

Wikimedia Foundation metrics meeting

28 July 2016



Agenda

A long, narrow alleyway between two-story stone buildings. The buildings are constructed from rough-hewn, light-colored stone blocks. The windows are rectangular with blue frames and some have metal grilles. The ground is paved with dark stones. The alleyway leads to a bright opening at the far end. The overall atmosphere is historical and rustic.

Welcome

Community update

Metrics

Feature

Research

Product demo

Questions and answers

Welcome!

Requisition hires:

- Gretchen Yen - Executive - SF

Contractors, interns & volunteers:

- Daniela Andrade - Communications - SF
- Helen Jiang - Product - SF
- Blanca Flores - Communications - SF



Anniversaries

Aaron Schulz (8 yrs)

Arthur Richards (6 yrs)

Michael Beattie (5 yrs)

Jeff Green (5 yrs)

Niklas Laxström (5 yrs)

Tilman Bayer (5 yrs)

Lynette Logan (4 yrs)

C Scott Ananian (3 yrs)

Dennis Porter (3 yrs)

Nick Wilson (3 yrs)

Tighe Flanagan (3 yrs)

Bryan Davis (3 yrs)

Kristen Lans (2 yrs)

Josephine Gulingan (2 yrs)

Katie Francis (1 yr)

Trey Jones (1 yr)

Sheree Chang (1 yr)

Mikhail Popov (1 yr)

Emerald Ross (1 yr)

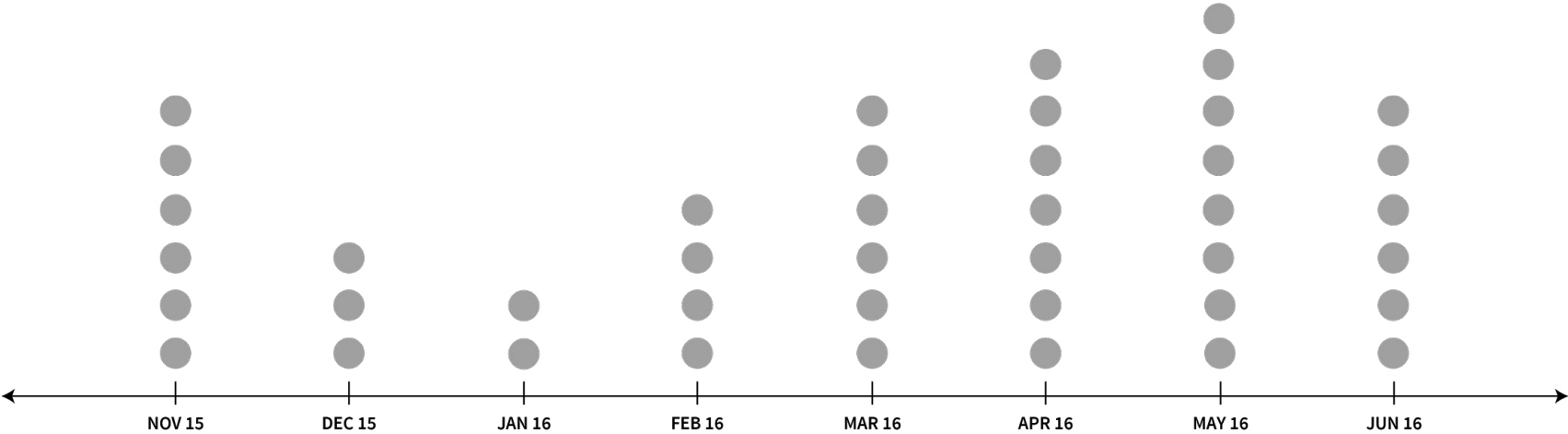
Community update



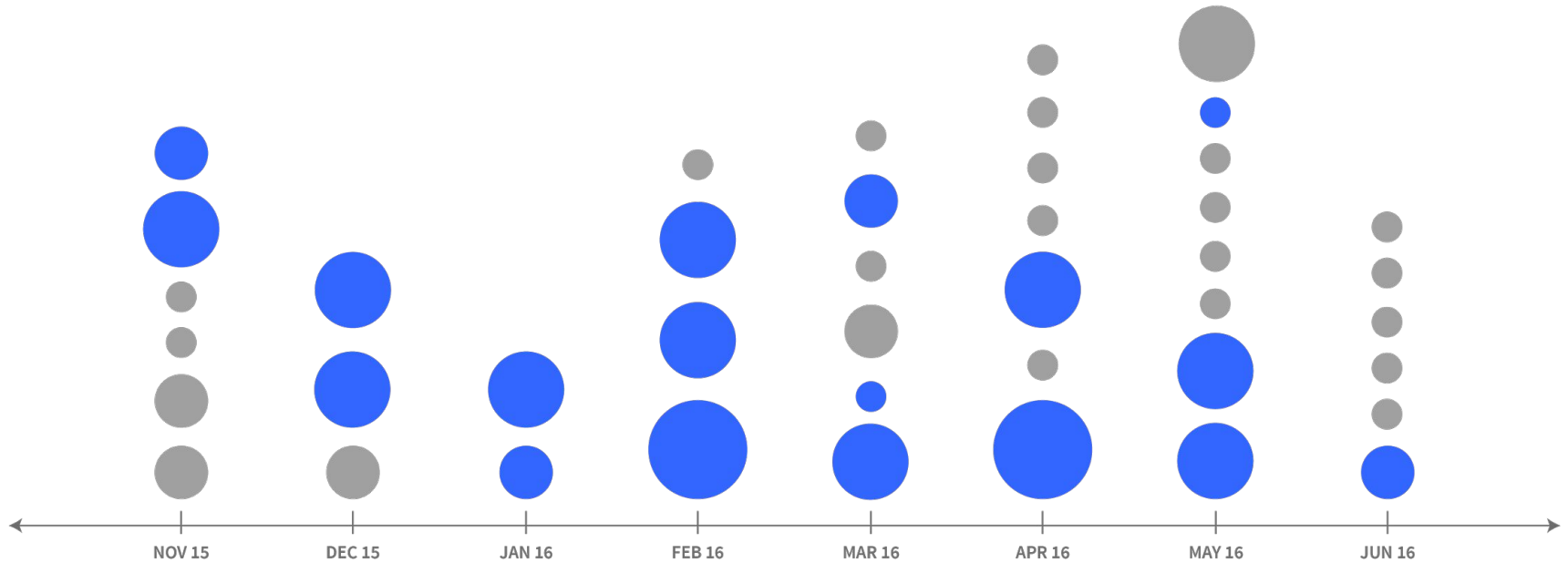
In the last 8 months we consulted with
communities many times.

We tracked 42 of these consultations.

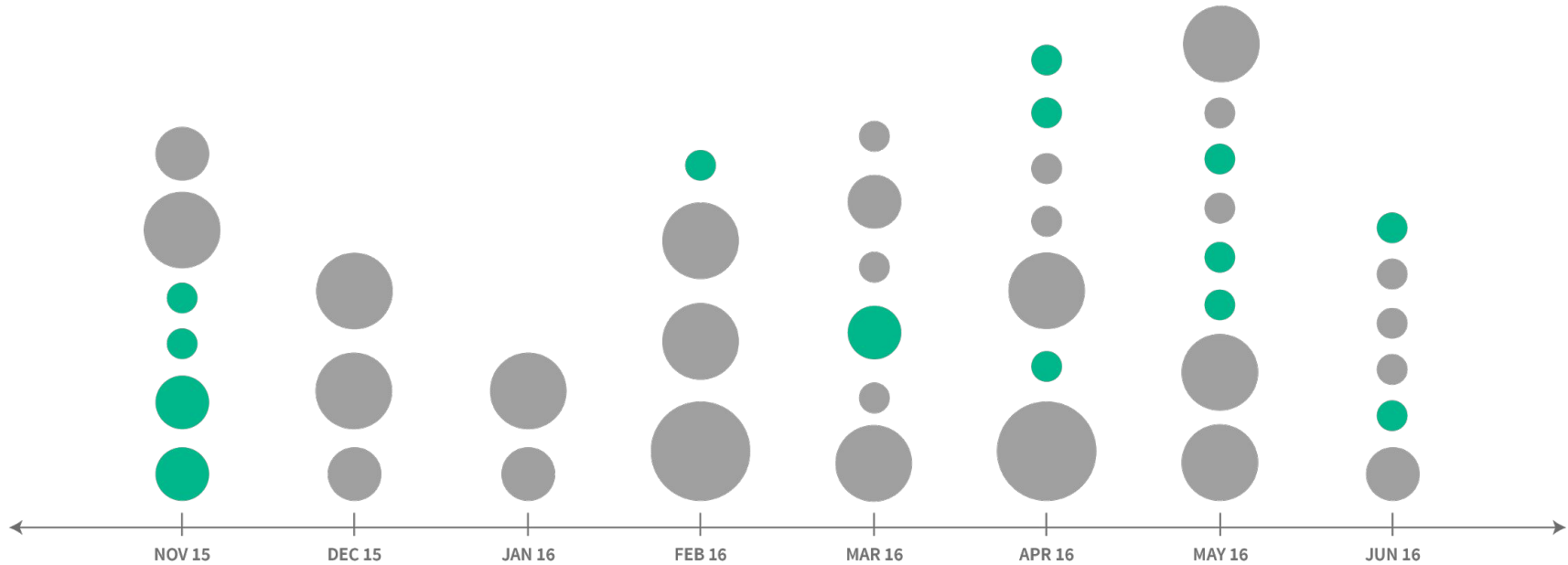
This is how consultations were distributed in the past 8 months



