saturn, microscope), geography (maps of Australia, Europe, Africa, South America), history/culture (pyramids, sphinx, Egyptian head, kangaroo), and general concepts (Eiffel Tower, globe, camera, microscope, microscope, microscope).

...measure the impact and reach of projects and features for editors, **communities** and WMF.

**ACTRIAL**



2017-10

**Analytics**7.1 [\[LINK\]](#)**Outcome 1: Foundation staff and community have better tools to access data.****Objective 1: Wikistats 2.0 redesign.**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Add Edit count to Data Lake per event for all events for all wikis since the beginning of time. (A, B)</b>	LAST QUARTER	
	TechOps	This was part of our work geared to have data for Wikistats that we moved ahead to get <b>ACTRIAL</b> metrics.
	NEXT QUARTER	

**STATUS: OBJECTIVE ON TRACK**

# Daily Pages Created By Autopatrolled Users on English Wikipedia

add project or language



Wikipedia

English EN

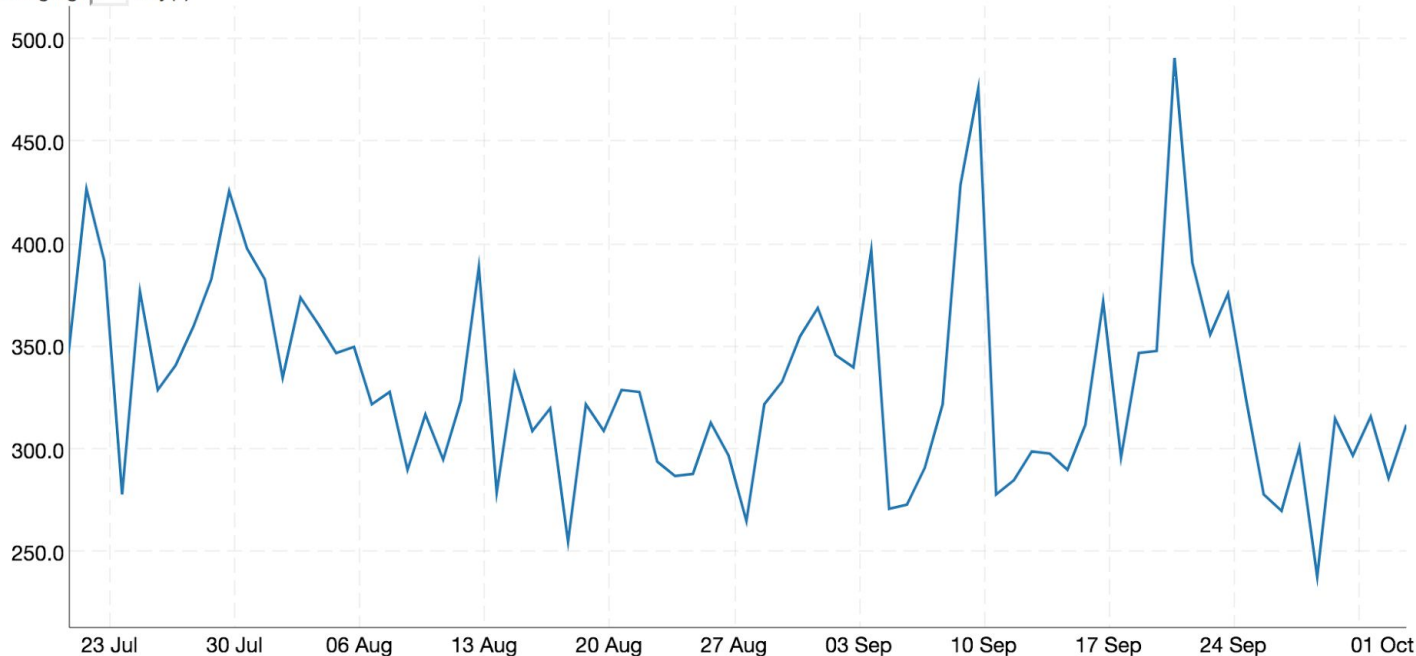
+ Add Metric

Daily Pages Created

Daily Pages Created in the Main namespace (ns=0) by autopatrolled users

Daily Pages Created in the Main namespace (ns=0) by autopatrolled users

Averaging: 1 day(s)



2017-10

**Cloud Services**7.3 [\[LINK\]](#)

## Outcome 3: Data services

### Objective 1: Provide reliable and available access to dumps

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Begin migrating customer-facing Dumps endpoints to Cloud Services (C)</b>	LAST QUARTER	
		Rack and provision new Dumps servers Migrate one customer use case to them
	NEXT QUARTER	
		Complete provisioning of Dumps servers Migrate one customer use case to new Dumps servers

**STATUS: OBJECTIVE BEHIND SCHEDULE**

The background features a light gray collage of various icons. These include architectural structures like the Brandenburg Gate and the Eiffel Tower, geographical maps of continents, scientific symbols like a planet with rings and a telescope, and other elements like a stork, a person with a headlamp, a person with a hard hat, a large letter 'W', and various animals like a llama and a kangaroo.

...measure the impact and reach of projects and features for editors, communities and **WMF**.

2017-10

**Analytics**7.1 [\[LINK\]](#)**Outcome 1: Foundation staff and community have better tools to access data.****Objective 2: Better visual access to EventLogging data**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Spike: load data in superset, test and productionize if pertains</b>	LAST QUARTER	
		We got an install of superset running but it is not productionized.
	NEXT QUARTER	

