AGENDA

- Expectations for this meeting
  - Focus on health metrics
  - Iterating and improving (combined meeting, new data, etc.)
  - Further suggestions welcome
- Context: better use of data
- Segmentation project
- Audiences highlights
- Reading health metrics
- Contributors health metrics - an introduction
Data improvements

● In order to use data more effectively to plan and measure our interventions, we are investing in our data pipeline and approach

● Current initiatives:
  ○ Better use of data annual plan program
  ○ Consolidating “product analytics” team, hiring manager
  ○ Hiring data engineer
  ○ Segmentation project
  ○ DACI and stewardship

● For the next fiscal, see our department’s annual plan
Close up on segmentation

**Phase 1:**
**Snapshot**

*Complete:* June 1

*Purpose:* Targeting, Analysis

*Description:* Initial exploration of data available across a small number of dimensions

*Use case:* would allow PM to find the perfect wiki to test a new homepage feature on

*Deliverable:* Big sortable spreadsheet with 20+ dimensions you can filter and sort by.

**Phase 2:**
**Manual Segmentation**

*June 14, 2018*

*Diagnostic*

Pull out 3-6 key dimensions, create 3 bands within each of them (e.g. small, medium, large)

Use to support target audience selection for next FY’s features.

These bands added to every dashboard.

**Phase 3:**
**Emergent Segments**

*TBD based on Analyst availability*

*Targeting, Correlations*

Begin work using dummy data, when Phase 1 is complete, plug in real data and cluster wikiprojects across all dimensions covered in Phase 1

Strategy/Product wants to develop interventions that target specific wiki-clusters

Baseline snapshot of emergent segments, documentation

---

Project Brief
Potential dimensions (a sample - we have 38 atm)

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project</td>
<td>Which project family (Wikipedia, Wiktionary, etc.) the wiki belongs to</td>
</tr>
<tr>
<td>Lifestage</td>
<td>Number of years in existence</td>
</tr>
<tr>
<td>Size</td>
<td>Number of articles</td>
</tr>
<tr>
<td>Admins</td>
<td>Currently active admins</td>
</tr>
<tr>
<td>Speakers</td>
<td>Global population of speakers for this language</td>
</tr>
<tr>
<td>Speakers to Editors</td>
<td>Ratio of speaking population to editors</td>
</tr>
<tr>
<td>Devices</td>
<td>Unique devices per month</td>
</tr>
</tbody>
</table>
Metrics highlights

- This quarter, the year-over-year decline in total pageviews became smaller, and even vanished afterwards in April.
  - Apparent causes include last summer’s partial rollout of the page previews feature and a decline in Google referrals since last year. The impact of the former grew in April with the completion of the Previews rollout, but Google referrals may have recovered in April.

- In January, global new editor retention hit its highest level since 2007
  - Apparent cause is the trend towards higher and higher January and September spikes on the English Wikipedia
Total Wikimedia pageviews, by access method

Pageviews per month (average in Q3 2017/18)

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Desktop</th>
<th>Mobile web</th>
<th>Apps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic</td>
<td>15.9 billion</td>
<td>7.4 billion</td>
<td>8.3 billion</td>
<td>238 million</td>
</tr>
<tr>
<td>YoY</td>
<td>-3.8%</td>
<td>-11.4%</td>
<td>+3.7%</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Corrected for artificial IE7 traffic from Pakistan and some other Asian countries.
In April, shortly before completing the rollout of the feature, we began to systematically measure Page previews alongside pageviews, as a major new way of reading Wikipedia content.

There is now about one preview for every four pageviews on desktop.

We are planning to combine both into a new “page interactions” metric, but do not yet have a quarter’s worth of data.
Long-term trends
Globally, total readership has been slightly declining since 2013 (-2.3% yearly, largely attributable to several one-time effects).

In the last two quarters we have seen a year-over-year decrease that was faster than this 5-year trend (-3.8% in Q3, -5.2% in Q2). However, this might have stopped, with April 2018 reaching the levels of April 2017 again.
See below for some apparent causes.