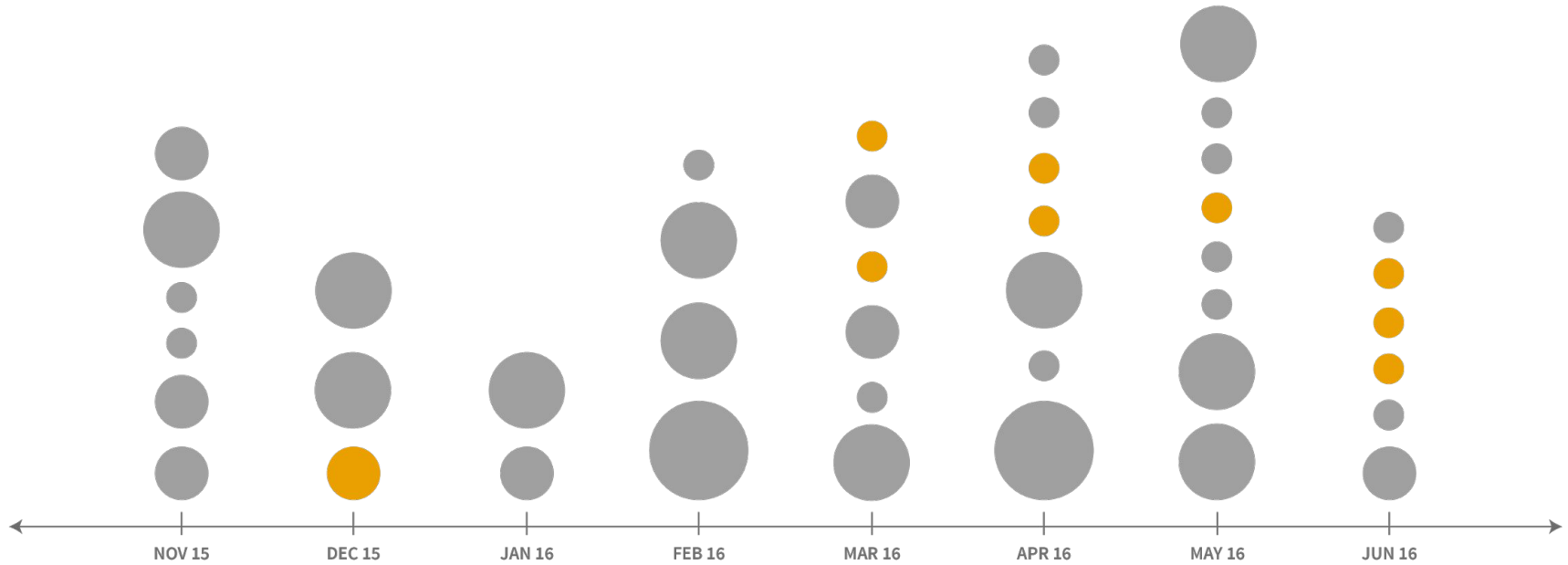
18 were conversation consultations (43%)



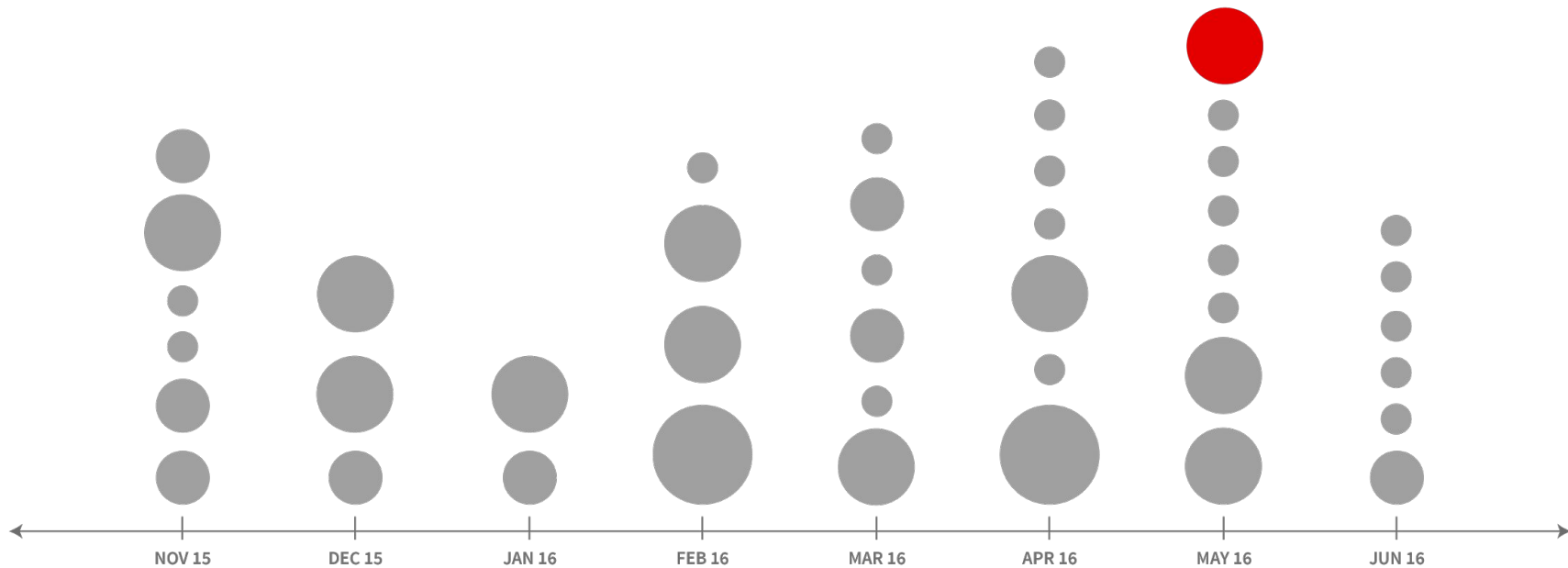
14 were surveys (33%)



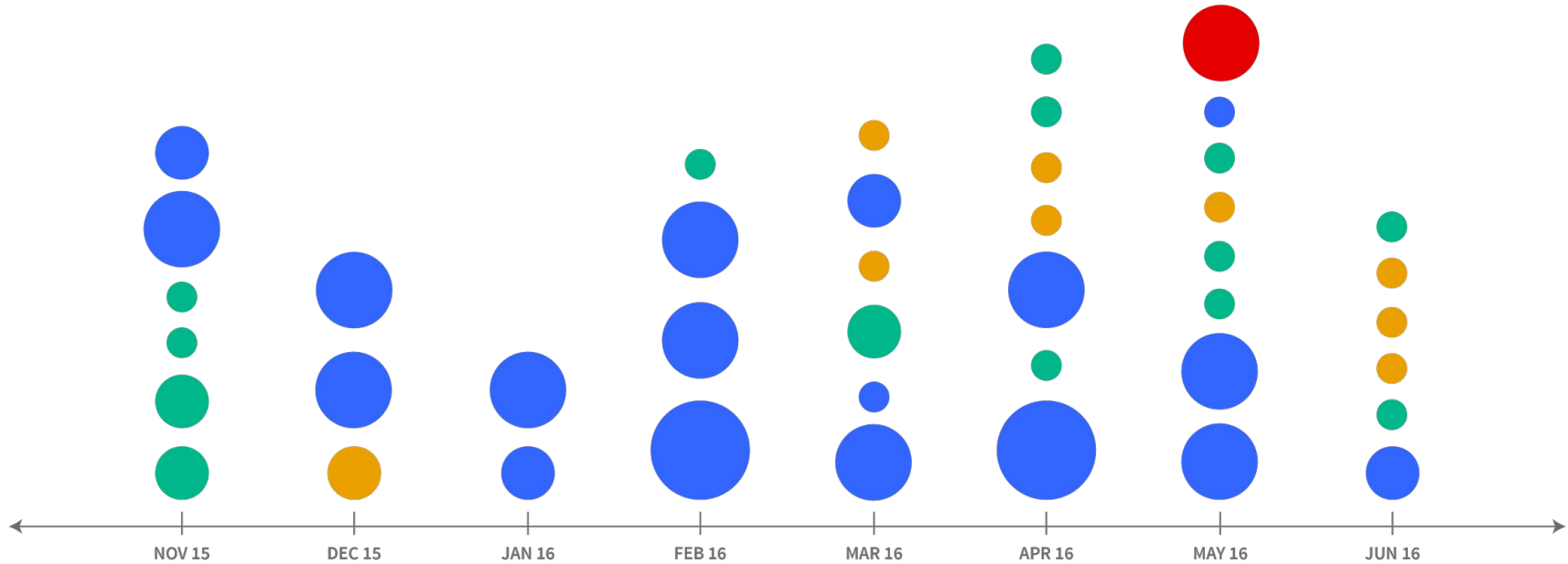
9 were events (22%)



1 was a community vote (2%)



... all together form a colorful map of consultations.