**STATUS: OBJECTIVE DELAYED**

Run Query Save

Sort Descending

Chart Options

X Axis

Y Axis

Y Axis Label: Unique Devices Left Margin: auto

Y bounds Y Log Scale

Y Axis Format: +, | +12,345.4321 Y Axis Bounds: Min 150000000

Advanced Analytics

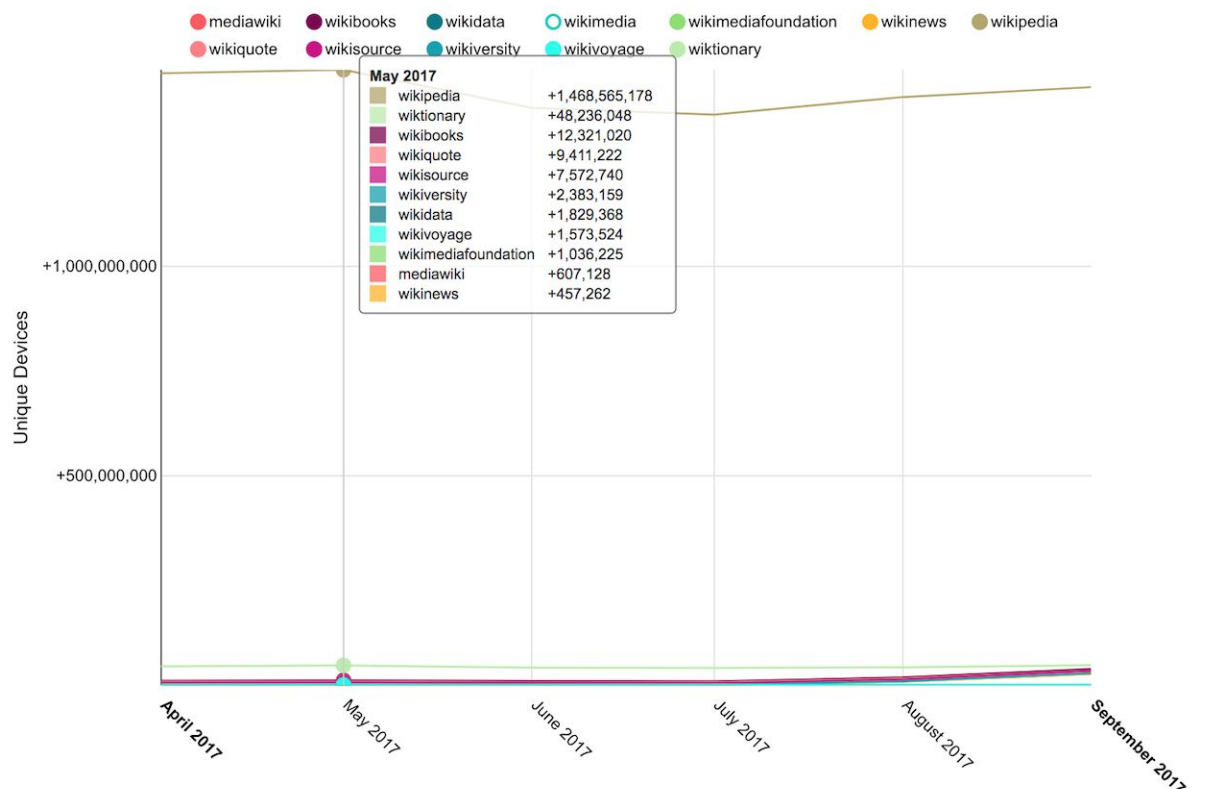
Rolling: None Periods: Min Periods:

Time Shift:

Period Ratio: Period Ratio Type: growth

# Unique Devices Per Project Monthly

00:00:00.55 View Query



2017-10

**Analytics**7.1 [\[LINK\]](#)**Outcome 1: Foundation staff and community have better tools to access data.****Objective 2: Experiments with real-time data and community support for new datasets available**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Productionize Real-time ingestion of data into druid</b>	LAST QUARTER	
<b>Release new datasets.</b>	NEXT QUARTER	
	TechOps	Productionize Spark streaming for banner data  Release clickstream dataset

**STATUS: OBJECTIVE IN PROGRESS**

DIMENSIONS

- Time
- Anonymous
- Banner
- Bucket
- Campaign
- Country
- Country Matches Geocode
- Device
- Project
- Region
- Sample Rate
- Status Code
- Uselang

MEASURES

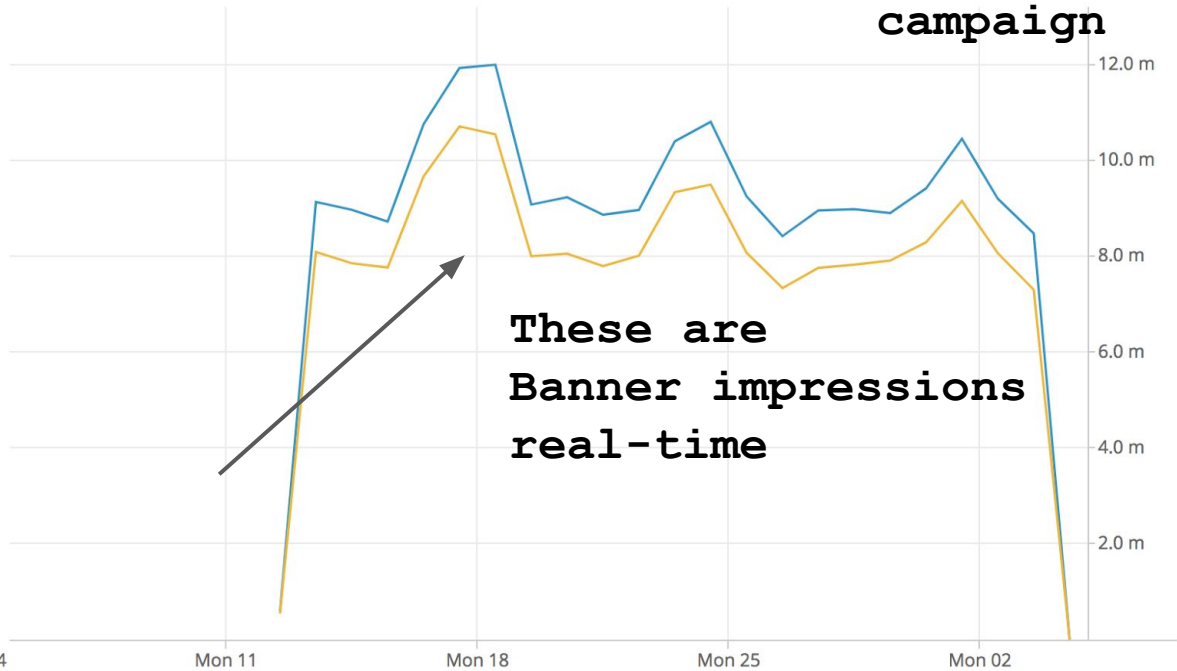
- Request Count
- Normalized Request Count

FILTER: Sep 4 - Oct 4, 11:15pm Campaign: C1718\_jaJP\_...

SPLIT: Device Time (Day)



Normalized Request Count: 378.8 m



Fundraising campaign

These are Banner impressions real-time

LEGEND

- iphone 201.3 m
- android 177.5 m

Show selection

PINBOARD

Click or drag dimensions to pin them



2017-10

**Analytics**7.1 [\[LINK\]](#)**Outcome 1: Foundation staff and community have better tools to access data.****Objective 2: Better visual access to eventlogging data**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>EventLogging refine, automagic table creation in hadoop (C, D)</b>	<b>LAST QUARTER</b>	
		Prototype scala code for refinement of eventlogging data in hadoop.
	<b>NEXT QUARTER</b>	
	TechOps	Start sunseting of MYSQL Backend  Productionize automagic hive table creation from eventlogging data.

**STATUS: OBJECTIVE IN PROGRESS**

## Outcome 2: Access to Wikimedia content and data with scalable APIs

### Objective 1: Revision storage scaling

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Gradual roll-out of Cassandra 3 &amp; new schema to resolve storage scaling issues and OOM errors. (B, D)</b>	LAST QUARTER	
	TechOps	Deployed and tested Cassandra 3 cluster. Finished implementation of new Cassandra storage schema. Migrated mobile HTML storage.
	NEXT QUARTER	
	TechOps	Migrate most use cases.

**STATUS: OBJECTIVE ON TRACK**



WIKIPEDIA  
La enciclopedia libre

Portada  
Portal de la comunidad  
Actualidad  
Cambios recientes  
Páginas nuevas  
Página aleatoria  
Ayuda  
Donaciones  
Notificar un error

Imprimir/exportar

Crear un libro  
Descargar como PDF  
Versión para imprimir

Herramientas

Lo que enlaza aquí  
Cambios en enlazadas  
Subir archivo  
Páginas especiales  
Enlace permanente  
Información de la página  
Elemento de Wikidata  
Citar esta página

No has iniciado sesión

Artículo **Discusión**

Leer **Editar** Ver historia

# Selección de fútbol de Canarias

Popup  
on eswiki

La **Selección de fútbol de Canarias**, es el equipo formado por jugadores canarios,<sup>1</sup> organizado por la Federación Canaria de Fútbol (FCF), entidad que gestiona el fútbol, el fútbol sala y el fútbol playa en el ámbito del **Archipiélago Canario (España)**.