Long-term, mobile has been increasing (+18.9% yearly since 2013) as desktop has been declining (-13.3% yearly since 2013).

Corrected for anomalous IE7 traffic from some countries and several other artifacts.
Total pageviews had been consistently down y-o-y since July (compare blue and purple), but picked up right at the end of the quarter (compare red and blue).

Trend chart includes corrections for some anomalies, e.g. artificial IE7 traffic from Pakistan and some other countries.
Desktop has been shrinking y-o-y in general again, but appears to have picked up since the end of the quarter, only interrupted by the completion of the Page previews rollout to de and en in April. This new feature is designed to save readers some clicks, so a decrease (3%-5% per the last A/B tests on de and en) is expected and to some extent desired.
Mobile growth remained slower since last quarter (see how blue and purple appear closer together since ca. November), but may also have picked up in April.

Note the annual “Christmas bumps.”

(As a desktop-only feature, page previews don’t affect mobile.)
Since late December, mobile pageview now outnumber desktop pageviews basically every single days (even on weekdays).
(Data corrected for an iOS bug from around Dec 20, 2016 on, and for anomalous IE7 traffic from some countries.)
“Seventy-five percent of the world’s online population is from the global South” (Whose Knowledge)

Ratio of total pageviews, Jan-Mar 2018 (definition)

77.2% Global North
22.8% Global South
Potential impacts on traffic:

Big:
- Page previews - see above
- Google search engine referrals - see below

Small or unknown:
- Singapore datacenter - more on next slides
- Facebook news media links increased views to certain articles about news media, but the overall effect is negligible (even among all FB referrals)
- September 2017 Nigeria campaign: small to undetectable longterm increase in PVs
- Spring 2017 Iraq campaign: 130% pageviews increase among Asiacell users
Singapore datacenter
The switchover brought significant performance improvements, but no discernible increase in pageviews and...
Singapore datacenter...no discernible increase in unique devices either.
Google referrals

In Q3, we started to systematically store the search traffic data provided by Google for future long-term analysis and prepared to launch a SEO project with consultants. Also, Google’s recently launched beta Search Console allows generating charts 16 months back, enabling direct year-over-year comparisons.

For the desktop domain of English Wikipedia, it confirms that the y-o-y decrease seen during Q3 abated in April:

en.wikipedia.org (Google Search Console beta)
Google referrals

For the mobile domain of English Wikipedia, Google’s remaining data indicates that there had been a temporary drop in Q3 2016/17:

en.m.wikipedia.org (Google Search Console beta)
Google referrals

The slight year-over-year decrease in pageviews referred from Google (observed for November-January in last quarter's meeting) remained for the rest of the quarter, but seems to have abated in April.

In Q3, 44% of web pageviews were directly referred from an external search engine, the vast majority of them (41%) from Google.
Unique devices visiting any Wikipedia project during a given month

NB:
- devices <> users
- Mobile/desktop breakdown only available for per-language uniques

Source: Superset
## Readers (core metrics)

### Appendix: Uniques

<table>
<thead>
<tr>
<th></th>
<th>Monthly unique devices (desktop + mobile web)</th>
<th>Pageviews/device mobile</th>
<th>Pageviews/device desktop</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Wikipedia</td>
<td>804 million</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Spanish Wikipedia</td>
<td>153 million</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>German Wikipedia</td>
<td>104 million</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Russian Wikipedia</td>
<td>95 million</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Japanese Wikipedia</td>
<td>91 million</td>
<td>9</td>
<td>15</td>
</tr>
</tbody>
</table>

Averages for January-March 2018
(Top 5 projects by overall unique devices)
Android app usage

<table>
<thead>
<tr>
<th>Category</th>
<th>Metric</th>
<th>YoY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly pageviews</td>
<td>166.5 million</td>
<td>N/A</td>
</tr>
<tr>
<td>Monthly users</td>
<td>5.08 million</td>
<td>N/A</td>
</tr>
<tr>
<td>Daily users</td>
<td>1.01 million</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Pageviews normalized to 30 days/month
Big updates and fixes to Android in-app analytics may have caused the dip in early March and April but we will be thoroughly investigating the cause.