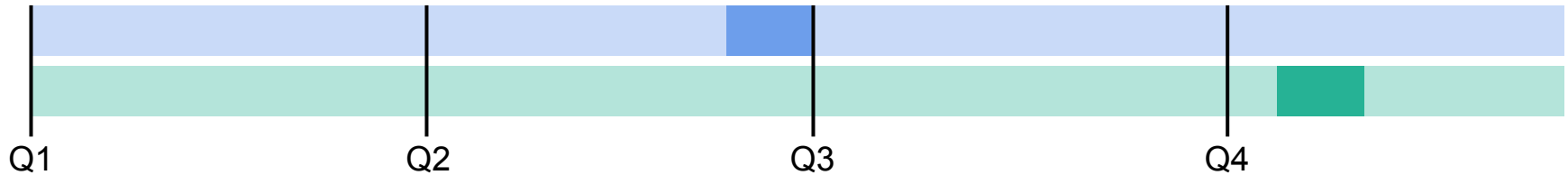


Conversations matter.

CE Insights: 2 surveys this year

One performance survey
in December

One pulse survey
in April/May



Do you need data from these communities to plan next year?

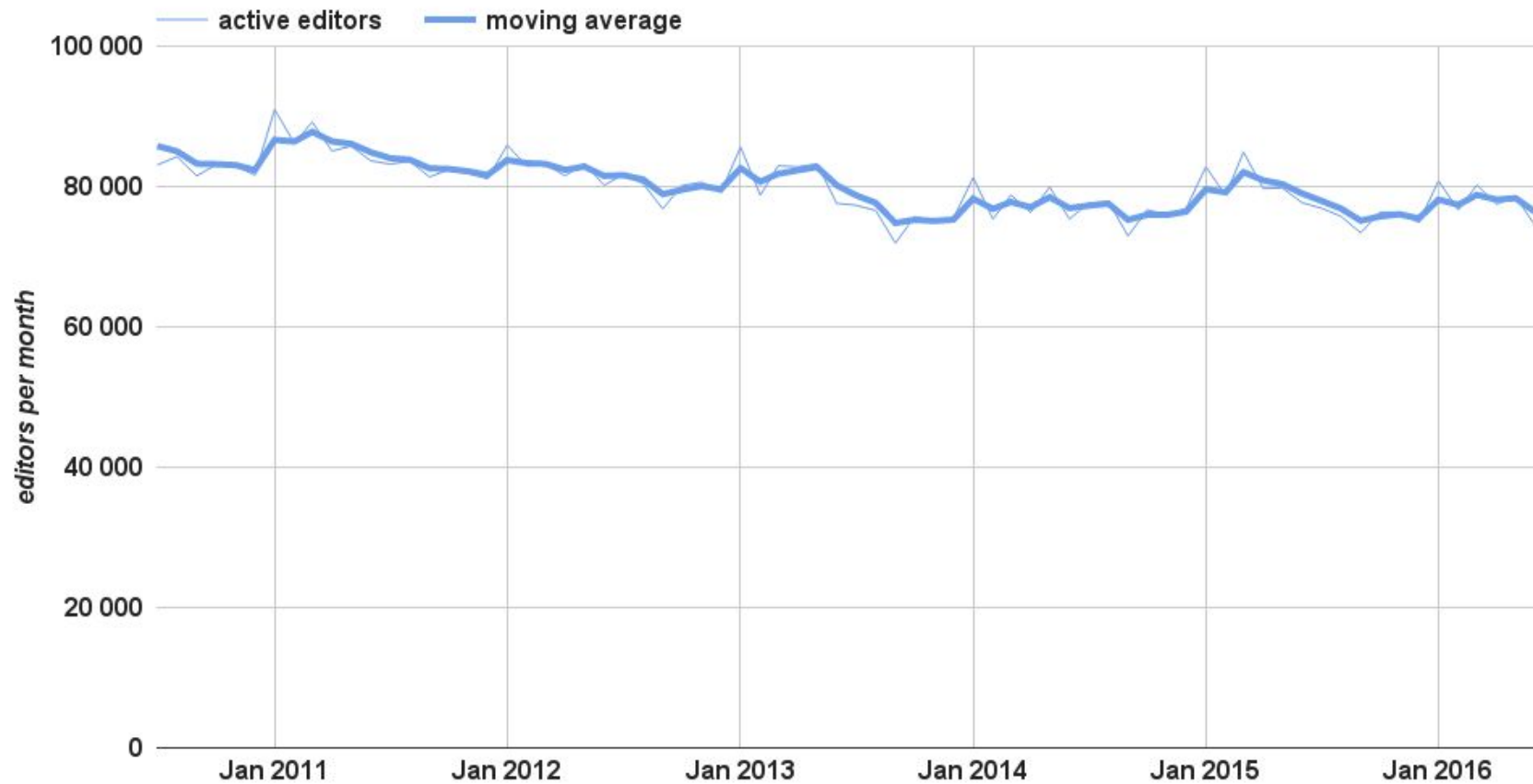
Letter of Intent due
August 5th!

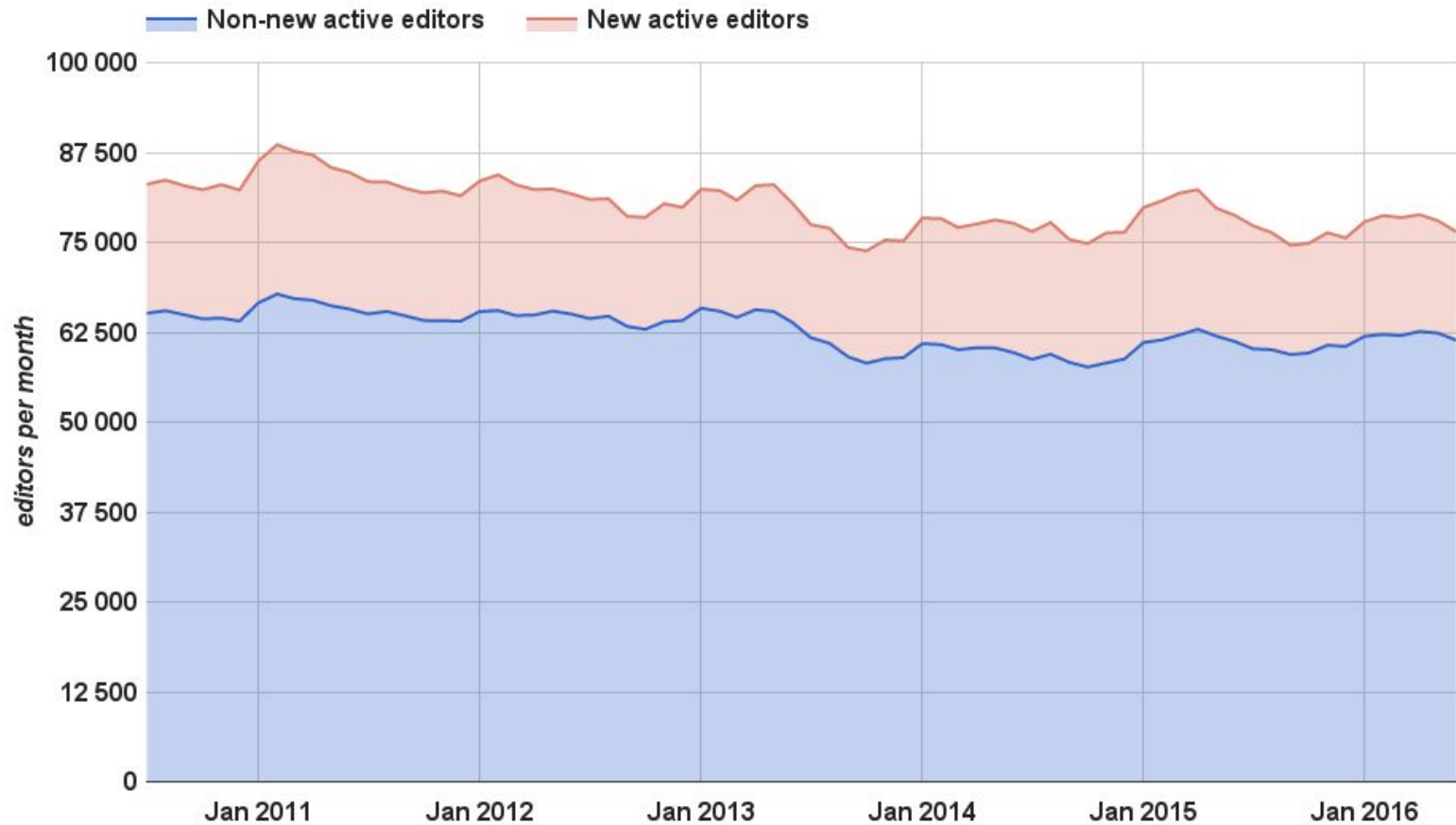
- Very active/involved editors
- Active editors who curate/create content
- Wikimedia Affiliates or program leaders
- Volunteer developers, technical collaborators
- External movement partners