Carece de reconocimiento por parte de la **FIFA**, pues sus estatutos (artículo 10)<sup>2</sup> estipulan que para que una selección pueda participar en competiciones internacionales oficiales debe o bien representar a un estado independiente o tener autorización expresa de la federación del país correspondiente (en este caso la Federación Española de Fútbol). Es por este motivo que sólo puede disputar partidos amistosos en categoría absoluta.

El combinado de Canarias sub-19, disputa anualmente la **Copa del Atlántico**, midiéndose a la selección de España y otras dos invitadas, en una competición que parte desde semifinale

## Índice [ocultar]

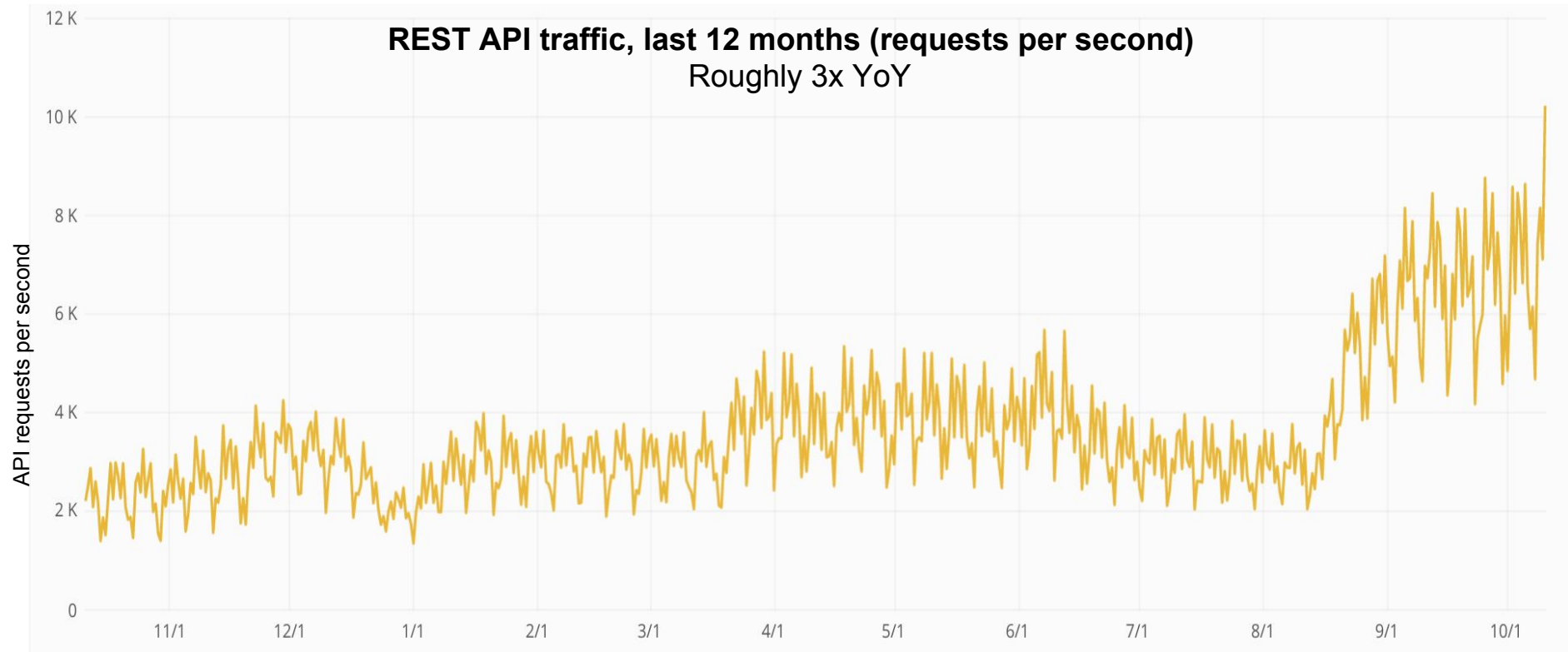
- 1 **Historia**
- 2 **Partidos**
  - 2.1 **Denominación no oficial**
  - 2.2 **Denominación oficial**
- 3 **Jugadores Internacionales**
- 4 **Categorías inferiores**
  - 4.1 **Selección sub-18**
  - 4.2 **Selección sub-16**
- 5 **Otras categorías**
  - 5.1 **Selección de fútbol amateur de Canarias**
  - 5.2 **Selección femenina de fútbol de Canarias**
- 6 **Véase también**

La Copa del Atlántico es un torneo internacional de selecciones de fútbol sub-18. Se disputa en la isla de Gran Canaria, España, desde el año 1964, y ya han tenido lugar 43 ediciones de manera oficial con el nombre actual, y 2 ediciones anteriores no oficiales. Generalmente se celebra en lo



