

Technology Department

Q2 FY20



WIKIMEDIA
FOUNDATION

Platform Evolution

We need to have a modernized, instrumented, and efficient platform to enable next generation engagement across the globe. As we engage, we need to incorporate next generation approaches around artificial intelligence to secure and enhance content, while ensuring the safety for all our curators and readers.

We continue to find success in understanding and assessing issues related to content and integrity and the deployment of artificial intelligence functionalities. We also have had a number of wins associated with protecting and improving the site (e.g. parsing speedups, hardware refreshes) as well as how we work better together (e.g. iOS app, Technical Conference, FAWG - Frontend Architecture Working Group, Anomaly Detection for censorship and outages) across Product and Technology teams.

We continue to be challenged by how best to define and capture the desired impacts we seek to achieve at the MTP level, as it is related to structured data and rich media as well as the annual plan key deliverables related to developer productivity measurement. We are actively working to align our teams and priorities around key infrastructure needs concerning data and our front line defenses.

Drill Down: Platform Evolution

The situation

Last tuning session we committed to coming back with a set of recommendations around our MTP metrics and the Annual Plan “reduce complexity” task today.

We have made progress on the tasks, but need guidance related to structured data and rich media goals.

We are struggling with balancing the outward-facing impact and long term view we know we need with our desire to feel like we have some say in affecting the outcome and what we are working on.

We are also facing challenges in meeting the growth in scale of Wikidata

The impact

While we are still moving ahead on the things we think are necessary to meet MTP goals around structured data and commons, there is a gap in how we measure that impact.

Wikidata and the Wikidata Query Service are reaching the limits of their current architecture.

We continue to see places where the initial goals have long been surpassed due to lack of definition and where Annual Plan reporting is green or yellow, yet the MTP is yellow or possibly red.

MTP Priority slides



WIKIMEDIA
FOUNDATION

Platform Evolution (1)



MTP Outcomes

We will build tooling for internal and external development and reuse of code and content

Key deliverables:

Content Integrity	
Machine Learning (ML) Infrastructure	
Technology and Product Partnerships	

MTP Metrics

Artificial Intelligence (AI) tools and workflows are utilized against 25% of all content

Baseline: All English Wikipedia content uses Artificial Intelligence tooling

25% of content (consumed or created) uses structured data

Baseline: Wikidata: 234 million pageviews with 59 million pages in the month of July

25% increase in rich media content created and consumed across projects

Baseline: Commons: 516 million pageviews; 72 million pages in the month of July

Partnerships (see Advancement deck)

Y1-Goal

10%

Q2-Status

**>25%
DONE**

Page views were not the right metric. **158M page views in Dec. 72m pages**

784M page views, 75.3M pages

Reporting on Q2 metrics for consistency.

In later slides, we cover our newly proposed MTP metrics as related to structured data and rich media and follows the newly refined definitions.



Platform Evolution (2)



MTP Outcomes

A secure and sustainable platform that empowers a thriving developer community with the ease of software -as-a-service tooling

Key deliverables:

Reduce complexity of the platform

MTP Metrics

25% increase in code quality
Baseline: to be determined

Increase in developer satisfaction (DevSat) and 20% decrease in overall outstanding code reviews

Baseline: Developer satisfaction: 3.8 / 5
Baseline: Code Review: 19 days

5% increase in growth and retention of technical communities
Baseline: <N/A>

Y1-Goal

5%

DevSat: 4%
Code Review: 15 days

See definitions (next slide)

Q2-Status

<1%

DevSat: We will release the survey after All Hands (mid-Feb)

CodeReview: 21 days

<N/A>

Reporting on Q2 metrics for consistency.

The newly proposed MTP metrics as related to technical contributions follows the newly refined definitions.



Platform Evolution *(new definitions)*

Structured data is **data that can be used in a programmatic fashion.**

It includes structured data associated to files in Wikimedia Commons and Wikidata entities:

Structured Data on Commons is for machine consumption to enhance media and data discovery.

