Technology Tuning Session Q3 FY20-21

Project Expedition: Refactoring the core of MediaWiki for easier development, better maintenance and more confidence

- *“Break up the big ball of mud”*
- Great progress in < 3 months
  - ~380 patches, 5 new value object types
  - Removed 300+ usages of old code
  - Lots of learning about how to untangle ~10+ (20?) years of code
  - Unlock future architectural changes

Platform Evolution

Overview

As we continue to expand and modernize our platform we are achieving steady results. However, there are areas of concern that we need to address in the coming quarter, particularly code review time, which has continued to backslide. Despite this, we are confident that the work on platform evolution is seeing tangible results and continues to add value.

Progress and Challenges

Progress: Machine Learning (ML) Platform server cluster is running and currently being configured in preparation for a MVP launch of Lift Wing. In addition, a new ML model governance strategy will formalize WMF roles and responsibilities around ML models we create and/or host.

Progress: Architecture released the first Proof of Value/Artifact focused on Structured Content.

Challenge: Code reviews continue to take longer.

OKRs

Content Integrity	
Evolutionary Architecture	

Actions

- Hiring of the Director of Anti-disinformation has been postponed. Anti-disinformation managers in Trust and Safety are being hired.

Department: Technology

Platform Evolution

MTP Outcomes

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

MTP Metrics	Y2 Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
A 80% increase in structured data used (uptake) across wikis. **Baseline:** 87M of pages across Wikimedia projects use Wikidata as of April 2020.	80% increase	16% increase in percentage of use from 4/20 baseline	26% increase from baseline	31% increase from 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	0.7% increase in percentage of use from 4/20 baseline	1.8% decrease from baseline	3% increase from baseline	-

Department: Technology

Platform Evolution Metrics

MTP Outcomes

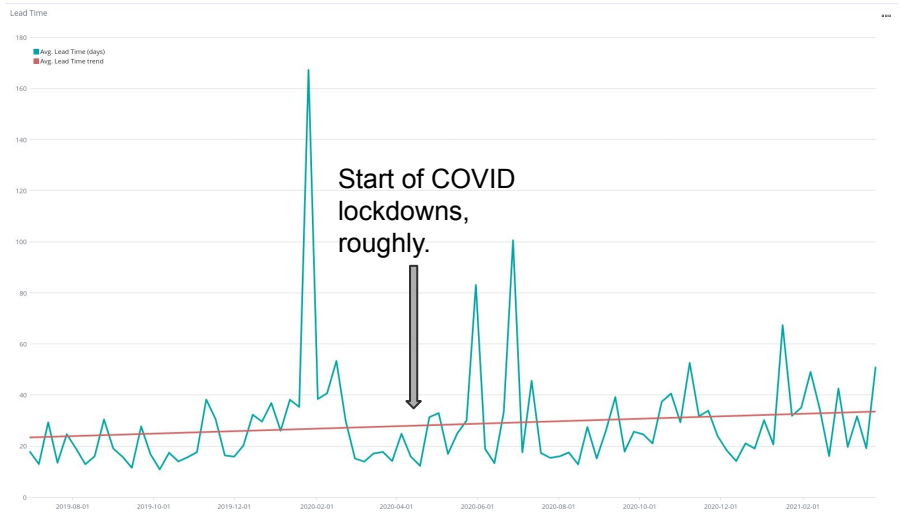
A secure and sustainable platform that empowers a thriving developer community with the ease of software -as-a-service tooling

MTP Metrics	Y2 Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
30% increase of tool maintainers **Baseline:** 1880 maintainers in Q2 (FY 19/20)	15% increase from baseline	8% (2033) increase from baseline	9.9% (2067) increase from baseline	12.5% (2122) increase from baseline	
10% (4.2 / 5) increase in developer satisfaction **Baseline:** 2019 developer satisfaction: 3.8 / 5	4% (4.0)	Next survey will be conducted in early Q3	Next survey will be conducted in early Q3	-13% (3.3)	
10% decrease in code review time **Baseline:** 19 days in June 2019	4% (18 days)	-5% (20 days)	-21% (22 days)	-79% (34 days)	
20% decrease in outstanding code reviews **Baseline:** 1088 code reviews in June 2019	4% (1043 reviews)	15% (918 reviews)	21% (851 reviews)	13% (940 reviews)	