# REST API traffic, last 12 months (requests per second)

Roughly 3x YoY



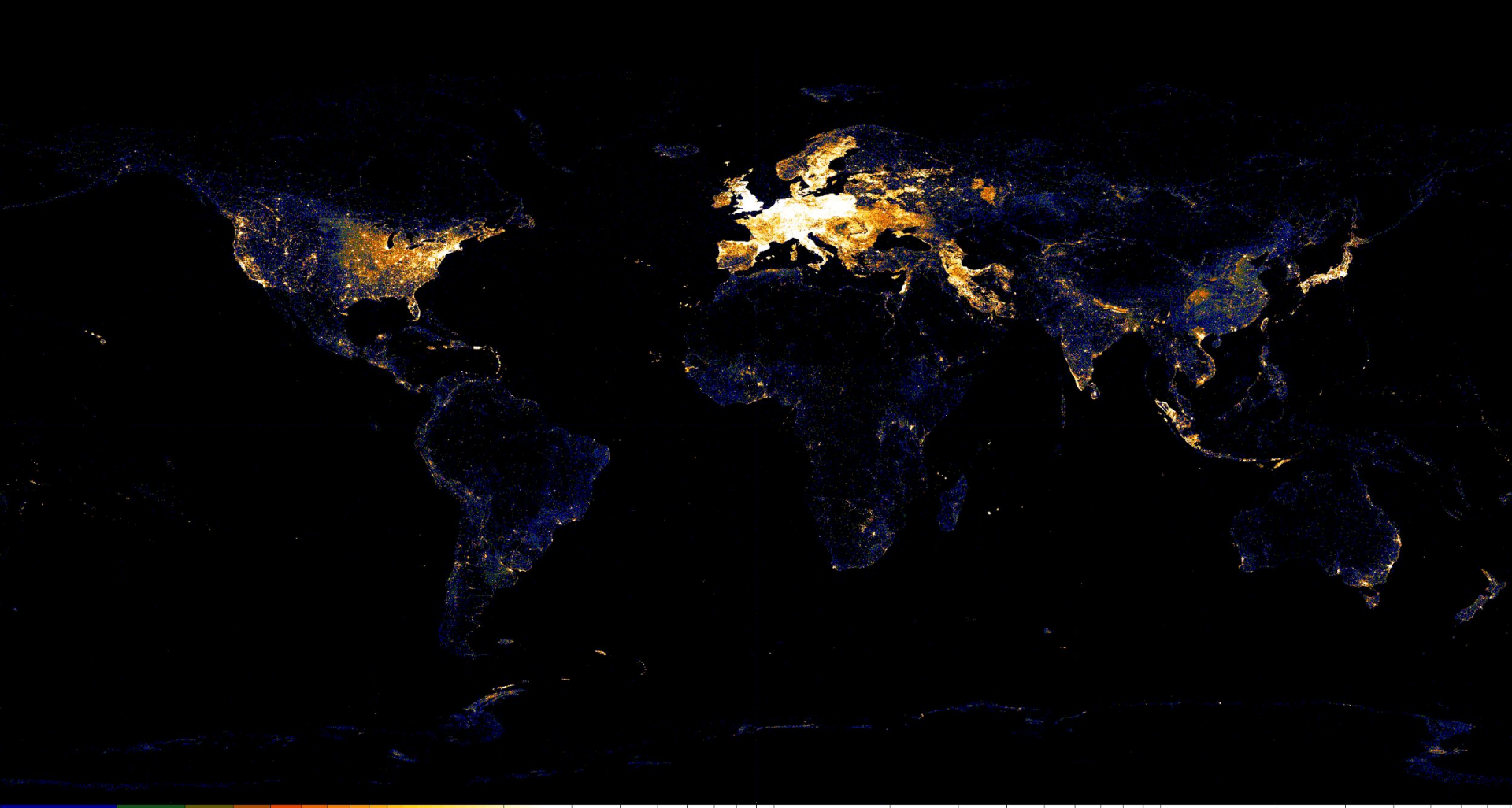
Q4 2016/17

Q1 2017/18

Q2 2017/18

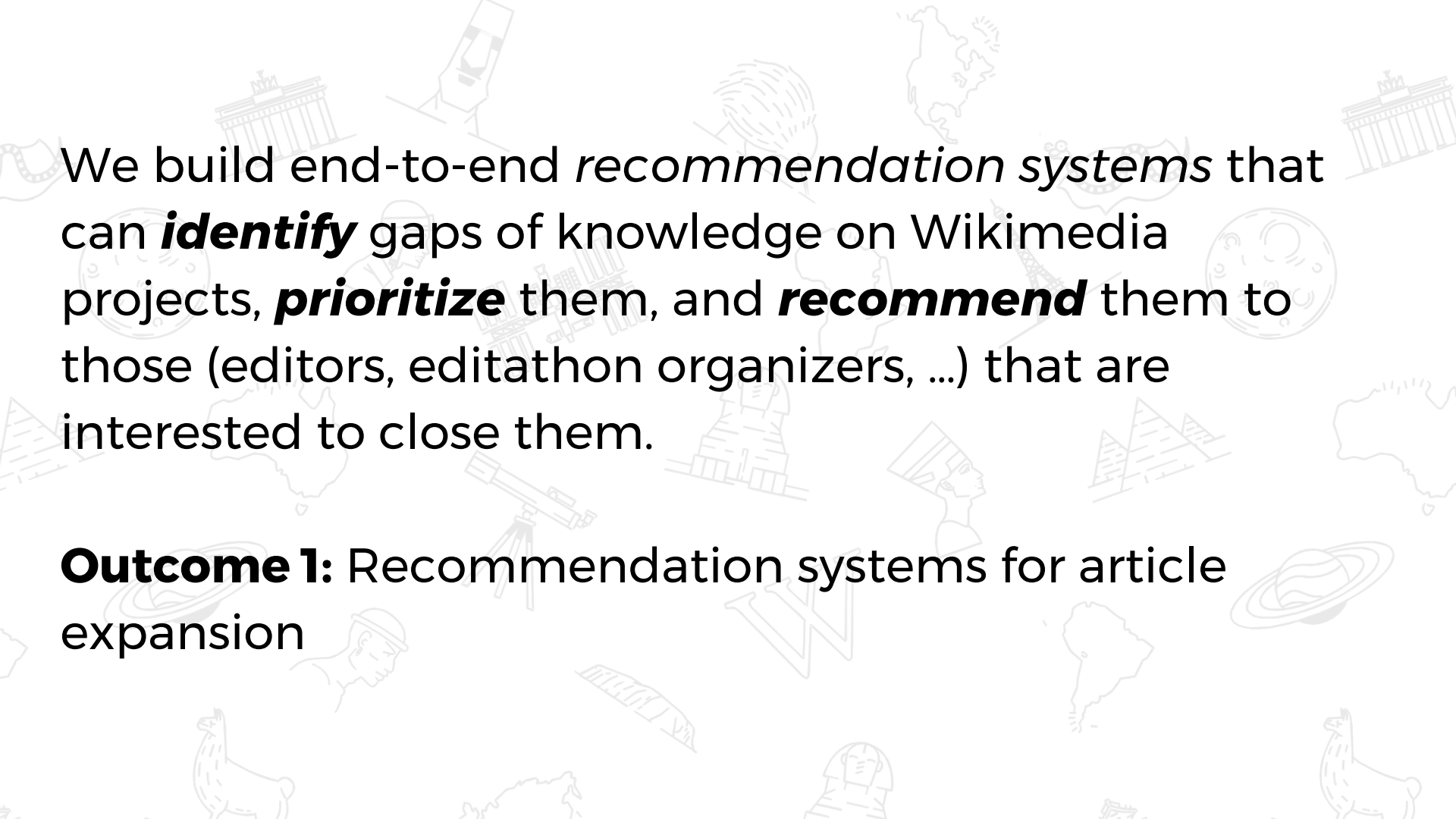


# **Program 9: Growing Wikipedia Across Languages via Recommendations**



The background features a light gray collage of various icons. These include architectural structures like the Brandenburg Gate, scientific instruments like a microscope and a telescope, celestial bodies like Saturn and the moon, geographical maps of continents, and other symbols like a globe, a person's head, and a large letter 'W'.

**How do we make  
access to all  
knowledge universal?**



We build end-to-end *recommendation systems* that can **identify** gaps of knowledge on Wikimedia projects, **prioritize** them, and **recommend** them to those (editors, editathon organizers, ...) that are interested to close them.

**Outcome 1:** Recommendation systems for article expansion





WIKIPEDIA  
The Free Encyclopedia

- Main page
- Contents
- Featured content
- Current events
- Random article
- Donate to Wikipedia
- Wikipedia store

#### Interaction

- Help
- About Wikipedia
- Community portal
- Recent changes
- Contact page

#### Tools

- What links here
- Related changes
- Upload file
- RTRC
- Special pages
- Page information
- Wikidata item

#### In other projects

- Wikimedia Commons

#### Languages

- brezhoneg
- Català
- Español
- हिन्दी
- Bahasa Indonesia

Article [Talk](#)

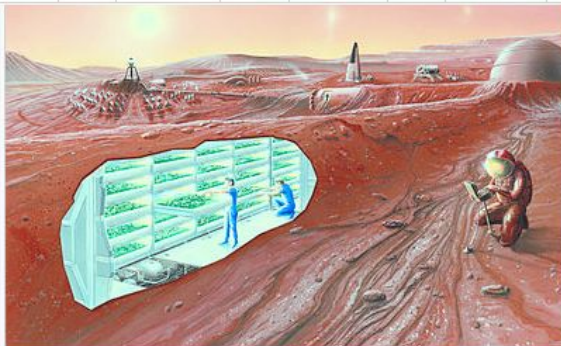
[Read](#) [Edit this page](#) [View history](#) [More](#)

Paragraph Cite Insert [Save changes](#)

**colonization**. Its surface conditions and the presence of **water on Mars** make it arguably the most **hospitable of the planets** in the **Solar System**, other than **Earth**. Mars requires less energy per unit mass (**delta-v**) to reach from Earth than any planet except **Venus**.

