

Technology Dept Tuning Session Q3 FY21-22



WIKIMEDIA
FOUNDATION

Platform Evolution



Overview

...

Progress and Challenges

New Marseille data center is up and running, eqiad data center expanded by 25%.

Machine Learning's Lift Wing launch into production is delayed, waiting on API Gateway to be finished. ORES still needs maintenance work to keep it from falling over.

The Global Equitable System Performance OKR has been deferred for the remainder of this FY.

Several losses of personnel introduced several delays to work streams.

OKRs

Data as a Service	
Product Platform	

Actions to do in Q4

- Continue the work and make plans to hire backfills to get work back on track



Platform Evolution



MTP Outcomes

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

MTP Metrics

An 80% increase in structured data used (uptake) across wikis.

Baseline: 87M pages across Wikimedia projects use Wikidata as of April 2020.

An 10% increase in non-text (e.g. Commons) content used across wikis.

Baseline: 31.2M items from Commons are used across Wikimedia projects as of April 2020.

**Y3
Goal**

80%
increase

**Q1
Status**

+39.5%
increase in
percentage of
use from 4/20
baseline

**Q2
Status**

+43.5%
increase in
percentage of
use from 4/20
baseline

**Q3
Status**

+47.1%
increase in
percentage of
use from 4/20
baseline

**Q4
Status**



Platform Evolution Metrics



MTP Outcomes

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

MTP Metrics

	Y3 Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
30% increase of tool maintainers Baseline: 1880 maintainers in Q2 (FY 19/20)	25% increase from baseline	17.4% (2207)	19.7% (2251)	21.5% (2291)	
10% (4.2 / 5) increase in developer satisfaction Baseline: 2019 developer satisfaction: 3.8 / 5	4% (4.0)	Next survey will be conducted in Q3	Next survey will be conducted in Q3	Survey responsibility transferring to Platform Product Management <i>(needs to be scheduled)</i>	
10% decrease in code review time Baseline: 19 days in June 2019	7% (17.6 days)	63% (31 days)	142% (46 days)	232% (63 days)	
20% decrease in outstanding code reviews Baseline: 1032 code reviews in June 2021	4% (991 reviews)	796 (-23%)	921 (-11%)	830 (-20%)	



MTP Priority OKRs



WIKIMEDIA
FOUNDATION

Data as a Service



Objective:

Wikimedia application data is easily discoverable and well-prepared to enable data-informed decision making, application development, and research by the community and the Foundation.

Wikimedia's new machine learning infrastructure - called *Lift Wing* - is waiting for connection to the API Gateway to launch its MVP. During that time work has continued, with major additional features and scalability added, in addition to over 100 models migrated from the old system.

Data Catalog evaluation has been completed and DataHub selected as the solution to implement.

We have completed the Discover and Define phases for all 3 use cases and are currently working on executing the defined execution plans for 2 out of the 3 use cases. The third use case for edit-save SLOs is At Risk due to higher priority effort to address User Agent changes.

Target quarter for completion: Q4 FY21-22



Data as a Service



Key Results

KR1 - Establish organizational data management structure. Build a browseable, shareable data dictionary, and describe 25% of known data elements.

Baseline: Existing informal documentation of misc datasets

KR2 - Enable efficient program evaluation and decision support through three novel use cases.

Baseline: Siloed analytics by team

KR3 - Build machine learning services. Operationalize a ML governance strategy for the Foundation, and create ways to understand, evaluate, and provide feedback on ML models.

Baseline: No MVP, No ML governance strategy in place

	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
	Data Discoverability	Resourcing, info gathering, planning On Track	Data Catalog evaluation in progress On Track	Data Catalog evaluation complete On Track	
	3 Integrated Analytics towards Self-Service	0 out of 3 On Track	0 out of 3 On Track	0 out of 3 2 On Track 1 At Risk	
	Lift Wing MVP ML Governance Strategy	MVP Infra: ~70% MVP Models: ~80% MVP API: ~10% Governance: ~5%	MVP Infra: ~90% MVP Models: ~90% MVP API: ~10% Governance: ~10%	MVP Infra: 100% MVP Models: 100% MVP API: ~10% Governance: ~25%	



Product Platform



Objective:

Our platform and processes are ready and able to invite all of the world's population to join us.

The Structured Data team completed a new resource plan for completing the grant requirements. We also completed a draft update that will be delivered to the Sloan Foundation requesting a removal of the requirement that images be added to 5 million pages.