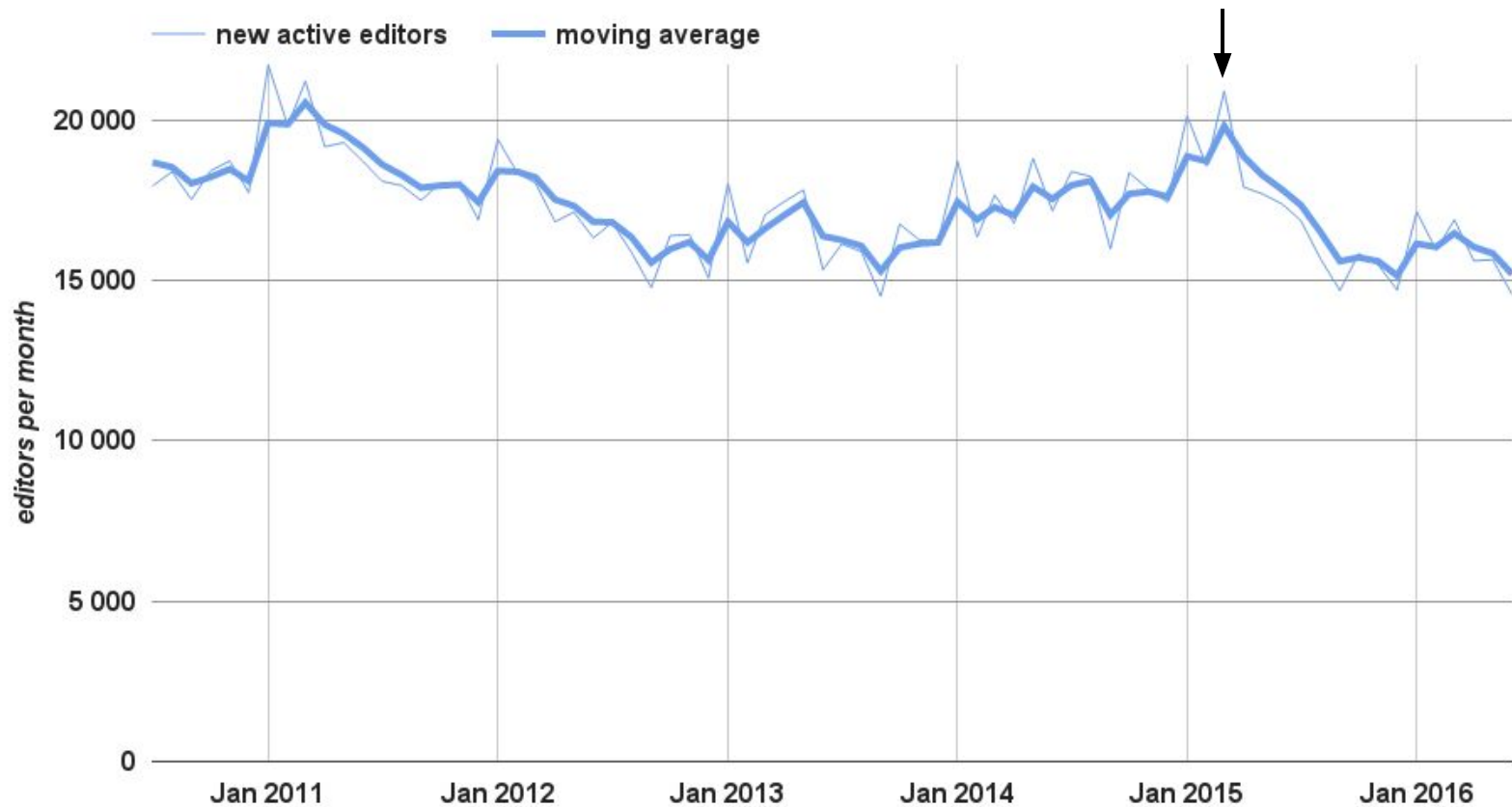
Talk with your team about [meta:CE Insights](#)