Permanent human habitation on a planetary body other than the Earth is one of science fiction's

most prevalent themes. As technology has advanced, and **concerns about the future of humanity on Earth** have increased, the argument that **space colonization** is an achievable and worthwhile goal has gained momentum.<sup>[1][2]</sup> Other reasons for colonizing space include economic interests, long-term scientific research best carried out by humans as opposed to robotic probes, and sheer curiosity.



An artist's conception of a human Mars base, with a cutaway revealing an interior horticultural area

## Relative similarity to Earth

### Sections you can add

[Relative similarity to Earth](#)

[Differences from Earth](#)

[Conditions for human habitation](#)

[Radiation](#)

[Transportation](#)

[Equipment needed for colonization](#)

[Robotic precursors](#)

[Mission concepts](#)

[Economics](#)

[Possible locations for settlements](#)

[Planetary protection](#)

2017-10

**Research**9.1 [\[LINK\]](#)**Outcome 1: Article expansion recommendations for editors and editathon organizers****Objective 1: Build, improve, and expand algorithms for article expansion recommendations**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Clean up the category system for machine consumption</b>	<b>LAST QUARTER</b>	
	Formal collaborators (EPFL, Stanford)	<ul style="list-style-type: none"> <li>Developed a computationally cheap technique for creating a hierarchical WP category network</li> <li>We learned (by exhaustive trial and error) that building an IS-A hierarchical network of Wikipedia categories for machines is not (at least reasonably easily) possible</li> </ul>
	<b>NEXT QUARTER</b>	
	Formal collaborators (EPFL, Stanford)	<ul style="list-style-type: none"> <li>Save others: A blog post or some other form of publication about the approaches we tried for cleaning up the category system and failed (stretch)</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

2017-10

**Research**9.1 [\[LINK\]](#)**Outcome 1: Article expansion recommendations for editors and editathon organizers****Objective 1: Build, improve, and expand algorithms for article expansion recommendations**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Building the section recommendation system</b>	<b>LAST QUARTER</b>	
	Formal collaborators (EPFL, Stanford)	<ul style="list-style-type: none"> <li>• Intra-language: Built the first version of section recommendation algorithms using a probabilistic model based on WP's category network, LDA models, and ALS collaborative filtering</li> <li>• Initiated inter-language signal based models</li> </ul>
	<b>NEXT QUARTER</b>	
	Formal collaborators (EPFL, Stanford)	<ul style="list-style-type: none"> <li>• Build a model to do section alignment across languages</li> <li>• Build a section recommendation model using only the inter-language signals</li> <li>• Fusion of inter-language and intra-language models (Stretch)</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

2017-10

**Research**9.1 [\[LINK\]](#)**Outcome 1: Article expansion recommendations for editors and editathon organizers****Objective 4: Build Labs API for article expansion recommendations**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Build a tool and improve the API for article expansion recomms</b>	<b>LAST QUARTER</b>	
	<b>NEXT QUARTER</b>	
	Formal collaborators (EPFL, Stanford) Baha Pau	<ul style="list-style-type: none"> <li>Build a tool for article expansion recomms, run an online experiment</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

2017-10

**Research**9.1 [\[LINK\]](#)**Outcome 1: Article expansion recommendations for editors and editathon organizers****Objective 1: Build, improve, and expand algorithms for article expansion recommendations****Objective 4: Build Labs API for article expansion recommendations**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Evaluating the section recommendation system and publishing</b>	<b>LAST QUARTER</b>	
	Formal collaborators (EPFL, Stanford)	<ul style="list-style-type: none"> <li>Tested the first results of section recommendations</li> </ul>
	<b>NEXT QUARTER</b>	
	Formal collaborators (EPFL, Stanford) Baha Pau	<ul style="list-style-type: none"> <li>Evaluate section recommendations with baselines</li> <li>Run an online experiment</li> <li>Submit a paper, write a blog post, assess the usefulness of data release</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

# Notable

- Ongoing conversations with UNICEF Office of Innovation about potential collaborations in the area of information poverty
- Ongoing work with Allen Institute for Artificial Intelligence and Primer on the creation of a rich dataset about women in computer science missing in Wikipedia or Wikidata.