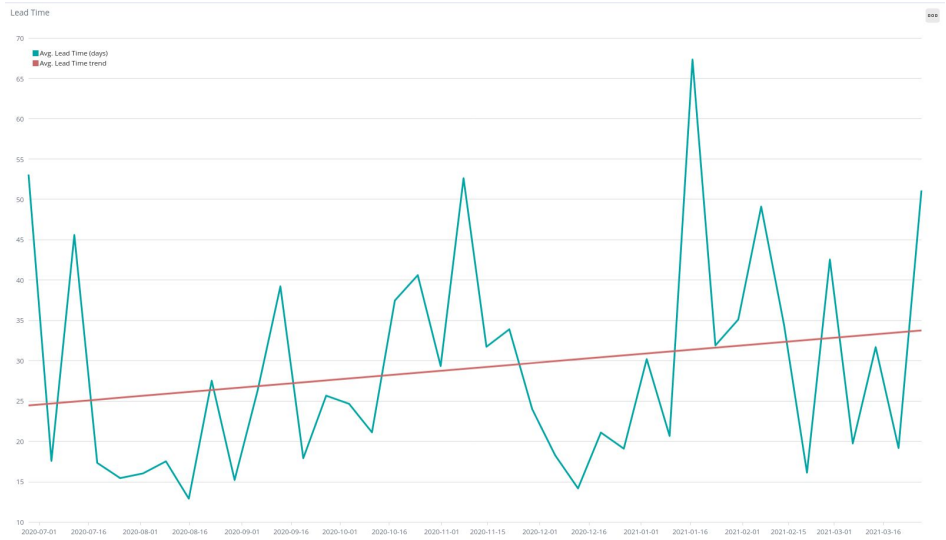
Department: Technology

Drill Down: Platform Evolution

Code review time (time to merge)



July 2019 - March 2021



July 2020 - March 2021

OKR slides

WIKIMEDIA
FOUNDATION

Content Integrity

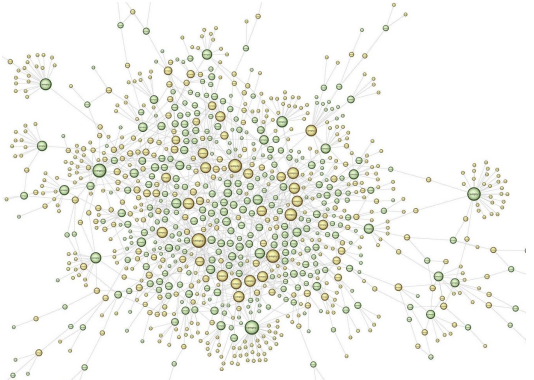
Objective:

The Wikimedia movement has the tools and knowledge to identify and respond to abuse, misinformation, and disinformation campaigns, in order to more efficiently and effectively improve the quality of content, defend the projects and retain the public's trust.

Created models to detect inconsistencies between Wikidata and Wikipedia and understand the diffusion of content across Wikipedia communities.

Designed the infrastructure for serving Machine Learning models to detect Abuse, misinformation, and disinformation.

Started the onboarding of the Research Scientist - Disinformation.
Paused the hiring for the Anti-Disinformation Director position.
Added a T&S Disinformation team with a focus on strategic high risk languages.



The OrangeMoody network of socks (ongoing)

Target quarter for completion: Q4 FY20-21

Department: Technology

Content Integrity

Key Results	Year Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
KR1: Develop 1 model to identify misinformation (by Q3) and deliver 4 milestones towards 1 model to understand the diffusion of content on the Wikimedia projects (by Q4). **Baseline:** 0 milestones	4	1	2	3	-
KR2: Complete 4 product integrations (e.g. link recommendations) of internally built machine learning models and services for improving content integrity **Baseline:** 0 models in the new ML platform	4	0	0	0	-
KR3: Coordinate and report quarterly on the disinformation activities across teams, establish a community support forum (by Q3), and participate in 10 external relationships (by Q4) including academia and industry partners. **Baseline:** 0	10	5	5	6	-

Department: Technology

Evolutionary Architecture

Objective:

A new Wikimedia Knowledge Platform is defined, building upon our key technologies, enabling our development teams to deliver value independently, and empowering our communities to share the world's knowledge in the spaces and formats which reflect their values and cultures.