Challenges:

The Search team has not has enough resources to do the work necessary to understand who uses search and how; we hope to spend more time on that next quarter. Additionally, there are not many Blazegraph alternative options, and we're going to have to make compromises moving forward. Lastly, trying to maintain a 99% update lag <10min SLO is very difficult and doesn't allow us enough room for other work -- but tracking this over the last 6+ months allows us to better adjust it.

The Search team is stretched thin between Search & Relevancy, and Query Services, leading to missed deadlines (WDQS Streaming Updater, WCQS production) and inability to address important work (Query Completion). WDQS Blazegraph migration KR4s have been revised from migrating off of Blazegraph in FY21-22 to identifying an alternative graph backend, and sketching out a technical migration plan.

KR4 target is shifted to Parsoid used for read views in Discussion Tools for at least some pages on at least one production wiki. The milestones set at the beginning of the year were risky and ambitious and have been downscoped for this fiscal year.



Product Platform



Key Results

	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
<p>KR1: Machines are able to recognize Wikimedia content and suggest relations to other Wikimedia content in at least 2 wikis, enabling experimentation with at least two new strategic features</p> <p>Baseline: 0 experimental features</p>	2 new experimental features released in 2 wikis	Requirements for 1 feature developed	Development in process for 1 feature; requirements for 2nd in process	1 feature making good development progress; draft PRD for two others completed	
<p>KR2: Scale our capability to surface and utilize relationships between content across two different use cases and the current WDQS service level continues to support editing of Wikidata</p> <p>Baseline: 25 10+ min server events in the past 30 days (11 Sept 2021 to 11 Oct 2021; manual count on this graph)</p>	WDQS & WCQS update lag < 10min 99% of the time	Updates to improve lag in progress	WDQS at 99.248% as of 19JAN22; WCQS Beta 2 was shipped on 01Feb2022, into production use	WDQS averaged 95.1% between Jan-Mar 2022.	
<p>KR3: Improve the satisfaction and velocity of staff and volunteer developer contributions to frontend components of Wikimedia projects by 10% through migration to the Vue frontend framework</p> <p>Baseline: 0%</p>	10% increase in satisfaction and velocity	0% increase	0% increase	0% increase	
<p>KR4: Reduce platform complexity by fully integrating Parsoid into the rendering pipeline so that by EOFY (a) all talk pages are rendered with Parsoid enabling advanced DiscussionTools features (b) over 95% of Main namespace pages on all non-language-variant wikis have insignificant visual diffs compared to current output (c) Parsoid HTML is used for read views on the main namespace on at least 1 wiki.</p> <p>Baseline: TBD</p>	100% talk pages rendered 95% Main namespace pages	Parsoid renders at least ~90% of pages from 20 wikis without visual diffs	Parsoid renders at least ~90% of pages from 20 wikis without visual diffs	Parsoid renders at least ~90% of pages from 20 wikis without visual diffs	

Efficacy & Resilience OKRs



WIKIMEDIA
FOUNDATION

Impact Culture: Business Intelligence

Objective:

Foundation leadership have access to organizational data, insights and dashboards for impactful decision making.

-
- MTP Insights Memo and corresponding dashboard delivered to leadership for reference during APP planning.
 - Consulted with Tech, Product, Comms, and Legal to compile and documented comprehensive list of past, current, and future recommended strategic metrics for consideration during FY 22-23 planning.
 - Theory of Change and SROI metric KR were on pause as leadership team worked through the “reimagining planning” exercise in Q3. We expect some of this work to take a new shape in FY 22-23 in line with the Quarterly Learning Sessions.

Target quarter for completion: Q4 FY 21-22



Impact Culture: Business Intelligence

Key Results

	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
<p>KR1: Each quarter, 50% of departments work with the GDI team on the theory of change; develop org-level business metrics, including Social Return on Investment (SROI) measures by Q2</p> <p>Baseline: No ToC or SROI measures established.</p>	<p>50% of departments work on Theory of Change</p> <p>SROI metrics defined</p>	On hold	On hold for reimagining planning	On hold for reimagining planning	
<p>KR2: By end of Jan 2022, design and roll out a data dashboard that can be used by at least 50% of Foundation departments for annual planning.</p> <p>Baseline: No dashboard.</p>	<p>Dashboard launched by Jan. 2022</p> <p>50% of departments use dashboard for annual planning</p>	60% of required metrics identified	Dashboard 90% complete and on track for end of Jan. 2022	<p>MTP Insights Memo and corresponding dashboard delivered to leadership for reference during APP planning.</p> <p>Used as a reference during <i>Metrics that Matter</i> cross-departmental conversions during NYC APP sessions.</p>	



Technical Community Building



Objective:

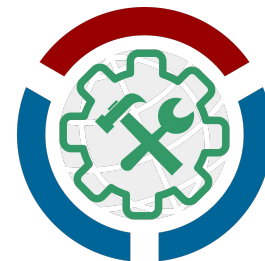
Our technical community is thriving and has a clear, consistent means to discover, build, and deploy applications that support community workflows, invent new forms of content creation and consumption, and leverage Wikimedia's APIs and data beyond the core wiki experience.

