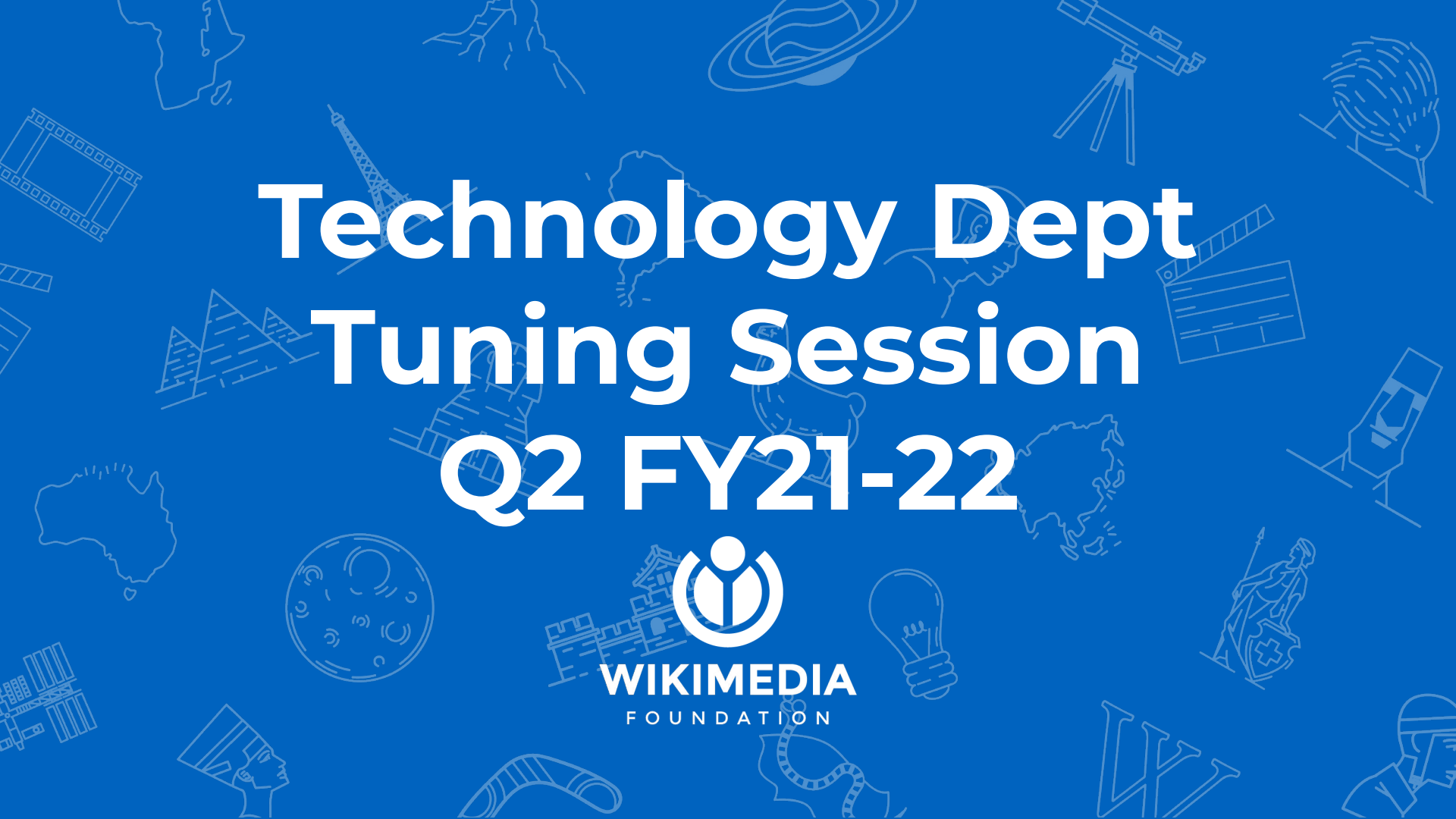
Technology Dept Tuning Session Q2 FY21-22



Platform Evolution

Overview

Created and published Execution Plans to understand and document our methodologies and resourcing for achieving our top level OKRs in the annual plan. Our new machine learning infrastructure achieved it's MVP and will soon be connected to the API Gateway.