Designing a modern mission aligned system

- Released proof of value exploring structuring Wikimedia content. (demo)(artifact)

Developing an Architecture Practice:

- Initial release of the Architecture Repository
- Hosted book club to explore/discuss modern event driven systems. Following up this in Q4 by hosting an unconference to explore this topic in the Wikimedia context

Measuring the impact of architecture:

- Analyzed results of architecture understanding survey

Target quarter for completion: Q4 FY20-21

Department: Technology

Evolutionary Architecture

Key Results	Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
KR1: Reached level 3 of the Architecture Maturity Model (AMM) for at least 85% of capabilities by Q4. **Baseline:** 4 capabilities at Level 1 (50%), 4 capabilities at Level 2 (50%)	move 85% of capabilities to Level 3	37.5% at Level 1 62.5% at Level 2	100% at Level 2 28.5% at Level 3	100% at Level 2 12.5 % at Level 3 12.5 % at Level 4	-
KR2: Engineers have an improved understanding of current architecture and target architecture, enabled by the outputs of reaching level 3 for the AMM Architecture Development capability, “Gap analysis, modeling and iterative strategy for reaching a target architecture are completed”, measured quantitatively through surveys in Q1, Q2 and Q4. **Baseline:** 69% of respondents had some level of architecture understanding	Baseline TBD	Baseline survey pushed to 3rd week of October	Survey sent Jan 2021	8 out of 26 Participants had a high level of understanding 10 out of 26 had some level of understanding	-
KR3: 4 new service components in production delivering new capabilities, enabling and adopting industry standard best practices for architecture, engineering and deployment, allowing for the mitigation of risks for both development teams and operational stakeholders, building trust in our development processes. 1 service in be end of Q2, 2 by end of Q3, and 4 by end of Q4 **Baseline:** Zero services	4 services decoupled	50% (2 services)	50% (2 services)	75% (3 services)	-

Department: Technology


Wikimedia Architecture
Phoenix Proof of Value
0:02 / 5:40

Efficacy & Resilience OKRs

Front Line Defenses

Objective:

Our infrastructure and data are staffed, secured and provisioned appropriately in each area to successfully prevent or handle malicious attacks, the unavailability of one system component, or the unavailability of a staff member.

Our 2nd EMEA region data center faced some initial setbacks in Q2/Q3, as the project was more complex than originally anticipated from a Legal standpoint. Following a thorough analysis that the Legal team provided, the C-team gave the green light for a buildout in the proposed locations in late March, with **Marseille** as the chosen location of the buildout. The project team has been working full steam in April to put the buildout back in its original time frame. We are **on track** to go live in Marseille in Q1 FY21-22, meeting the original plan for this project.

Three more positions were filled in Q3: a Privacy Engineer in the Security team, an SRE in the Traffic team, and an additional SRE focused in our network in the SRE Infrastructure Foundations team. This places the total number of "*SPOF hires*" to **four** for this FY, meeting the key result for this objective.

Continue to add maturity to our cybersecurity capabilities to address risk and threat stemming from disinformation, Security incident response, vulnerability management and privacy.

Target quarter for completion: Q1 FY21-22

Department: Technology

Front Line Defenses

Key Results	Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
KR1: We will plan (by end of Q3), build (Q4) a second EU data center, to better serve EMEA users (by end of FY21/22 Q1). **Baseline:** one EU data center	Two EU data centers	1 EU data center (planning not started yet)	1 EU data center Shortlist of cities prepared	1 EU data center City and site identified; contract negotiations in progress	-
KR2: The number of staff members solely responsible for a system or piece of software in production is reduced by 50%, from 8 to 4, by end of Q4. **Baseline:** 0 SPoF positions filled	Hire for 4 SPoF positions	0% filled 1 hiring process in final stages	25% filled 4 SPOF positions in active hiring	100% filled All 4 SPOF positions in plan hired	-
KR3: Implement a threat identification and risk treatment program (Q1) and deploy effective countermeasures for the top 2 threats and risks per quarter thereafter. **Baseline:** Partial risk management policy and 0 threat intelligence feeds.	6 threat and risk countermeasures deployed	Completed Q1 items	4 countermeasures deployed	2 countermeasures deployed	-

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.