Q3 avg: 7.55%
Q2 avg: 18.25%
Readers (core metrics)

iOS app usage

<table>
<thead>
<tr>
<th>Monthly pageviews</th>
<th>71 million</th>
</tr>
</thead>
<tbody>
<tr>
<td>YoY: N/A*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Daily downloads</th>
<th>3.6k</th>
</tr>
</thead>
<tbody>
<tr>
<td>YoY: -24.8%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Monthly users</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily users</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Pageviews normalized to 30 days/month
*Includes a small correction for the remaining effects of a pageview-affecting bug that arose in December 2016, which also (together with another bug from Q2016/17) makes year-over-year comparisons unreliable.
7-day retention of Wikipedia iOS app users

Q3 avg: 16.5%
Q2 avg: 16.3%

NB: Data from opt-in users only
new editor retention
Global new editor retention

- Retention rate decreases from 2005 to 2017.
- The retention rate hovers around 7.5% after 2011.

Contributors

May 2018
May 2018

**Contributors**

May 2018

Contributors

English Wikipedia new editor retention by month
Contributors

new active editors

0 5,000 10,000 15,000 20,000

2015 2016 2017 2018

editors per month
### New editor gainers and losers

<table>
<thead>
<tr>
<th>Project</th>
<th>2015–18 change in monthly new active editors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wikidata</td>
<td>↑ 379</td>
</tr>
<tr>
<td>Commons</td>
<td>↑ 105</td>
</tr>
<tr>
<td>Persian Wikipedia</td>
<td>↑ 90</td>
</tr>
<tr>
<td>Tamil Wikipedia</td>
<td>↑ 75</td>
</tr>
<tr>
<td>Estonian Wikipedia</td>
<td>↑ 24</td>
</tr>
<tr>
<td>Hindi Wikipedia</td>
<td>↑ 20</td>
</tr>
<tr>
<td>Marathi Wikipedia</td>
<td>↑ 19</td>
</tr>
<tr>
<td>Bengali Wikipedia</td>
<td>↑ 15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Project</th>
<th>2015–18 change in monthly new active editors</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Wikipedia</td>
<td>↓ 2 417</td>
</tr>
<tr>
<td>Spanish Wikipedia</td>
<td>↓ 446</td>
</tr>
<tr>
<td>Turkish Wikipedia</td>
<td>↓ 235</td>
</tr>
<tr>
<td>German Wikipedia</td>
<td>↓ 199</td>
</tr>
<tr>
<td>Portuguese Wikipedia</td>
<td>↓ 176</td>
</tr>
<tr>
<td>Russian Wikipedia</td>
<td>↓ 162</td>
</tr>
<tr>
<td>Italian Wikipedia</td>
<td>↓ 136</td>
</tr>
<tr>
<td>Chinese Wikipedia</td>
<td>↓ 123</td>
</tr>
</tbody>
</table>
May 2018

Contributors

existing active editors

editors per month

2015 2016 2017 2018
Rank-activity plot of editors

Number of yearly edits vs. User's rank by edits made
May 2018

Contributors

Cumulative edits made by editor activity rank

Percent of edits made by editors at this rank or higher

User's rank by edits made

- 10%
- 90%
- 100%

- 10
- 100
- 1,000
- 10,000
- 100,000
- 1,000,000
May 2018

Contributors

Non-bot, non-reverted edits made by users at different activity frequency
May 2018

Contributors

nonbot edits

edits per month

25 M
20 M
15 M
10 M
5 M
0 M

2015 2016 2017 2018
Contributors

non-bot edits

edits per month

25 M
20 M
15 M
10 M
5 M
0 M

Contributors

total content (main namespace pages + files)
Contributors

Wikidata entities

- 2015
- 2016
- 2017
- 2018

May 2018
Questions?

Monthly core metrics updates: [mw:Wikimedia Audiences]