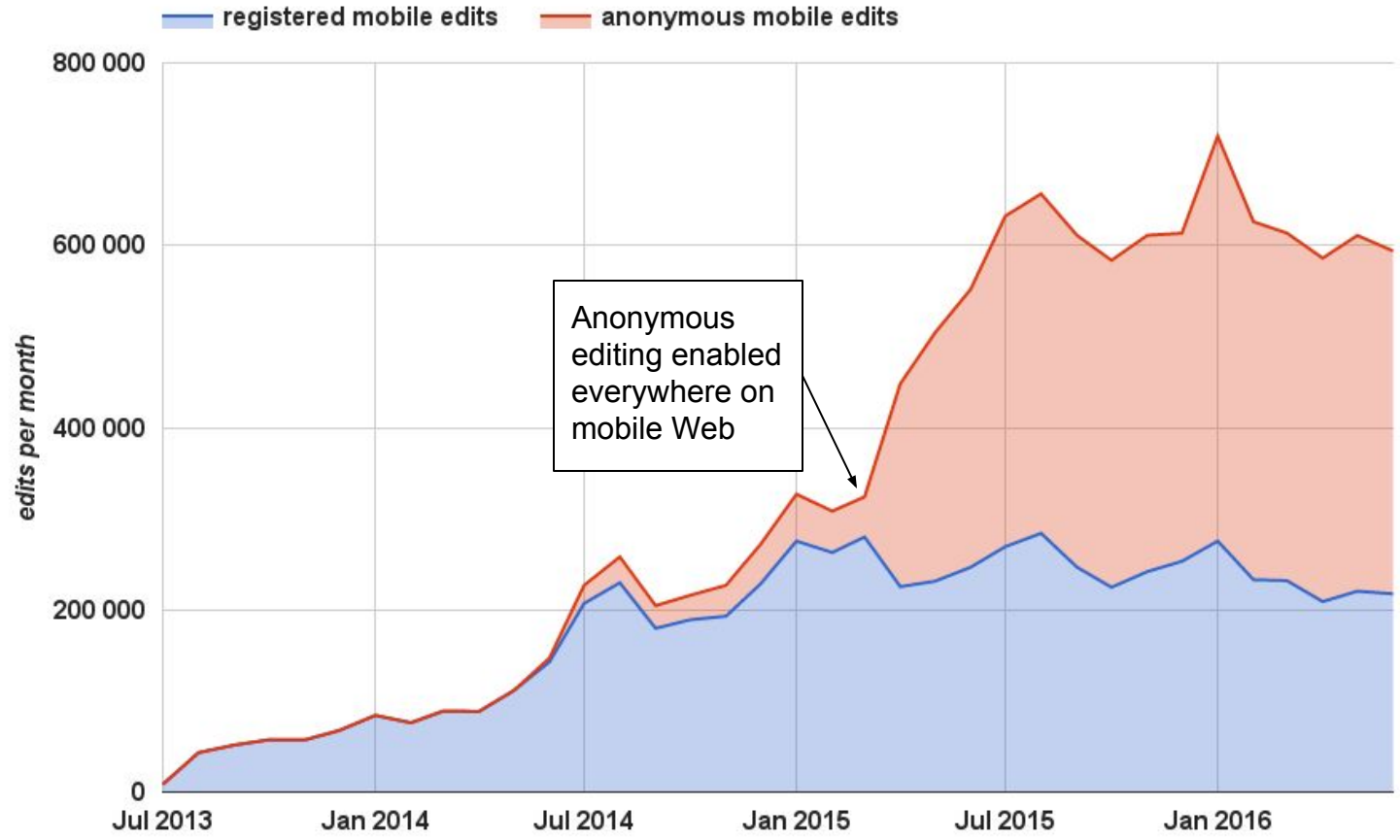
Metrics

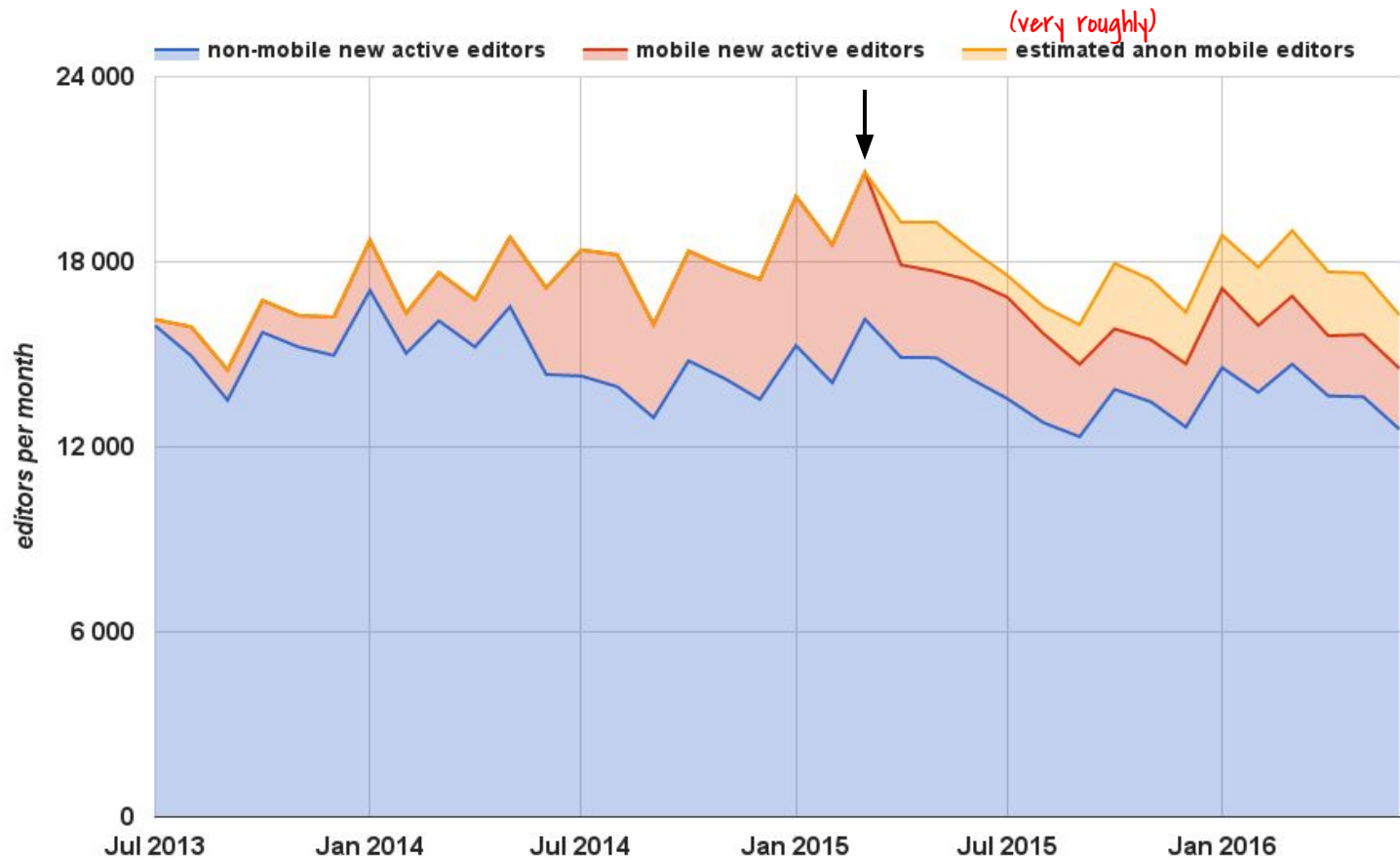


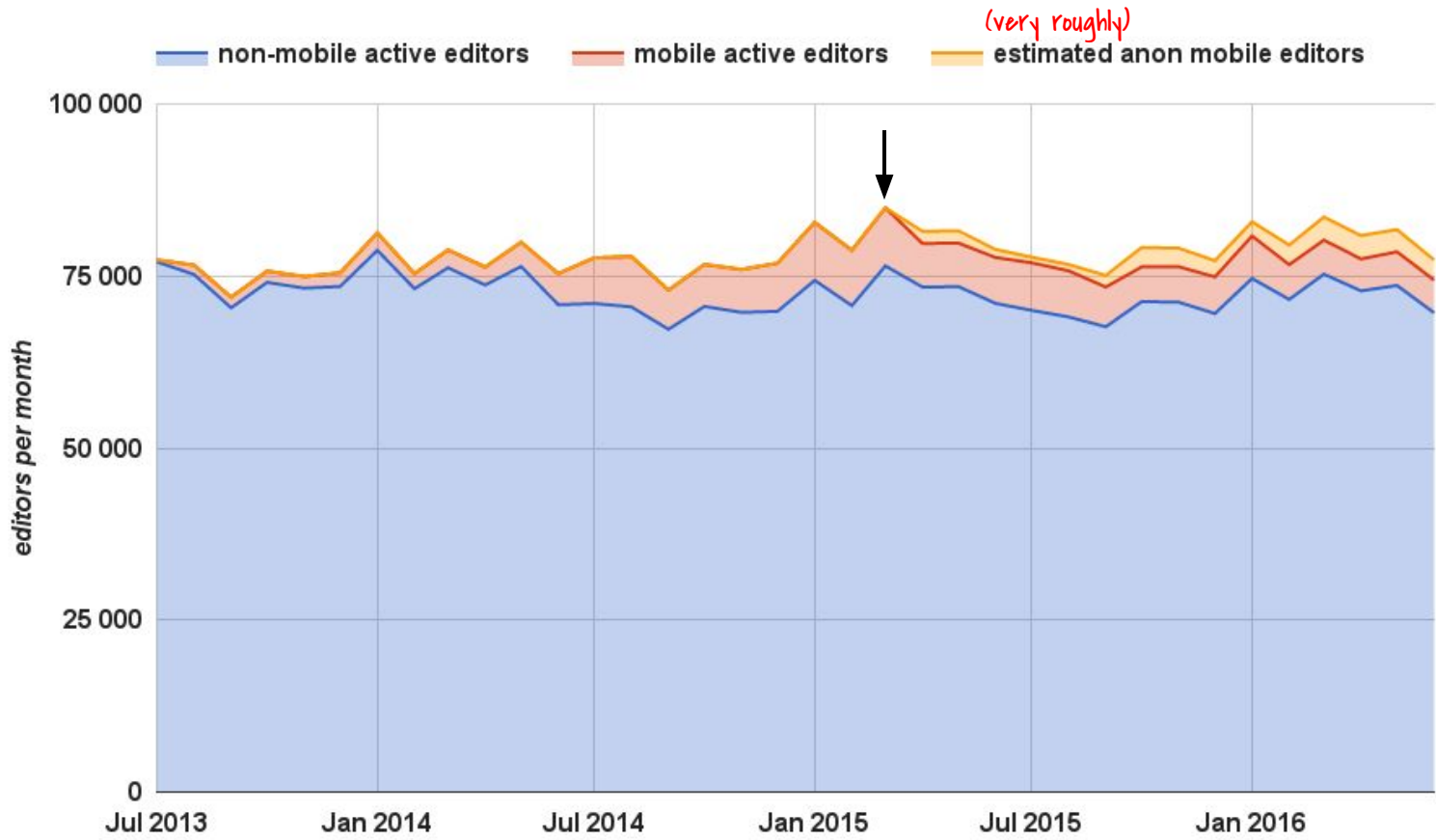












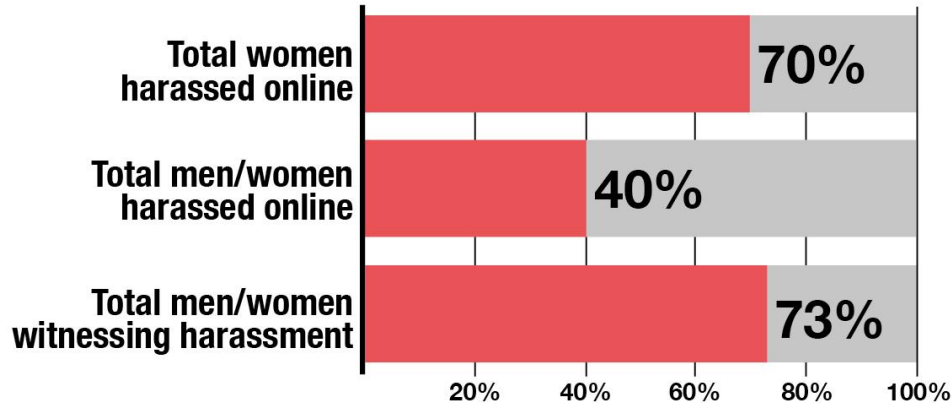
Feature





Community culture: Harassment and healthy environments

Harassment is a challenge across the internet



Harassment obstructs community health, growth, and inclusion

- Impact on people
 - 78% found distressing;
 - 27% extremely distressing
 - Women twice as likely to find harassment extremely or very distressing
- Widely misunderstood, poorly defined, sometimes denied
- No silver bullet or solutions: combination of tools, management, culture

*Pew Research Center study on Online Harassment (2014)

**UN (2015), "Cyber Violence Against Women and Girls"

How does it affect Wikimedians?