Toolhub catalogue project: We are making good progress and have implemented faceted search and further functionalities in Q3.

Small Wiki Toolkits initiative: The South Asian workshop series are well underway, with 3 workshops conducted in Q3 and more to come in Q4. The workshops series with Kurdish communities will kick off in early April. While we're receiving good feedback from participants, we still need to explore ways to measure long term impact once this workshop round is completed.

Technical documentation: The multi-layer complexity of our technical documentation as well as the need for clear and scalable processes has become very visible in our research and interviews with staff.
We have incorporated the input on key technical documents from staff, and are moving the community input phase into Q4.

Target quarter for completion: Q4 FY 21/22

New 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: Communities find the tools that they need through the new Toolhub catalog system (technical plan Q1, working prototype Q2, MVP Q4), and developers create more diverse solutions using a user-extensible, container-based default deployment process (working prototype Q4) in Toolforge. **Baseline:** n/a	Toolhub MVP, working prototype for new deployment process	Q1 goal complete	Q2 goals complete	Q3 goals, 95% complete	-
KR2: A 10% increase in number of tool maintainers (4% by end of Q2) speaks to a vibrant ecosystem of technical contributors, supported by an iterative model and practice of community and capacity building (Q1 draft, Q4 MVP) which has been refined through 3 initiatives in focused outreach to technical communities (Q1-Q4). **Baseline:** 1974 tool maintainers	10% increase in number of tool maintainers, MVP model of community and capacity building	3% (2033), draft concept	4.9% (2067)	7.5% (2122)	-
KR3: An organization strategy for key technical documents informs a continuing roadmap based on a process of drafts (Q1) refined by consultation with staff (Q2) and community input (Q3) and a prototype of a single entry point to lower barriers finding existing documentation (Q4). **Baseline:** n/a	Prototype for a single entry point for technical documentation	Q1 goals complete	Interviews with staff in progress, not completed	Community input moved to Q4	-

Department: Technology

Production Quality

Objective:

Engineering teams at WMF have a shared understanding of development velocity, production health, and code quality, and they develop commitments and workflows for improving overall code health.

The purpose of this Objective is to streamline and strengthen our development, deployment, and hosting practices to improve the daily lives of our technical community.

Most of this work is highly cross-team and cross-department dependent *(Service Level Objectives and Code Health Objectives {SLOs & CHOs})* and we want to thank those who have participated thus far for their engagement. We look forward to more engagement in the remainder of the year.

We are working with an outside contractor to do an MVP setup of GitLab for use by a few early adopters by the end of Q4.

Last year we set the ambitious goal of 95% of production traffic being served from our new Kubernetes cluster and we stayed on track for 3 quarters. We don't expect to hit the final 95% goal by the end of next quarter. None of the work we did in Q3 or will do in Q4 came as a surprise, but the timelines for the work were unknown. This work represents one of the most significant changes to how we run software in production, and we must be thoughtful about the path forward

Target quarter for completion: Q4 FY20-21

Department: Technology

Production Quality

Key Results	Goal	Q1 Status	Q2 Status	Q3 Status	Q4 Status
KR1: Evangelize, implement tooling for (by end of Q3) and define and implement Service Level Objectives (SLO) and Error Budgets for our top 10 services and systems (Q4: 2 services; 2021-22 Q1: 4 services; 2021-22 Q2: 4 services) and report and iterate on them with product owners on a quarterly basis to optimize the balance of speed of innovation and reliability. **Baseline:** 0 services with SLOs & Error Budgets defined	10 services with SLOs and error budgets	0 services Developed SLO definition worksheet, Technical SLO baseline work completed	1 service	2 services plus 3 inflight	-
KR2: Educate teams on and define Code Health Objectives for our production-deployed code repositories (20% in Q1, 40% in Q2, 70% in Q3, 100% in Q4) and report and iterate on them with code owners on a quarterly basis. **Baseline:** 0% coverage of CHOs *(Code Health Objectives)*	100% coverage	17% of deployed repos have CHOs identified	28% of deployed repos have CHOs identified	23% of deployed repos have CHOs identified-	-
KR3: Evaluate and shepherd a consultation and best practices recommendations on a potential move to GitLab for code review by the end of Q2. In parallel, 70% of Wikimedia developed application layer production service traffic is served by the Deployment Pipeline in Q3 and 95% by Q4 **Baseline:** 27k request per second (rqs) *(baseline was measured in FY1920Q4 - eqiad usage)*	Code Review decision, 95% of traffic	GitLab consultation in progress and on-track; 54% of traffic	GitLab consultation completed; 68% of traffic	Working with GitLab contractors on prod install 68% of traffic	