Discoverable Docs (KR1): We made good progress on the implementation of the Developer Portal and developed a UX research study with the Design Research team. Reduced team capacity and higher than anticipated coordination efforts with doc owners slowed down the docs review progress. Documentation updates will continue in Q4.

Curated Pathways (KR2): We kicked off the Small Wiki Toolkits global [workshop series with focus on script and bot development](#) as well as the main planning phase of the 2022 *hybrid* Hackathon. Newcomers from a broad range of regions are expected to participate this year.

Toolhub (KR3): Accomplishments in Q3 include new features for creating and viewing lists, a recent change feed to support patrolling, and backend features for annotations that users can add to toolinfo records.

Target quarter for completion: Q4 FY 21/22



[Toolhub](#)



[Small Wiki Toolkits project logo](#) by Sethiasta, cc-by sa 4.0



Technical Community Building



Key Results

KR1: A single entry point into our technical documentation, combined with accurate landing pages, enables technologists of all levels to discover Wikimedia's technical areas and engage in communities of practice (acceptance criteria for accuracy defined by Q1, 100% of landing pages meet criteria by end of Q3, launch of developer portal by Q4)

Baseline: Content draft and research was done in FY 20/21

KR2: At least 10 underserved wiki communities have a set of technical skills and capabilities enabled through globally localized partnerships and curated pathways for learning and contribution, in support of scaling the ecosystem of tool developers (10%)

Baseline: 2162 Tool maintainers; communities: 7

KR3: The Community is empowered to improve data coverage in Toolhub and can assess the reliability of tools for adoption based on a set of quality signals (research by Q3, code and community outreach by Q4).

Baseline: 0

	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
	Developer portal is launched 100% of landing pages meet accuracy criteria	Q1 goals completed	Q2 goals completed	landing page updates continue in Q4	
	10% increase in tool maintainers 10 communities	2.4% (2207) increase in tool maintainers	4.7% (2251) Increase in tool maintainers	6.5% (2291) Increase in tool maintainers	
	Release of Toolhub 1.1; adoption	On track	Proposed update to KR to reflect changes	On track	



Platform Excellence: Resilience



Objective:

Our services, infrastructure and data are resilient to and/or quick to recover from unexpected malicious or non-malicious events

In Q3, we hit **two major milestones** in the Resilience objective, by seeing two separate projects for this year get across the finish line (**KR1**). First, we launched our new **Marseille data center**, a huge accomplishment for a year-long project that was delayed by several unexpected challenges, obstacles and [curve balls](#)! As part of this launch, we performed a **failover test** of all EMEA traffic at **full load**, temporarily steering away traffic from our older Amsterdam site, verifying that the new site is able to handle **all EMEA traffic** by itself when necessary, and thus successfully meeting our Resilience objective.

In this quarter, we *also* saw the launch of our other data center project: the **expansion** of our **Ashburn data center**, one of our two “main” sites. This constitutes a **25% growth** of available (installed) space, the biggest since the launch of the site in **2011**, securing the necessary room for our current and future levels of infrastructure growth.

For **KR3**, in Q3 we hired a new Project Manager for the Security team, and deployed new security services: a new **security awareness** service, as well as a new **pen testing** service.

Target quarter for completion: Q4 FY21/22



Platform Excellence: Resilience



Key Results

KR1: Wikimedia's infrastructure is scaled to address known compute, storage and traffic capacity risks, by adding a new data center in EMEA (by end of Q2), expanding our main data center by at least 20% (by end of Q2), and by documenting two new capacity plans (by end of Q4)

Baseline: One EMEA data center; 32 [racks](#) of space in our main data center; zero documented capacity plans

KR2: Service and security operational issues are detected, escalated, remediated and communicated to stakeholders and the movement, as measured by a 20% incident score improvement

Baseline: 6.5/15 (*based on Q2's incidents*)

KR3: Security and privacy services are enterprise wide, centrally coordinated, scalable and resilient in a way that empowers all users to make good security and privacy decisions, measured by a 10% increase in consumption of consultation services and a 30% decrease in operational services

Baseline: Defined on a per service basis

	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
	One new data center in EMEA Expand main data center by >20% Two new capacity plans	Contracts signed and orders placed Awaiting equipment	Marseille site up, not serving user traffic yet 48 racks (+50%) of committed data center space	Marseille up, serving user traffic (goal met) 40 racks (+25%) of available data center space (goal met)	
	20% incident engagement improvement	Scorecard defined	Scorecard published	Baseline established Q3 score: 6.8 (+4.6% ↑)	
	10% increase in consultation services 30% decrease in operational services	Defined on a per service basis	Most services are level trending or trending toward & more consultative with the exception of Appsec & Vendor Security	All services are trending toward operational vs consultancy.	