Progress and Challenges

Structured data and non-text usage across wikis continue to increase. The requirements phase is complete for the new organizational data management structure and the Data Catalog evaluation is in progress.

Structured Data team has continued to see significant staff turnover, slowing velocity and putting the Product Platform's KR1 at risk. WDQS Blazegraph migration has stalled as a new methodology is found. WCQS Beta 2 was nearly completed in Q2, was released to production in early Q3.

We do not expect to work on the Global Equitable System Performance OKR for the remainder of this FY.

OKRs

Data as a Service	
Platform Excellence: Global equitable system performance	
Product Platform	

Actions to do in Q3

- Lift Wing (our machine learning model serving platform) will be ready to launch

Department: Technology

Platform Evolution

MTP Outcomes

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

MTP Metrics	Y3 Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
An 80% increase in structured data used (uptake) across wikis. **Baseline:** 87M pages across Wikimedia projects use Wikidata as of April 2020.	80% increase	+39.5% increase in percentage of use from 4/20 baseline	+43.5% increase in percentage of use from 4/20 baseline		
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.	10% increase	+14.5% increase in percentage of use from 4/20 baseline	+18.1% increase in percentage of use from 4/20 baseline		

Department: Technology

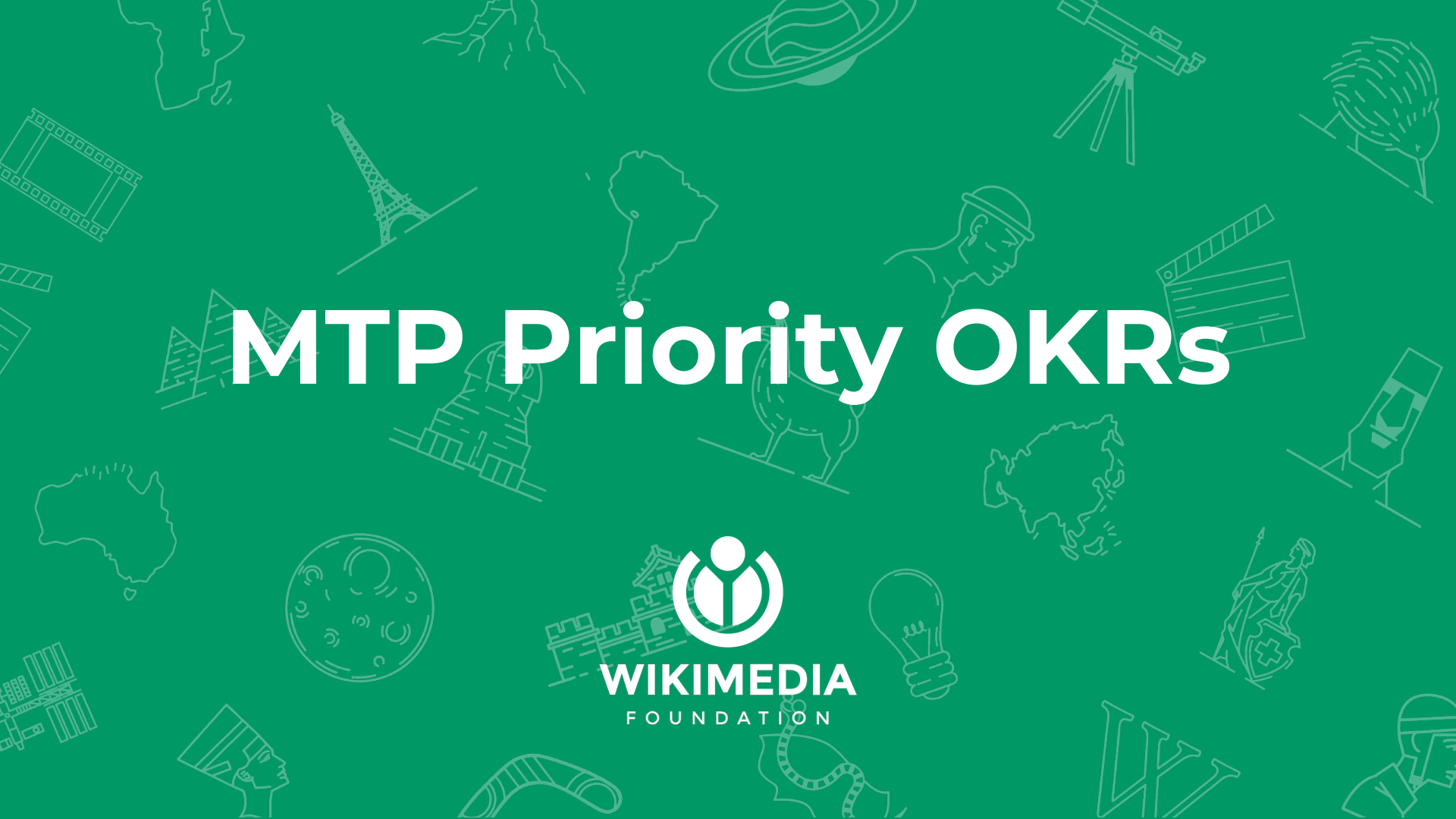
Platform Evolution Metrics

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)		
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		
10% decrease in code review time **Baseline:** 19 days in June 2019	7% (17.6 days)	63% (31 days)	142% (46 days)		
20% decrease in outstanding code reviews **Baseline:** 1032 code reviews in June 2021	4% (991 reviews)	796 (-23%)	921 (-11%)		

MTP Priority OKRs



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* - has reached its MVP and is about to launch. The infrastructure is already hosting and serving real, production models. The final step to launch is working to connect it to the soon-to-be-launched API Gateway.

Data Management requirements are complete and the Data Catalog evaluation is in progress for a number of shortlisted Open Source solutions.

We have completed the Discover and Define phases for all 3 use cases. We have shared outcomes and achieved buy-in on next steps with stakeholders. In Q3 all 3 use cases will be in active development.