Service Level Objectives

Reporting period:
Dec 1 2020 - Feb 28 2021

Service	SLO		Performance	SLO met?
etcd	Availability	**99.9%** of requests are successful (i.e. without errors)	**100%**	✓
	Latency	**99.8%** of requests under 32 ms	**99.994%**	✓
API Gateway	Availability	**99.9%** of requests proxied successfully (i.e. without errors caused by the Gateway)	**99.988%**	✓
	Read latency	99% of read requests under **2000 ms**	99% < **480 ms***	✓
	Write latency	99% of write requests under **1500 ms**	insufficient data*	?

* see speaker notes

WIKIMEDIA
FOUNDATION

Drill Down: Production Quality

The situation

Old wording:
Evangelize, implement tooling for (by end of Q1) and define and implement Service Level Objectives (SLO) and Error Budgets for our top 10 services and systems (Q2: 2 services; Q3: 4 services; Q4: 4 services) and report and iterate on them with product owners on a quarterly basis to optimize the balance of speed of innovation and reliability.

New wording:
Evangelize, implement tooling for (by end of Q3) and define and implement Service Level Objectives (SLO) and Error Budgets for our top 10 services and systems (Q4: 2 services; 2021-22 Q1: 4 services; 2021-22 Q2: 4 services) and report and iterate on them with product owners on a quarterly basis to optimize the balance of speed of innovation and reliability.

The impact

In the early groundwork and tooling stages, we uncovered a lot of complexity that we had to resolve in order to set up the later stages to be viable.

Now that that work is done, it's real progress toward completing the annual KR, it just wasn't anticipated.

Covid and staffing issues also slowed us down.

The recommendation

We will push KR1 deadlines out by 6 months - as we had discovered problems and fixed them.

We are now back on track.

Department slides

WIKIMEDIA
FOUNDATION

Anniversaries & New Hires

8 years
Greg Grossmeier

7 years
Gilles Dubuc
Leila Zia

6 years
Corey Floyd
Tyler Cipriani
Eric Evans
Joseph Allemandou

5 years
Emanuele Rocca
Luca Toscano
Deb Tankersley
Guillaume Lederrey

4 years
Anthony Borba
Jean-René Branaa

3 years
John Bennett
Brooke Storm
Sam Reed
Kate Chapman
Valentin Gutierrez

2 years
Holger Knust
David Sharpe
Brennen Bearnes
John Bond
Dominic Walden
Birgit Müller

1 year
Hugh Nowlan
David Pifke
Jennifer Cross
Mandy Mooney
Mat Nadrofsky

Research
Pablo Aragón

#wikilove

Legal

Superb work helping with discovering and planning for new data center sites; helping Machine Learning team understand governance of models

Abstract Wikipedia team

Engineer swap was fantastic!

APP Team

Much appreciation for the updates to the process and the space for strategic thinking and sharing

Recruiting / Julie Brown

Big thanks for landing that SRE candidate for the DE team… truly rescuing victory from the jaws of defeat