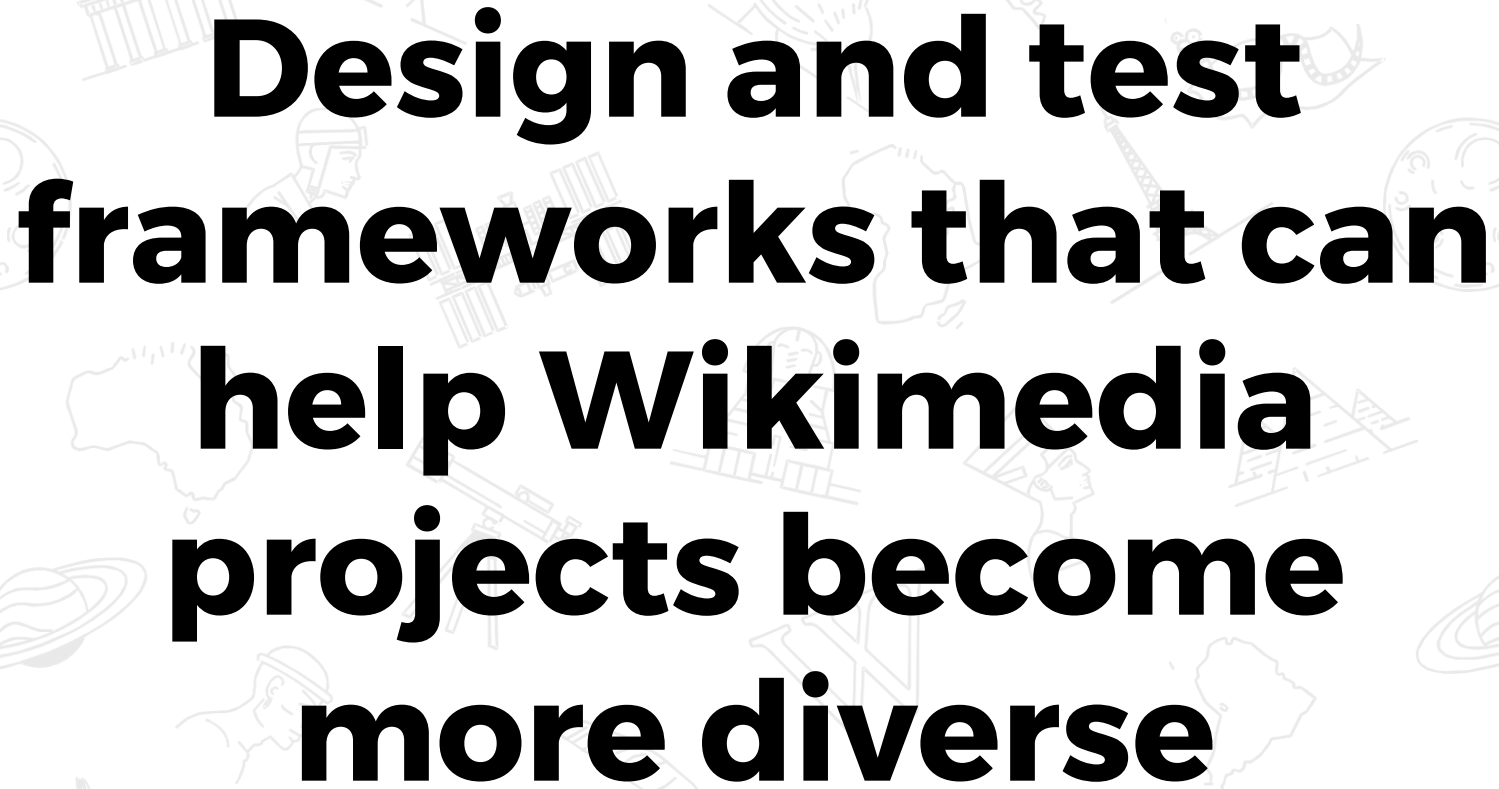


# Program 12: Grow Contributor Diversity



**Only 10-15% of  
Wikipedia editors are  
estimated to be  
female, and this does  
not stop here.**





**Design and test  
frameworks that can  
help Wikimedia  
projects become  
more diverse**

## Outcome 1: We improve Wikipedia's contributor diversity after designing and testing potential intervention(s)

**Objective 1: Identify the underlying (potential) causes of lack of representative contribution from certain demographics**

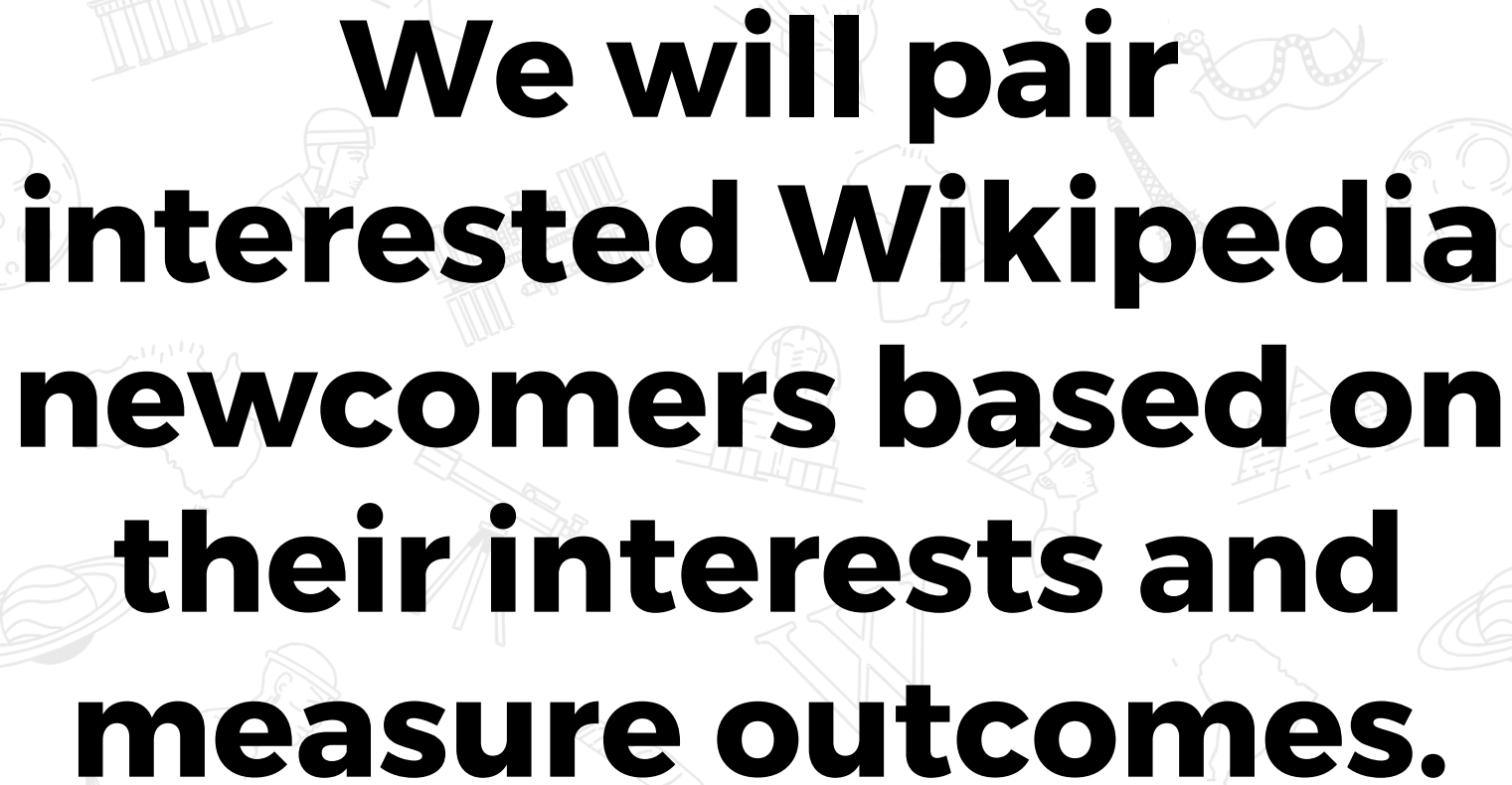
**Objective 2: Design frameworks to change the current socio-technical infrastructure to address at least one of the underlying causes of lack of representativeness**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<p>Create one or more formal collaborations for this research</p> <p>Perform a literature review and potentially run survey(s) to identify the self-reported causes of imbalanced representation in contribution</p> <p>Initiate the design of a framework to address one of the identified/hypothesized causes of imbalance in contributor demographics, if time permits. (Stretch)</p> <p>Run a couple of quick surveys to get a better sense of where in the pipeline we start losing diversity. (Stretch)</p>	<b>LAST QUARTER</b>	
	<p>Formal collaborators (EPFL, ETH, TSE)</p>	<ul style="list-style-type: none"> <li>● Created formal collaborations</li> <li>● Finalized the literature review</li> <li>● Designed a high level framework for testing for confidence</li> </ul>
	<b>NEXT QUARTER</b>	

**STATUS: OBJECTIVE ON TRACK**

# The literature suggests:

- Women and men of the same ability differ significantly in terms of self-confidence levels and taste for competition.
- There are important and systematic differences in those dimensions between the two gender with important consequences.
- These differences predict orientation choices among high school students, industry choices among young business professionals, and part of the earning differential.
- Those gender differences tend to disappear when women are given a chance to work in **teams**.



**We will pair  
interested Wikipedia  
newcomers based on  
their interests and  
measure outcomes.**



# Program 11: Improving Citations across Wikipedia Projects

We will increase the verifiability of Wikimedia contents by conducting research aiming to improve how citations and sources are stored, accessed and vetted.