Target quarter for completion: Q4 FY21-22

Department: Technology

Data as a Service

Key Results	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
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	Data Discoverability	Resourcing, Info gathering, Planning On Track	Data Catalog Evaluation in progress On Track		
KR2 - Enable efficient program evaluation and decision support through three novel use cases. **Baseline**: Siloed analytics by team	3 Integrated Analytics towards Self-Service	0 out of 3 On Track	0 out of 3 On Track		
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	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%		

Department: Technology

Global equitable system performance ↘

Objective:

People in emerging communities have the same level performance for their contribution workflows as those in established communities.

Due to gaps in staffing in the Technology Department, we do not expect to work on this OKR for the remainder of this FY.

Department: Technology

Global equitable system performance ↘

Key Results	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
KR1 - Plan and resource 3 projects based on our strategy for increasing performance for contributors globally, specifically addressing caching and content composition for logged in users. **Baseline:** 0 projects planned and resourced	3 projects	0%			
KR2 - Complete 3 performance projects which decrease the time it takes to complete existing contr markets (Based on Largest Cont perceived load speed) **Baseline:** 0 projects which resul within a laboratory regional tes					
KR3 - Make 3 product decisions made by using performance data gathered from testing against standard emerging community performance personas. **Baseline:** 3 product decisions made	3 product decisions	0%			

Due to gaps in staffing in the Technology Department, we do not expect to work on this OKR for the remainder of this FY

Department: Technology

Product Platform

Objective:

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

Foundational work has begun to enable KR1 and KR3 including new technology to support necessary data platforms (KR1) and initial development of a shared component library to support the move to Vue (KR3).

Parser consolidation work continues with active work to address rendering, performance, and functionality gaps.