Deterrents to contribution identified by our most recent Editor Survey:

- Criticism of their work from others **27.6%**
- Editors are “not fun to work with” **23.6%**
- Harassment from other editors **10.8%**

What does this mean to the projects?

- How does harassment affect our community: What forms does it take, how does it affect motivation, how effective are our current practices?
- How does the community want to address the problem?
- What alternatives exist: How do other online communities address the issue? What does the research have to say?
- How do Wikimedia's technical infrastructure and social values contribute to the problem? What can be improved?

We surveyed the issue

- Evaluated external research and practices
 - With external researchers and community, started the [Online harassment resource guide](#) (a literature review of existing academic research)
- Reviewed other community policies, published summaries and findings
 - [Online Community Conduct Policies](#)
- Collected data on the problem (2015)
 - [Harassment consultation](#) and workshop
 - [Harassment survey](#) (in 16 languages)

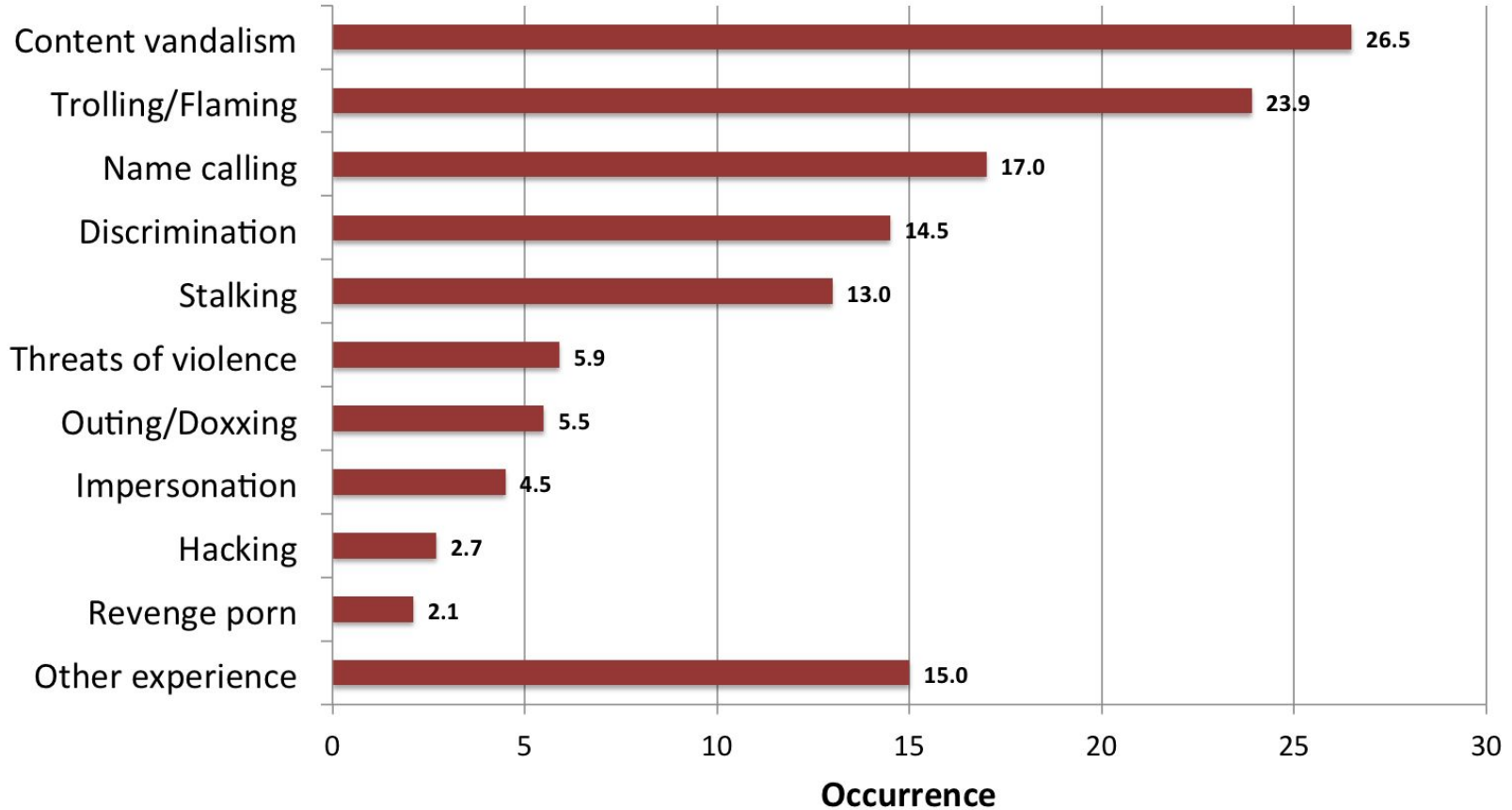
Consultation and survey goals

- Collect baseline information
- Gauge the community's perception and self-assessment of harassment
- Show fail points
- Invite the community to engage
- Raise awareness of the issue
- Drive changes in process or policy
- Outline measures of success

Key survey findings

- Women and marginalized genders are more likely to experience harassment on our projects than men.
- 88% of people reported being upset by their experience with harassment, with 35% describing it as “very” or “extremely” upsetting.
- More than **50%** of people who reported being harassed reported subsequently **decreasing their participation in Wikimedia.**
- 42% of responders found their efforts to address harassment ineffective.
- 51% of people who witnessed (rather than experiencing) harassment did not become involved. **67%** of those who did attempt to intervene reported experiencing some level of **retaliation.**

Forms of harassment



"Your Writings are shit !"

"you don't know the first thing about..."

"nasıl bu kadar onursuz olup ama ksine görünüşte iyi ve salak olabiliyorsın?"

(How can you be such dishonorable and at the same time seem like a good and stupid person?)

"All queers will be shot! You fucking faggot, I hope you burn in hizzell!"

"HIER RIECHT ES NACH KACKE!"
("HERE IT SMELLS OF SHIT")

"DIE CIS SCUM"

***"Arrête avec tes bandeau à la ***
et va jouer à la guerre ailleurs"***

("Stop with your dumb banners and go play little soldier elsewhere")

"Was fürn SCHEIB!!!"
("What SHIT!!!")

**"I am going to kill
your grandchildren"**

"Soy un hijo de la grandisima puta"

(text impersonating an editor:

"I am the child of the biggest whore")

"son of a bitch"

"Oh you aRE ALL I GIVE YOU PERMISSION YOU
DICK..... THEY TELLYOU"

***"What entitles a feminized nebbish like
you to delete a book that you haven't even
read"***

"Reichstag Barkstein"

"With Jews you win"

**"Dieser benutzer hat keine arbeit und schreibt sinnlosen
kl_se unter jeden mist den man aus spass kommentiert.**

MfG Your asshole lickin' OG Loc"

*(This user is unemployed and writes senseless crap below all shit that other
people comment on for fun.. With regards, Your asshole lickin' 'OG Loc)*

"Halt die Fresse, du Arschloch"

(Shut up, you asshole)

"You're an old fart
and so is your friend"

The good news:
This is an opportunity to lead

The community supports change

2015 strategy survey's second-most supported point

Reduce harassment issues and the gender gap to facilitate a safe, welcoming, and supportive environment for contributors and editors.

Where does the community want change?

Five key areas of improvement:

1. More data on the problem
2. Technical-based solutions
3. Improvements to training for community members with advanced rights
4. Improved dispute resolution practices
5. Improved policies and enforcement

What is underway?

- Completed and implemented an [Event Ban Policy](#) to support local organizers to protect event participants
- Created [Friendly Spaces Expectations](#) in Meta grantmaking space; working with community on proposed technical spaces [Code of Conduct](#)
- Delivered Community Capacity Development [conflict management](#) pilot module (Ukrainian comm.)
- Partnered with thought-leaders on finding information and solutions ([Danielle Citron](#) at WikiConference USA, Congresswoman [Katherine Clark](#), the Berkman Center [workshop on misogyny](#), participating in the Berkeley [roundtable on online harassment](#))
- Began designing [an early-detection tool for harassment](#), called Detox, along the lines of our machine-learning projects around vandalism detection

What's next?

Inspire Campaign: 280 proposals in 30 days; 700 participants

- "Protect user space by default" - Explores protection levels for user pages, which are a common location for abusive edits
- "Create a help page in all Wikipedias" - Proposes coordinated improvement of help documentation regarding harassment across all projects
- "Study how current mechanisms handle harassment" - A research proposal: Do we know where current processes break down? Where they are effective?

Training community leaders: translatable online training modules

- Best practices for handling harassment reports
- Event organizers training

Questions we face

- What are our responsibilities to our users?
 - What are our obligations to address harassment as a matter of values?
- What is our appetite for meaningful change, and what does that look like?
 - What are the risks of action? What are the risks of inaction?
- What would global policies or enforcement look like?
 - How would we address this in an inclusive, global fashion?
- What role should we play in shaping the broader internet discourse?