## Outcome 1: Quantitative research is available to help Wikipedia and Wikidata contributors focus and prioritize their sourcing efforts.

**Objective 1: Estimate what proportion of content in Wikipedia or Wikidata is unsourced and in need of citations.**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Publish a rich dataset on citation contexts to support additional research on Wikipedia sourcing gaps</b>	LAST QUARTER	
	<ul style="list-style-type: none"> <li>Formal collaborators (Drexel)</li> <li>Scoring Platform</li> </ul>	<ul style="list-style-type: none"> <li>Finalize the parsing strategy and complete the preparation of raw citation context data</li> </ul>
	NEXT QUARTER	
	<ul style="list-style-type: none"> <li>Formal collaborators (Drexel)</li> <li>Scoring Platform</li> </ul>	<ul style="list-style-type: none"> <li>Formally publish and announce the availability of the dataset.</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

# A structured dataset of citation contexts

**Amy Jean Klobuchar** ([/ˈkloʊbəˈfɑːr/](#); born May 25, 1960) is the **senior United States Senator** from **Minnesota**. She is a member of the **Minnesota Democratic-Farmer-Labor Party**, an affiliate of the **Democratic Party**. She is the first woman to be elected as a senator for Minnesota and is one of twenty-one women serving in the current **United States Senate**.

She previously served as **County, Minnesota**

Senate Web site (2007). "U.S. Senator for Minnesota Amy Klobuchar: Biography". Archived from the original on February 21, 2007. Retrieved 2007-02-23.



As an attorney, she worked with former **vice President Walter Mondale**.<sup>[1]</sup> She has been called a "rising star" in the **Democratic Party**.<sup>[2][3]</sup>

```
{
  "name": "bio",
  "raw_content": "<ref name=bio>{{cite web| author= Senate Web s
Senator for Minnesota Amy Klobuchar: Biography| year = 2007| url
http://klobuchar.senate.gov/biography.cfm| accesdate= 2007-02-2
https://web.archive.org/web/20070221013430/http://klobuchar.sena
|archivedate = February 21, 2007}}</ref>",
  "templated": true,
  "cite_template": "cite web",
  "urls": [
    "http://klobuchar.senate.gov/biography.cfm|",
    "https://web.archive.org/web/20070221013430/http://klobuchar.sen
  ],
  "identifiers": [],
  "occurrences": [
    {
      "section": 0,
      "text_offset": 2033,
      "preceding_text": "s Senate]].\n\nShe previously served as
attorney]] for [[Hennepin County, Minnesota]], the most populous
As an attorney, she worked with former [[Vice President of the U
President]] [[Walter Mondale]].",
      "succeeding_text": " She has been called a \"rising star\"
Party.<ref>{{Cite news|url=http://m.startribune.com/huffington-p
the-smartest-u-s-senator/87658327/?section=politics/blogs|title=
Klobuchar the smartest U.S.",
      "header_level": 0,
      "header_text": null,
      "header_offset": null,
      "level_2_text": null,
      "level_2_offset": null
    },
  ],
}
```



## Outcome 3: Outreach activities continue to ensure community and technical efforts are aligned


### Objective 1: Fundraise for, and host the 3rd annual meeting in the WikiCite series

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Write a report for the funders and fundraise for WikiCite 2018</b>	LAST QUARTER	
	<ul style="list-style-type: none"> <li>Formal collaborators (CMU)</li> </ul>	<ul style="list-style-type: none"> <li>Completed the survey of WikiCite 2017 participants</li> <li>Wikimania session: <a href="#">Citations needed for the sum of all knowledge</a></li> </ul>
	NEXT QUARTER	
	<ul style="list-style-type: none"> <li>WikiCite organizing committee</li> <li>Advancement</li> </ul>	<ul style="list-style-type: none"> <li>Publish the WikiCite 2017 report</li> <li>Fundraise for the WikiCite 2018 event</li> <li>Additional outreach: talks accepted at FORCE '17, WikidataCon. Keynote at SWIB '17</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

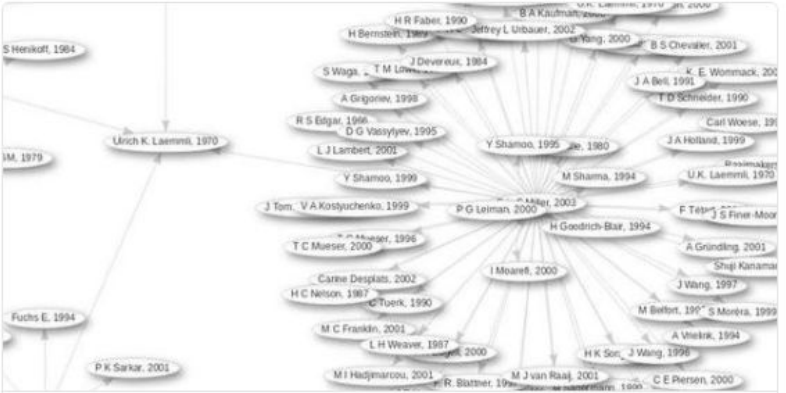
# Initiative for Open Citations

- **Progress:** over 45% of all indexed scholarly publications now have open references
- **36 million citation links** (made available thanks to the I4OC) **are now in Wikidata**
- **New members:**  
British Library • Microsoft Research • Harvard OSC • Citeseer • EigenFactor • ISSI
- We published an **open call to action** for stakeholders
- Additional **press coverage:**  
*Le Monde*, *EOS*, *British Library Blog*, *Hindawi Blog*

 **Le Monde Sciences**  
@lemonde\_science Follow

## Les « Open citations », des savoirs partagés librement sur Internet

[Translate from French](#)



**Les « Open citations », des savoirs partagés librement sur Internet**  
L'Initiative for Open Citations (I4OC) vise à rassembler des références dans une nouvelle base de données libre d'accès et d'utilisation. Elle est née sous l'impuls...  
[lemonde.fr](http://lemonde.fr)

# **Cross-Department Programs**

**(Just the Technology Department portions)**



**WIKIMEDIA**  
FOUNDATION

# Community Health

## Segment 3: Research on harassment

In the next fiscal year, we will continue research on harassment and abuse detection and deliver empirical insights and new models to better characterize toxicity in Wikimedia discussion spaces.