Structured data in Wikidata is consumed by machines and humans alike.

Non-text content are **images, audio, video, documents** (pdfs), **and data** (e.g. JSON, et. al).

Platform Evolution: Proposal



MTP Metrics

Y1-Goal

Q2-Status

Original: 25% of content (consumed or created) uses structured data
Proposed: A X% increase in structured data used (uptake) across wikis.
For example, X=25% means we would need to use ~9m more items over the course of the MTP.

Help Wanted:
Depends on X

Help Wanted

Baseline (Preliminary Calculations):

Uptake: ~19m entities from Wikidata are currently used across wikis. 56% of pages across the Wikimedia projects make use of Wikidata [1]. It grew by 18% in 2019.
Coverage: 27% of Wikidata entities are used on wikis.

Original: 25% increase in rich media content created and consumed across projects
Proposed: A Y% increase in non-text (e.g. Commons) content used across wikis.

Help Wanted:
Depends on Y

Help Wanted

Baseline (Preliminary Calculations): ~57 million files exist in Commons; 15% growth in file uploads.

Uptake: 14.6m non-text items are used across wikis (12m are images). We don't know the growth rate.

Coverage: 25% commons files are used in wikis

[1] http://wdcm.wmflabs.org/WD_percentUsageDashboard



Platform Evolution *(new definitions)*

Our complex ecosystem of contributors across a wide variety of fields can't be measured by just one metric.

We propose to focus on target groups in both established and emerging communities, and to measure community health through a set of metrics as it relates to sustainability as well as the number of technical contributors.

Initially, we will focus on:

- Growth of independent developers who contribute code patches
- The time that it takes from code contribution to codebase commit
- Growth of tool maintainers
- Developer satisfaction
- Code quality

We expect to visit:

- Number of tools that have co-maintainers
- [Healthy percentage of edits coming from tools and bots](#) in 3 target emerging communities. Target communities and definition of “healthy” are to be determined in Q4
- Target communities have at least 5 technical contributors to share the work



Platform Evolution *(new)*



MTP Outcomes

A secure and sustainable platform that empowers a thriving developer community with the ease of software -as-a-service tooling

Key deliverables:

Improve developer productivity and efficiency to accelerate innovation.

MTP Metrics

X% increase of independent developers who submit patches to production code
Baseline: 228 (July 1, 2018 - Jan 1, 2019)

10% decrease in time from code contribution (patch submission) to codebase commit.
Baseline: 19 days

30% [increase of tool maintainers](#)
Baseline: 1180 maintainers

Increase in developer satisfaction measured and 20% decrease in outstanding code reviews
Baseline: Developer satisfaction: 3.8 / 5

25% increase in code quality
Baseline: 0%

Y1-Goal

Additional time needed to define X

2% / 18 days

5%

4%

5%

Q2-Status

248 devs submitted patches in Q2

21 days

1189

We will release the survey after All Hands (mid-Feb)

<1%



Drill Down: Platform Evolution

The recommendations

- 1) Provide feedback on proposed options for changes
- 2) Extend Technology's MTP/OKR Working Group with representation from the Product Dept
- 3) Prioritize hiring of data team leadership and data product management
- 4) Invest more in shared definitions and collaboration with WMDE to own shared metrics and scaling needs around storage and Wikidata Query Service
- 5) Spend and/or divert more resources to meet scaling requirements in Wikidata and Wikidata Query Service
- 6) A Product Data Analyst works with Analytics Engineering to define, automate and own the calculation and refinement of the structured data and non-text (rich media) calculation.

Key Deliverable slides



WIKIMEDIA
FOUNDATION

Content Integrity



Objective:

Secure and protect the platform and our communities, in the free knowledge movement against the spread of disinformation and bad-actor risk

A portion of the high priority work delivered includes:

- Completion of the 2nd phase of API Token project... tokens allow us to limit damage caused by specific bad actors
- Development of process, policy supporting a repeatable mechanism for delivering security services.
- Expansion of capabilities across 3 out of 4 security services
- We are significantly down the road on securing and implementing additional DDoS protection services

As a result, we've seen...