Challenges:
The Structured Data team has continued to see significant staff turnover, slowing velocity and putting KR 1 at risk for the year. We are working to create a new resource plan and timeline to meet the requirements of the SDAW grant. Risk has been escalated to CPO team and CRO.

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 KRs have been revised from migrating off of Blazegraph in FY21-22 to identifying an alternative graph backend, and sketching out a technical migration plan.

Target quarter for completion: Q4 FY21-22

Departments: Product & Technology

Product Platform

Key Results	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
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 **Baseline:** 0 experimental features	2 new experimental features released in 2 wikis	Requirements for 1 feature developed	Development in process for 1 feature; requirements for 2nd in process		
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 **Baseline:** 25 10+ min server events in the past 30 days (11 Sept 2021 to 11 Oct 2021; manual count on this graph)	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		
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 **Baseline:** 0%	10% increase in satisfaction and velocity	0% increase	0% increase		
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. **Baseline**: TBD	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		

Efficacy & Resilience OKRs

Impact Culture: Business Intelligence

Objective:

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

- Developed core elements of MTP Insights “dashboard” to measure progress on key metrics since 2019.
- 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.
- Initial dashboard on track to be shared with WMF leadership by end of Jan. 2022.
- Theory of Change and SROI metric KR is on pause as CEO Maryana leads WMF through reimagining annual planning, which will have significant implications on how we might prioritize and carry out this work moving forward.

Target quarter for completion: Q4 FY 21-22

Department: Technology

Impact Culture: Business Intelligence

Key Results	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
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 **Baseline**: No ToC or SROI measures established.	50% of departments work on Theory of Change SROI metrics defined	On hold until January	On pause for reimagining planning		
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. **Baseline**: No dashboard.	Dashboard launched by Jan. 2022 50% of departments use dashboard for annual planning	60% of required metrics identified	Dashboard 90% complete and on track for end of Jan. 2022		

Department: Technology

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 technical implementation of the Developer Portal. Work on updating key technical documents has begun. Reduced technical writer capacity from Q3 on requires us to distribute work differently within the Developer Advocacy team. Impacts to the timeline are not expected.

Curated Pathways (KR2): Q2 focus of the engineering work has been on updating PAWS, and to continue improving Toolforge workflows. We also designed a new workshop concept to scale and streamline the Small Wiki Toolkits initiative. The workshops series is organized in collaboration with members of multiple smaller wiki communities and will take place in Q3/4.

Toolhub (KR3): Toolhub successfully launched in October. However, the plan for the FY 21/22 deliverables depended on Product Management support which did not materialize. We propose to adjust the focus of this KR to reflect the new timeline, and to incorporate insights gained since Toolhub's launch. More information follows in the drill down slide.

Target quarter for completion: Q4 FY 21/22

Toolhub

Small Wiki Toolkits project logo by Sethiasta, cc-by sa 4.0

Department: Technology

Technical Community Building

Key Results	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
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	Developer portal is launched 100% of landing pages meet accuracy criteria	Q1 goals completed	Q2 goals completed		
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	10% increase in tool maintainers 10 communities	2.4% (2207) increase in tool maintainers	4.7% (2251) Increase in tool maintainers		
KR3: Users can assess the reliability of tools for adoption, contribution or research based on a system of quality signals (co-maintainers, docs, recent editing usage, published source code, endorsements, etc) within Toolhub (research by Q2, code by Q3, adoption by Q4). **Baseline:** 0	Release of Toollhub 1.1; adoption	On track	Proposed update to KR to reflect changes		

Department: Technology

Drill Down: Technical Community Building

The situation

Old wording for KR3:

Users can assess the reliability of tools for adoption, contribution or research based on a system of quality signals (co-maintainers, docs, recent editing usage, published source code, endorsements, etc) within Toolhub (research by Q2, code by Q3, adoption by Q4).

New wording for 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).

The impact

Revised timeline:

The original plan for Toolhub depended on Product Management support which did not materialize. As an effect, research phase and prioritization will extend into Q3, and the implementation of Toolhub 1.1 to the end of Q4.