Opportunities to discuss will be forthcoming in public consultation. Meanwhile, if you have thoughts to share, please email ca@wikimedia.org.

Research



Modeling personal attacks on English Wikipedia

Goals

1. Develop an algorithmic approach to detect personal attacks on Wikipedia
2. Use these algorithms to extend the analysis of personal attacks on Wikipedia

Outline

1. Data Pipeline
2. Model Building
3. Analysis

Outline

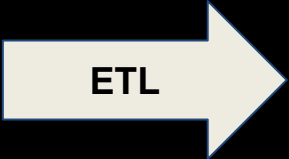
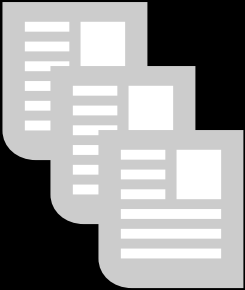
1. Data Pipeline

2. Model Building

3. Analysis

Data Pipeline

Revision History



Raw Comments



Labeled Comments



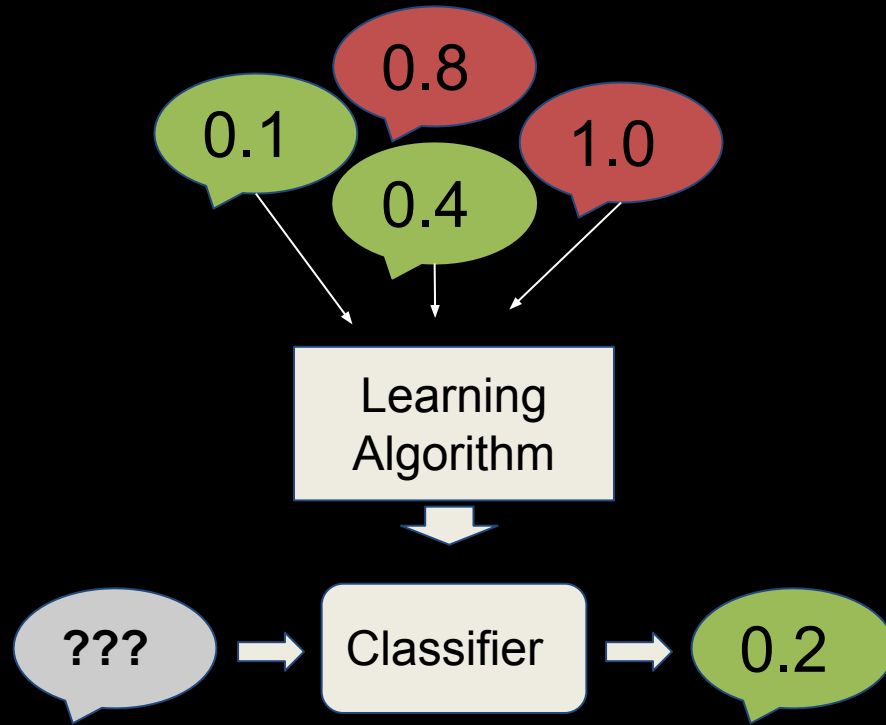
Outline

1. Data Pipeline

2. Model Building

3. Analysis

Model Building



Model Evaluation

Question:

How good is our classifier?

Idea:

Compare our model's judgements, to pooled judgements of a group of people.

Model Evaluation

Prediction Group Size	ROC AUC
1	0.854
2	0.911
4	0.941
6	0.950
8	0.961
10	0.963

Fix "Ground Truth group size at size 10

Model:
0.951

Demo

Available at: *wikidetox.appspot.com*

Demo

Select Input Type:

- Text
- Revision ID

Congratulations. I don't know whether you are aware of this fact or not, but you have shown your qualified stupidity.

Score

Results:

not attack: 0.18

attack: 0.82

Demo

Select Input Type:

- Text
- Revision ID

F#@\$ you, a\$\$h0l3

Score

Results:

not attack: 0.31

attack: 0.69

Demo

Select Input Type:

Text
 Revision ID

I will punch your lights out.

Score

Results:

not attack: 0.41
attack: 0.59

Select Input Type:

Text
 Revision ID

Let's drink punch.

Score

Results:

not attack: 0.83
attack: 0.17

Demo

Select Input Type:

- Text
- Revision ID

Your intellect is lacking

Score

Results:

not attack: 0.90

attack: 0.10

Outline

1. Data Pipeline
2. Model Building
3. Analysis

Analysis

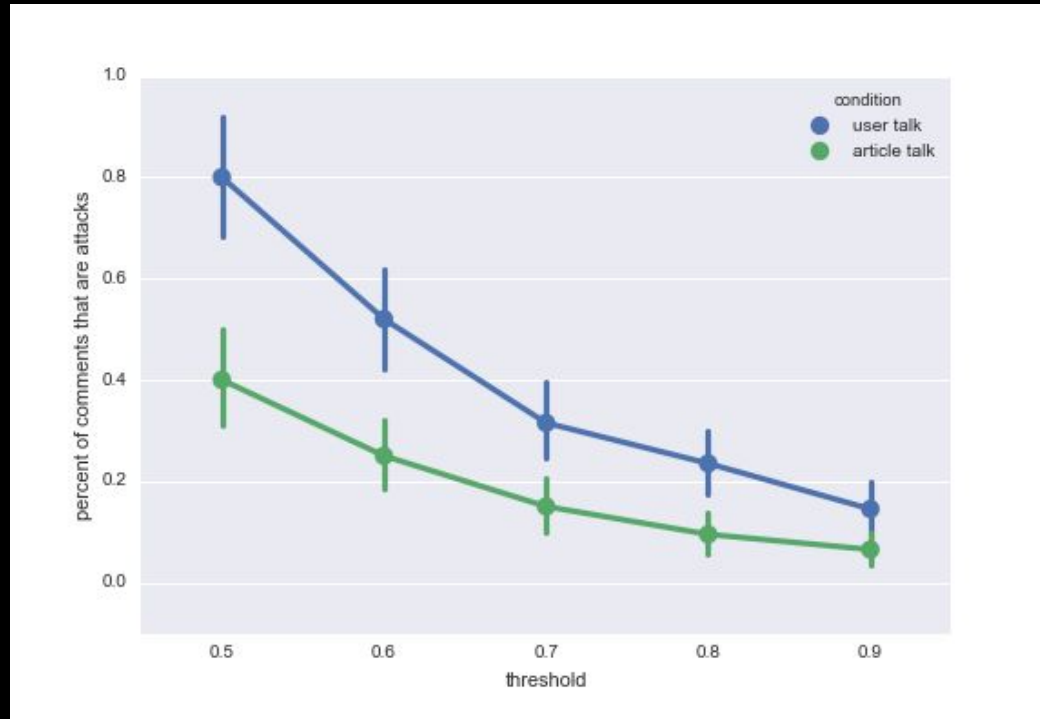
Goal:

Explore prevalence, dynamics and impact of personal attacks on English Wikipedia

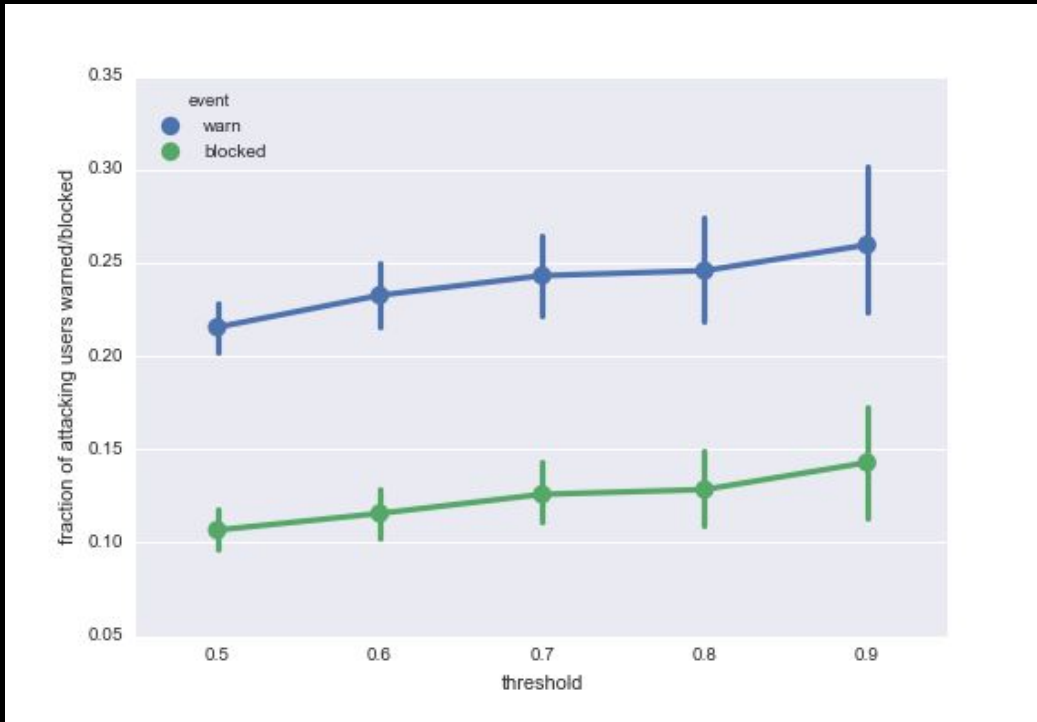
Input:

Complete historical data set of talk page comments + classifier scores

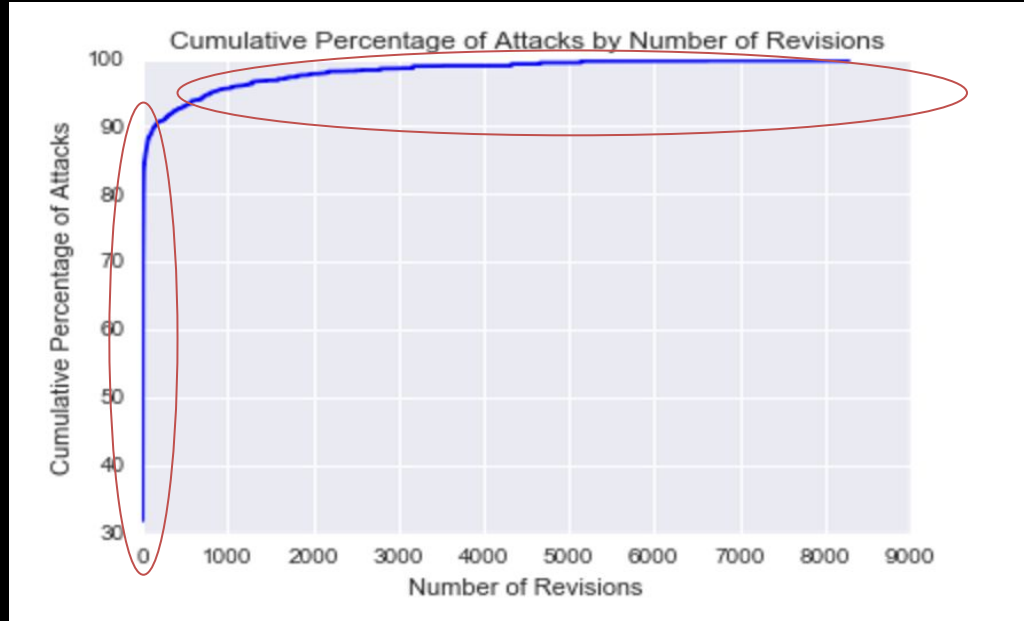
How many comments are personal attacks?



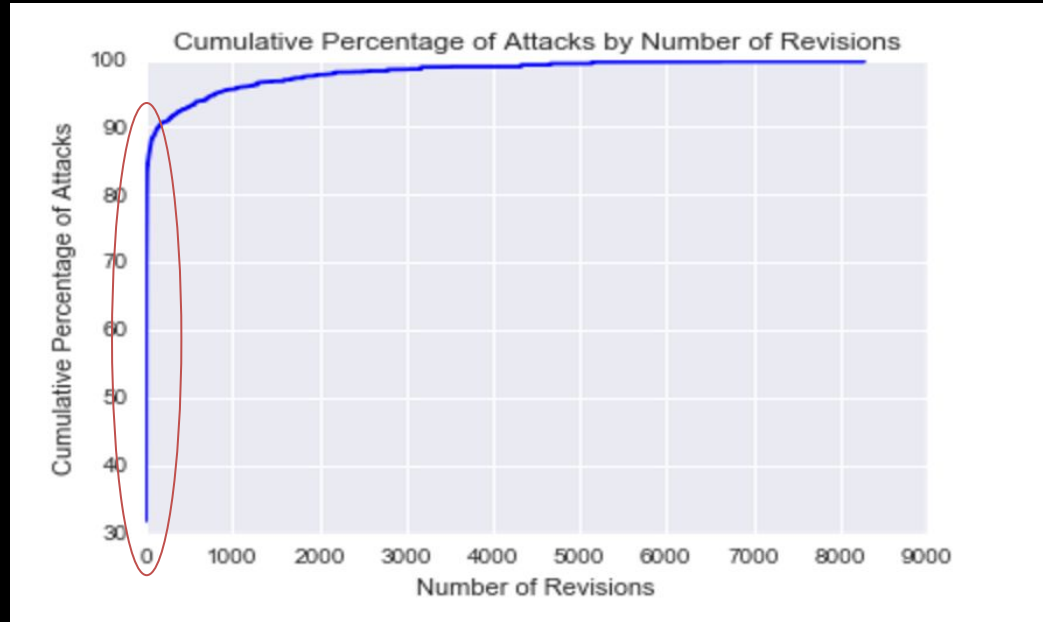
How many attackers have been warned/blocked?



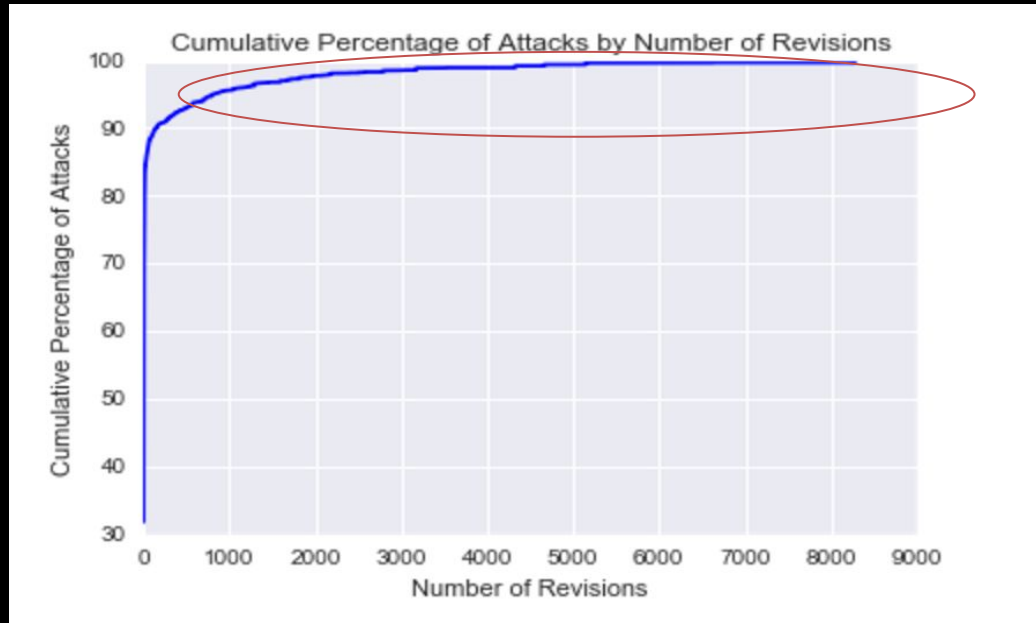
Two major types of attackers



75.7% of attacks come from users that have made fewer than 10 total revisions



9.3% of attacks come from users with over 200 total revisions



Next steps

- Improve Modeling
- Extend Analysis
- Release of Annotated Datasets
- Integration with ORES

For more information

meta.wikimedia.org/wiki/Research:Detox

Product demo



All-New Notifications Page

Helping Wikimedians
focus on what matters most

Notifications

Preferences

Recent activity


Pages with unread notifications

English Wikipedia

91

 Talk:ET61 moved

21

 User talk:ET13

16

 Talk:ET23 1

16

 Talk:ET4

8

 Talk:ET33

7

 Talk:ET2

6

Catalan Wikipedia

6

 Discussió:ET10

5

 Usuari Discussió:ET13

1

All

Read

Unread

Tuesday 5 July 2016

Mark group as read

 New topic created on **Talk:ET23 1: "Lorem ipsum dolor sit amet, consectetur adipiscing..."**.
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Praesent vel metus enim. Praesent condi...
 Etonkovidova  ET23 1 ... 2d

 New topic created on **Talk:ET23 1: "Lorem ipsum dolor sit amet, consectetur adipiscing..."**.
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Praesent vel metus enim. Praesent condi...
 Etonkovidova  ET23 1 ... 2d

Monday 4 July 2016

Mark group as read

 ET10 thanked you for your comment in **"May 23"**.
test on May 31
 ET10  ET2 ... 3d

 New topic created on **Talk:ET2: "July 04"**.
1
 ET1  ET2 ... 3d

Notifications System and Page

Maintained by: [Collaboration Team](#)

Notifications [Talk Page](#)

[Notifications Phabricator board](#)

[Notifications page Phabricator board](#)

Q&A



wikilove
LIVE