- An uptick in requests for security services

Target quarter for completion: Q4 FY19/20

Full-time equivalent budget

Actual: 3.27 FTEs



Key Results

Create 2 security *governance* services: risk management and security awareness

Create 2 security *engineering* services: application security and privacy engineering

Baseline: 0

Develop a means to limit and disable the API access of bad actors without interrupting the access of other contributors, integrate it into our platform, and measure the effect
Baseline: 0%

Y1-Goal Q2-Status

Create 2 security <i>governance</i> services: risk management and security awareness	4	3
Create 2 security <i>engineering</i> services: application security and privacy engineering Baseline: 0		
Develop a means to limit and disable the API access of bad actors without interrupting the access of other contributors, integrate it into our platform, and measure the effect Baseline: 0%	100%	33%

Content Integrity



Objective:

Secure and protect the platform and our communities, in the free knowledge movement against the spread of disinformation and bad-actor risk

A portion of the high priority work delivered includes:

- Through a collaboration between Research, Policy, Product, and WMF Chief of Staff (CoS), we now have a prioritized list of projects to address disinformation.
- Through CoS, Policy, and Research, we have built strong relationships with (potential) partners in academia and industry.
- We have identified our first set of Formal Collaborators in the disinformation space and are finalizing the proposal with them.

As a result, we've seen...

- Additional conversation and alignment across the organization with regard to disinformation.

Target quarter for completion: Q4 FY19/20

Full-time equivalent budget

Actual: 3.27 FTEs



Key Results

Organize at least one brainstorming session with internal and external stakeholders to identify and prioritize actions to address disinformation (by the end of Q2)

Baseline: 0

Build 2 sets of Formal Collaborations to expand our capacity for working on prioritized disinformation projects (by the end of Q3)

Baseline: 0

Build a test model to address a specific type of disinformation (by the end of Q4)

Baseline: 0

Y1-Goal Q2-Status

	Y1-Goal	Q2-Status
Organize at least one brainstorming session with internal and external stakeholders to identify and prioritize actions to address disinformation (by the end of Q2)	1	5
Build 2 sets of Formal Collaborations to expand our capacity for working on prioritized disinformation projects (by the end of Q3)	2	In progress
Build a test model to address a specific type of disinformation (by the end of Q4)	1	In progress

Department:
Technology

Content Integrity



Objective:

Expand quality control Artificial Intelligence (AI) tooling to underserved communities, to make fundamental services available for consumption by tools that support the growth of high quality content and the maintenance of quality standards

A portion of the high priority work delivered includes:

- Topic models for routing review of new content to people with the right expertise.
- New advanced counter-vandalism support for under-served wikis (e.g. Norwegian Wikipedia)

Target quarter for completion: End of Q3 FY19/20 for KR's 1 & 3; yet to be determined on the blocked KR

Full-time equivalent budget

Actual: 3.27 FTEs



Key Results

Deploy 8 new quality control AI tools to Wikimedia Projects, increasing availability of AIs for tooling
Baseline: 0

Recruit 4 new campaign coordinators to advertise the availability of the new AI tools, increasing rate of consumption of those tools
Baseline: 0

Improve 3 AI tools in a statistically significant way based on community feedback, ensuring utility of AIs for quality control work
Baseline: 0

Y1-Goal

Q2-Status

Key Results	Y1-Goal	Q2-Status
Deploy 8 new quality control AI tools to Wikimedia Projects, increasing availability of AIs for tooling Baseline: 0	8	6
Recruit 4 new campaign coordinators to advertise the availability of the new AI tools, increasing rate of consumption of those tools Baseline: 0	4	Blocked on coordinatin g resources to engage community
Improve 3 AI tools in a statistically significant way based on community feedback, ensuring utility of AIs for quality control work Baseline: 0	3	7

Department:
Technology

Machine Learning Infrastructure



Objective:

Consolidate and simplify workflows — building, training, and deploying models to enable machine learning aided product augmentation and research.