Culture, Equity & Team Practices



Objective:

Culture, Equity and Team Practices: Modern and Inclusive Practices. Modern software practices and inclusive communication practices will remove or significantly reduce barriers to entry and collaboration, and streamline deployments for all technical contributors in the wiki ecosystem, enabling a faster path to innovation.

We co-developed the Technology department's Communication Guidelines - Draft 1 and we gathered feedback for it from the [ERC team representatives](#) in Technology, Managers and Senior Managers in the department, Tech's Leadership (the Directors and our VP), as well as Talent & Culture.

Target quarter for completion: Q4 FY21/22



Culture, Equity & Team Practices



Key Results

	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
<p>KR1: At least one user-journey, and all its supporting infrastructure services, are covered by SLOs, and those SLOs are used to drive engineering decisions including when to roll back deployments. For all supporting services within this slice developed at the Foundation, including MediaWiki, we measure and track lead-time as an indicator of the objective of maximizing the speed of innovation.</p> <p>Baseline: 0 user-journey SLOs, 3 infrastructure service SLOs</p>	<p>1 user journey</p> <p>N infra services (with SLOs defined)</p>	<p>1 user journey selected</p> <p>4 infra svc SLOs defined</p>	<p>1 user journey selected</p> <p>5 infra svc SLOs defined</p>	<p>1 user journey selected</p> <p>5 infra svc SLOs defined</p>	
<p>KR2: 100% of non-inclusive language is removed from our documentation and code repositories according to https://www.mediawiki.org/wiki/Inclusive_language_list.</p> <p>Baseline: 0</p>	<p>100% of non-inclusive language removed</p>	<p>0%</p>	<p>0%</p>	<p>0%</p>	
<p>KR3: Deliver 3 milestones towards one recommendation resulting from the Effective and Responsible Communication discovery phase (by Q4)</p> <p>Baseline: 0</p>	<p>3</p>	<p>0</p>	<p>1</p>	<p>2</p>	
<p>KR4: 100% of development teams across Technology and Product departments build using the DEI framework.</p> <p>Baseline: 0%</p>	<p>100% of teams</p>	<p>0%</p>	<p>7/10 teams piloting</p>	<p>7/10 teams piloting</p>	

Service Level Objectives

Reporting period:
Dec 1, 2021 – Feb 28, 2022

Service	SLO	
etcd	Availability: 99.9% of requests are successful	✓
	Latency: 99.8% of requests under 32 ms	✓
Docker Registry	Availability: 99% of requests are successful	✓
	Manifest read latency: 95% of manifest reads under 2000 ms	✓
	Tag read latency: 95% of tag reads under 2000 ms	✓
	Write latency: 95% of manifest writes under 3000 ms	✓

Service	SLO	
API Gateway	Availability: 99.9% of requests proxied successfully	✓
	Latency: <i>SLO temporarily withdrawn; see speaker notes</i>	
Varnish	Latency–availability: 99.9% of requests proxied successfully and with less than 100 ms added, in each data center (DC)	✓
Logstash	Availability: 99.5% of events indexed successfully in each DC	✓
	Latency: 99.5% of events ingested without consumer lag in each DC	✓

✓ **All services met their SLOs this quarter. See the speaker notes for numeric values.**

Appendix







WIKIMEDIA
FOUNDATION

Color Coding Legend

The color coding is intended to reflect the status of the key results (KRs). Completing KRs, as described in the [Annual Plan](#), is what advances an overall OKR status. The average progress on all KRs aligned to an objective is what dictates the current status.

If work is not scheduled for this quarter, the key result may be greyed out.

 COMPLETE	<ul style="list-style-type: none">All key results have been fully achieved and work is complete
 ON TRACK	<ul style="list-style-type: none">Work is underway and the key results are on track for scheduled completion with no concerns
 CONCERNS	<ul style="list-style-type: none">Work is underway but there are concerns which may impact the intended key results, further discussion required
 AT RISK	<ul style="list-style-type: none">Work has not started or has ceased and there are concerns which will impact the intended key results, further discussion required