2017-10

**Research**X-CH [\[LINK\]](#)

## Outcome 1: We aim to understand and model the characteristics of harassment in Wikimedia projects.

### Objective 1: Conduct research to characterize and model wikipounding

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Analyze editor-editor interactions for instances of wikipounding</b>	LAST QUARTER	
	<ul style="list-style-type: none"> <li>Community Tech</li> </ul>	<ul style="list-style-type: none"> <li>Conducted exploratory work to characterize instances of <b>wikipounding</b> using qualitative data to identify possible sources of signal</li> </ul>
	NEXT QUARTER	
	<ul style="list-style-type: none"> <li>Community Tech</li> </ul>	<ul style="list-style-type: none"> <li>Conduct research on negative editor-editor interactions that are not topic-specific</li> </ul>

**STATUS: OBJECTIVE ON TRACK**

## Outcome 1: We aim to understand and model the characteristics of harassment in Wikimedia projects.

**Objective 2: Prototype new models to facilitate sockpuppet detection and the classification of toxic discussions.**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Improve open algorithmic solutions to classify toxic discussions</b>	LAST QUARTER	
	<ul style="list-style-type: none"> <li>Formal collaborators (Jigsaw, Cornell)</li> </ul>	<ul style="list-style-type: none"> <li>Finalized the talk page parser, the toxicity scoring strategy and the labeling schema for studying the evolution of discussions.</li> </ul>
	NEXT QUARTER	
	<ul style="list-style-type: none"> <li>Formal collaborators (Jigsaw, Cornell)</li> </ul>	<ul style="list-style-type: none"> <li>Publish an annotated dataset of conversation structures on English Wikipedia talk pages and submit a paper on <b>early indicators of toxic conversations</b></li> </ul>

**STATUS: OBJECTIVE ON TRACK**

## Outcome 1: We aim to understand and model the characteristics of harassment in Wikimedia projects.

**Objective 2: Prototype new models to facilitate sockpuppet detection and the classification of toxic discussions.**

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<b>Design new open algorithmic solutions to identify sockpuppets</b>	LAST QUARTER	
	<ul style="list-style-type: none"> <li>Formal collaborators (Stanford)</li> <li>Support &amp; Safety</li> <li>Community Tech</li> </ul>	<ul style="list-style-type: none"> <li>Completed exploratory phase to scope the project with input from multiple teams</li> <li>Announced the project in various community spaces</li> </ul>
	NEXT QUARTER	
	<ul style="list-style-type: none"> <li>Formal collaborators (Stanford)</li> <li>Support &amp; Safety</li> <li>Community Tech</li> </ul>	<ul style="list-style-type: none"> <li>Conduct a survey of CheckUser strategies</li> <li>Prototype new models to <b>identify potential sockpuppet accounts</b> on Wikipedia</li> </ul>




# Structured Data on Commons

## Segment 4: Programs

We will develop better understanding of existing needs for **Structured Commons**, including, but not limited to, media uploads, Wikidata usage for various programmatic applications, and the broader partnership applications of Wikidata beyond the Wikimedia ecosystem.




An image created as part of the Carpathian Ethnography project 



One of the images with improved metadata at the University of São Paulo Veterinary Museum 



Digitization of an unpublished manuscript of a participant of Napoleon's invasion of Russia at the Center for Retrospective Digitization, Göttingen, Germany (2008) 



## Outcome 2: Develop a better understanding of existing needs for Structured Commons

### Objective 1: Identify needs and priorities for this stakeholder group

What is your objective / workflow?	Who are you working with?	What impact / deliverables are you expecting?
<ul style="list-style-type: none"> <li>Complete and report GLAM stakeholder research interviews (A)</li> <li>Conduct and report survey of GLAM Commons contributors' workflows, wants, and needs (A)</li> </ul>	<b>LAST QUARTER</b>	
	Community Programs, Technical Collaboration, Niharika Ved (intern)	<ul style="list-style-type: none"> <li>Conducted 11 interviews with GLAM participants</li> <li>Presented incremental findings at SDC offsite</li> <li>Drafted GLAM participant survey</li> </ul>
	<b>NEXT QUARTER</b>	
	Community Programs, Technical Collaboration, Niharika Ved (intern)	<ul style="list-style-type: none"> <li>Analyze interview data, report findings (MetaWiki, meetings with SDC stakeholders, blog post)</li> <li>Release survey, collect results, analyze data and report findings</li> </ul>

# Support existing curation tools & workflows

Q images uploaded today with missing attribution ✕ Search

[Content pages](#) [Multimedia](#) [Translations](#) [Everything](#) [Advanced](#)

[Search categories](#) • [View other tools](#)

Create the page "**Images uploaded today with missing attribution**" on this wiki!

## Uploading Media

One file upload  
Uploading multiple files  
Multiple entry points

[File:Design ideas for Upload Wizard.pdf](#)

the workflow of **uploading** a file on Commons. Date 13 May 2014, 16:51:59 Source Own work Author Pau Giner Includes the following **images**: Cheetah\_Feb09\_02 (2,258 × 1,725 (82.21 MB)) - 08:35, 19 February 2017



[File:Karelia today.png](#)

Map of **today** Karelia The map was originally **uploaded** to w:**Image**:Karelia **today**.png by w:User:Jniemenmaa. (453 × 600 (32 KB)) - 06:46, 7 December 2014



[File:Honda Today.jpg](#)

This **image** was moved from File:Honday **Today**.jpg The Honda **Today** Originally **uploaded** by ja:利用者:Kuha455405 to ja:画像:**TODAY**.JPG. The original description

Gadget	Number of users	Active users
Cat-a-lot	6,542	1,694
QuickDelete	4,491	928
CropTool	2,518	883
GoogleImages	2,531	638
Tineye	2,739	554
purgetab	2,560	533
QInominator	2,014	484
edittop	3,209	471
AddInformation	2,734	465
popups	2,756	449
CatScan2-link	1,755	438
UserMessages	2,482	424
Gallerypreview	4,768	391
Glamorous	1,662	358
UTCLiveClock	3,008	356
ThumbnailPurger	2,478	349
Geotoolbox	1,930	323
GalleryDetails	2,718	319
markAdmins	1,036	315
WhatsIsThat	2,227	307
GlobalUsageUI	1,213	302

# Wikidata Integration: Learning from current approaches

The **WikiProject sum of all paintings** is a WikiProject to get an item for every notable painting.\*

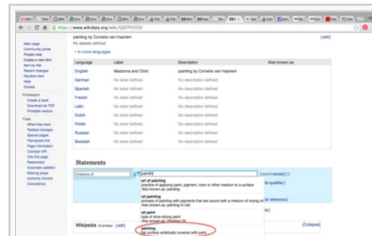
**Contents** [\[show\]](#)

## Background [\[edit\]](#)

On Wikipedia we have a lot of experience breaking down large amounts of work into smaller chunks that can be completed. Examples of this are the windmill project on the Dutch Wikipedia and the many lists of historic buildings that were created for [Wiki Loves Monuments \(Q1353202\)](#). For paintings we can use the same strategy: We break it down into smaller parts that can be handled and then we can utilize Wikidata to celebrate the sum of all paintings through linked open data!

## Scope [\[edit\]](#)

All or everything is too broad and has no limit. This limit is to be able to handle the processing of the data - let's not overwhelm ourselves (yet). The scope can always be expanded - in the future we will have the sum of all sculptures, drawings, you get the picture! The scope of this project is [notable](#) paintings. A painting is considered notable if one of these options is true:



Add a painting! See also [this video](#) by [Shonagon](#)

July 2017

**Research****Structured Data****Outcome 2: Develop a better understanding of existing needs for Structured Commons****Objective 2: [Perform research to] allow project dev. among WM communities and identify gaps in existing tools****What is your objective / workflow?****Who are you working with?****What impact / deliverables are you expecting?**

- Interview 6-10 volunteer contributors to Commons to understand current workflows and impact of structured data (A)

**NEXT QUARTER**

Community Programs, Technical Collaboration, Multimedia, Discovery

- Conduct interviews, begin data analysis, share incremental results with Multimedia & Discovery teams

**STATUS: OBJECTIVE ON TRACK**