There hasn't been any work on this project in 2nd quarter due to lack of resources. We do foresee problems ahead due to our staffing constraints in Q3.

Target quarter for completion: Q4 FY19/20 if new ML/Scoring FTEs are hired

Full-time equivalent budget

Actual: 2.10 FTEs

Key Results

Y1-Goal

Q2-Status

Deploy a fully open sourced solution for GPU-enhanced computation infrastructure, improving training times of algorithmic image analysis models by 50% Baseline: 0	100%	100%
Speed up model training by providing models with easier access to feature data through a feature store. Baseline: 0	100%	Q3 work



Department:
Technology

Machine Learning Infrastructure



Objective:

Enable our communities to more easily detect the hidden algorithmic biases in current Machine Learning (ML) solutions

A portion of the high priority work delivered includes:

- Complete set of design assets delivered for MediaWiki integrations
- Jade API deployed to beta cluster
- Consultations with tool developers about providing data to Jade

Target quarter for completion: End of Q3 FY19/20 for deployment of MVP

Full-time equivalent budget

Actual: 2.10 FTEs

Key Results

Build, deploy, and establish baseline metrics for infrastructure that enables Wikipedians to correct the algorithmic predictions around quality of content to 4 wikis.

Baseline: 0%

Increase the rate of community-based false-positive reporting in damage detection models by 100 times

Baseline: 1 report per day

Y1-Goal

4

Q2-Status

50% of the MVP complete

100

0 - Blocked on deployment of the MVP



Platform Evolution



Objective:

Build a reliable, scalable, and comprehensive platform for building services, tools and user facing features that produce and consume event data.

A portion of the high priority work delivered includes:

- Developed a new extension (StreamConfig) so users to configure their data streams without deployment. Example: dynamically changing sampling rates
- We have an EventGate instance just for error logging connected to logstash
- We have a new public endpoint for schemas such developers do not need to modify live schemas to deploy new instrumentation: <https://schema.wikimedia.org/#!/>

Target quarter for completion: Q4 FY19/20

Key Results

5% of analytics and production events migrated to the new event platform. *The percentage of production and analytics events will increase every quarter until older systems can be fully deprecated.*

Baseline: 1%

Client Error Logging is deployed to 1 wiki and error stats are displayed on our operation dashboards.

Baseline: 0

By June 2020, all production and consumption of **new event data** originated in our websites is flowing through this new event platform.

Baseline: 0%

Y1-Goal

5%

Q2-Status

100% of production events and 5% of analytics events migrated

On track. Backend in place, waiting to deploy frontend mediawiki code

100%

On track. Migration of new events starts Q3



Reduce Complexity of the Platform



Objective:

Maintain and evolve developer tooling, testing infrastructure, validation environments, deployment infrastructure, and supporting processes.

Some of the high priority work delivered includes:

- A proposed solution was prepared by Release Engineering team for the Deployment Pipeline compliant CI system in Q2, however, we are also investigating and evaluating other potentially cheaper options before committing.
- 1 service migrated to the Deployment Pipeline and 3 others started.

Target quarter for completion: Q4 FY19/20

Full-time equivalent budget

Actual: 26.53 FTEs

Key Results

Release Engineering and SRE teams create a plan to implement a Deployment Pipeline compliant CI system.
Baseline: 0%

Maintain and improve the Continuous Integration and Testing services
Baseline: Standard Deviation of 12 mins

Developers have a consistent and dependable deployment service.
Baseline: 1 issue per quarter

Reduce infrastructure gaps in the areas of backups, disaster preparedness, observability, infrastructure automation and team structure & support.
Baseline: 0% of planned work done

Y1-Goal

100%

<13.2 mins

Address new reports within 1 month

100%

Q2-Status

60%

4.62 mins

0 incidents this past quarter

85%



Reduce Complexity of the Platform



Objective:

We will improve developer efficiency for all developers: new and experienced, internal and external.