Refined focus:

Insights gained at the launch of Toolhub in October informed the need to set a stronger emphasis on workflows that empower the community to improve both data quality and quantity. Conversations with users confirmed that recognizing “data signals” for reliability is yet to be learned, as users tend to use tools that others recommend - independent of indicators like multiple maintainers, or published source code.

The recommendation

We propose to prioritize features that enable the community to improve data coverage in Toolhub, and focus on quality signals that reference best open source practices. Community outreach activities in Q4 will focus on data-editing, and providing information on how to recognize signals for reliability.

The new wording of KR3 reflects both the revised timeline, and the refined focus.

Platform Excellence: Resilience

Objective:

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

Fluctuating lead times directly related to the global supply chain issues and chip shortages have further delayed our two data center projects (**KR1**) in Q2. We were able to **mitigate** some of the effects by temporarily repurposing existing hardware and putting additional pressure on our vendors. We have **revised our timeline** to start serving traffic out of the new Marseille (EMEA region) data center by the **end of Q3** (previously was end of Q2) and go live with the expanded Ashburn data center by **mid-Q3** (also from end of Q2).

For our **incident response** initiative (**KR2**), we ***remain on-track***. In Q2, we finalized and communicated our **incident scorecard** to the broader SRE team, to help establish metrics for assessing our progress. Our **incident retrospective process** was **revamped**, establishing new templates, format and rituals. We also completed the **rollout of the new status page** to **better communicate** our operational status to stakeholders and the movement at large. **In Q3**, we are retroactively scoring Q2's incidents to establish our baseline, and experimenting with incident process improvements, measuring our score improvement as we go.

For our security services initiative (**KR3**), in Q2 the security team collected the data we needed to **further refine** our service portfolio, set expectations for the deployment of **checklists**, defined where automation can occur and where can we start using **risk assessments** to accelerate our service delivery model.

Target quarter for completion: Q4 FY21/22

Department: Technology

Platform Excellence: Resilience

Key Results	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
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	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 main data center space		
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:** To be defined ~~in Q1/Q2~~ in Q3 with a Q2 baseline	20% incident engagement improvement	Scorecard defined	Scorecard Published		
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:** To be defined in Q1	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		

Department: Technology

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 expanded the Effective and Responsible Communication team to 22 individuals. **We developed a Core and Extended team model** for the group of 22 individuals to be able to work with one another efficiently and effectively. We defined roles and responsibilities for each group. We kickstarted the first work session of the expanded ERC team. While we met the quarterly goal for this project, we faced reduced capacity for change in the department during Q2.

Target quarter for completion: Q4 FY21/22

Culture, Equity & Team Practices

Key Results	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
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. **Baseline:** 0 user-journey SLOs, 3 infrastructure service SLOs	1 user journey N infra services (with SLOs defined)	1 user journey selected 4 infra svc SLOs defined	1 user journey selected 5 infra svc SLOs defined		
KR2: 100% of non-inclusive language is removed from our documentation and code repositories according to https://www.mediawiki.org/wiki/Inclusive_language list. **Baseline:** 0	100% of non-inclusive language removed	0%	0%		
KR3: Deliver 3 milestones towards one recommendation resulting from the Effective and Responsible Communication discovery phase (by Q4) **Baseline**: 0	3	0	1		
KR4: 100% of development teams across Technology and Product departments build using the DEI framework. **Baseline**: 0%	100% of teams	0%	7/10 teams piloting		

Department: Technology

Service Level Objectives

Reporting period:
Sep 1 – Nov 30, 2021

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: *SLO temporarily withdrawn; see speaker notes*	
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**	• All key results have been fully achieved and work is complete
 **ON TRACK**	• Work is underway and the key results are on track for scheduled completion with no concerns
 **CONCERNS**	• Work is underway but there are concerns which may impact the intended key results, further discussion required
 **AT RISK**	• Work has not started or has ceased and there are concerns which will impact the intended key results, further discussion required