Improving our developers' day-to-day effort is a force-multiplier for all the work they do in supporting the Foundation and the Movement's goals.

A portion of the high priority work to improve productivity includes:

- FAWG & related Architecture gains
- TechConf focused on developer productivity
- WMCS edit data dashboard
- New and improved Kubernetes cluster in Toolforge
- MediaWiki REST API and history APIs for Product

Target quarter for completion: Q4 FY19/20

Full-time equivalent budget

Actual: 26.53 FTEs

Key Results

Determine a baseline set of metrics to assess developer efficiency, including time to first merge (new devs), time to first review (new devs), average time to merge (fully ramped devs), and average time to review (fully ramped devs), by end of Q2.

Baseline: Determining baseline

Improve all baseline developer efficiency metrics by 10% by the end of the year.

Baseline: 3.8/5 or 76%

Improve Cycle Time by 10% year over year.

Baseline: 11.6 days

Y1-Goal

Q2-Status

	4	4
	10% increase	No change, update in February
	10% decrease	11% decrease (10.3 days)

Reporting on Q2 metrics for consistency.

The next 2 slides cover our newly proposed annual plan metrics in regards to reducing the complexity of the platform.



Department:
Technology

Reduce Complexity: Proposal



We propose the following changes to the “Reduce Complexity” Key Deliverable based on our Q1 Tuning session and our analysis of our goals as they relate to the MTP

- A change in title from “Reduce Complexity of the Platform” to “Improve Developer Productivity”
- Keep all existing key results (as seen on the previous three slides) that are designed to track **internal** developer productivity
- Add the following key results to **start** tracking **external** developer engagement and productivity

Key Results

Successfully run Wikimedia’s technical internship and outreach programs — [Google Summer of Code](#) (GSOC), [Google Season of Docs](#) (GSOD), [Outreachy](#) and [Google Code-In](#) (GCI) — that are measured by number of completed projects in GSOC, GSOD, Outreachy and the number of completed task instances in GCI.

Develop, test and evaluate different formats for building technical capacity in emerging communities.

Y1-Goal

20 projects completed, 700 tasks completed, 14 projects promoted for the 2020 rounds

At least 3 formats are developed and tested, year 1 evaluation

Q2-Status

15 projects completed in [GSOC](#), [GSOD](#), [Outreachy](#). [600 task instances completed](#) in GCI

2 formats developed, ongoing testing



Supporting work



P-O15-D2

Wikimedia Technical Conference



Supporting wins: Q2

Sharing and Communication

Bi-Annual [Research report](#) evangelizing our work and building research partner relationships

Create new Data Map ensuring the privacy of our users

WM Tech social media: Driving content on both [@wikimedia-tech](#) and [@mediawiki](#); more tech talks to raise visibility for our technical stories

Documented new architecture for Continuous Integration (CI)

Better Infrastructure

Rebuilt Amsterdam datacenter and reduced our footprint there

New infrastructure for backup and monitoring with a >99% success rate

Implementing a new tool for alert escalation and paging

Large purchase of servers for Dallas datacenter refresh with CPU upgrades and significant cost savings

Improved cloud storage capabilities

Growth and Improved Capabilities

Technology Working Groups

Core Platform Team Clinic Duty to better support other teams

Public release of Media API for communities and GLAMs to better understand the usage of media on Wikimedia Commons

Supporting themes: Q3

Implement Working Group improvements for hiring, onboarding, engineering practices, culture and OKRs

Hire more Engineering Managers

Enable Data Management and Governance

Improve quality of pageview data via better bot identification

Invest in Architecture and Sustaining Engineering

Create a shared understanding of the needs for Machine Learning infrastructure

Launch a Technical blog

Further improvements in alert escalation, paging and incident response