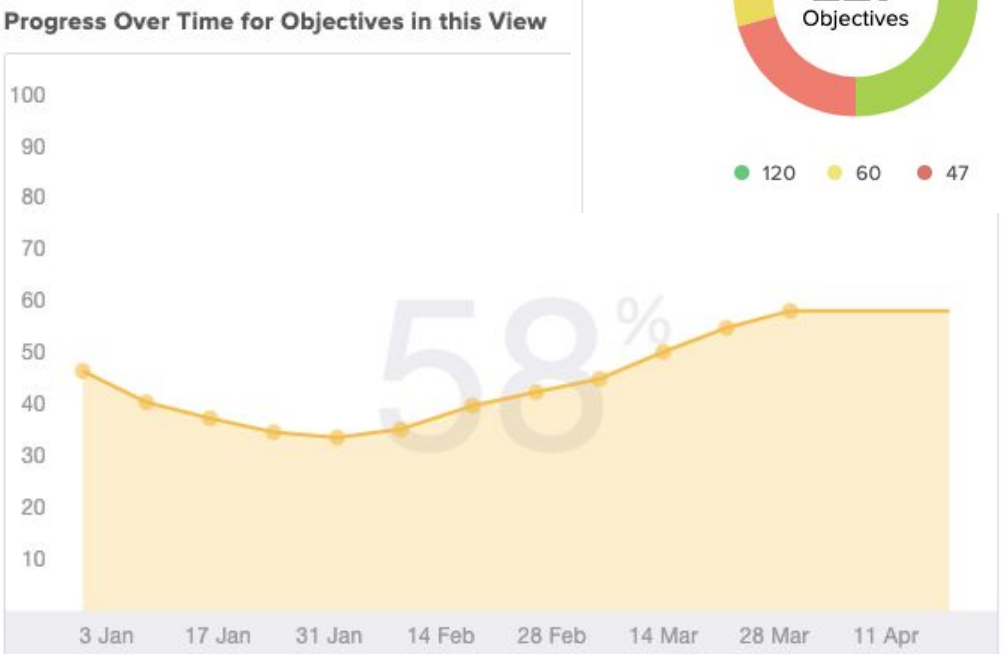
Product Dept

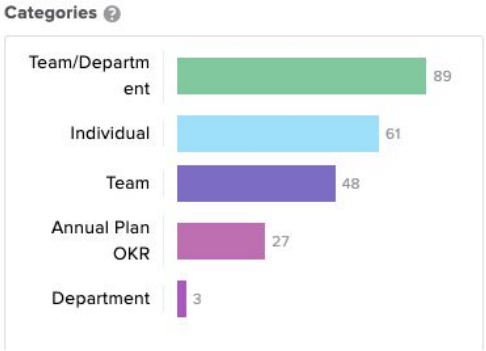
Service Level Objectives *(SLOs)* draft work is now more user oriented; collaboration on Technical Decision Making process — making it a reality

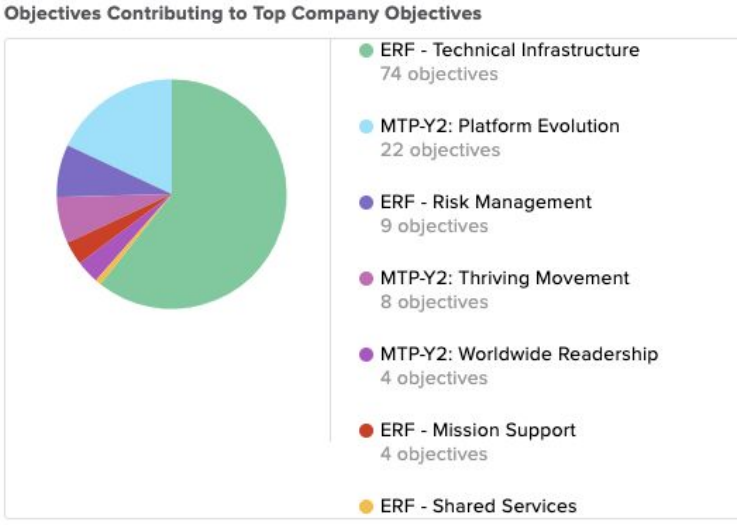
Partnerships

For tirelessly helping to push forward research in collaboration with external platforms, which is not always *(ever?)* a smooth process

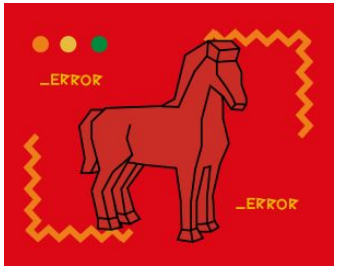
Department OKR Status Q3





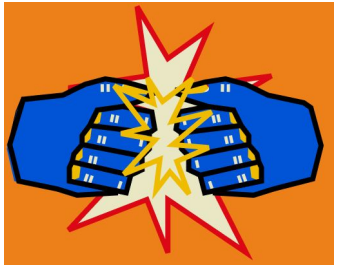


Challenges



Challenges on dependant teams on prioritizing finishing active-active; complexity of multi-dc setup for applications; lack of testing environments

Keeping pace with ever growing development teams and new features; even more patches pushed into production with occasional train stoppages



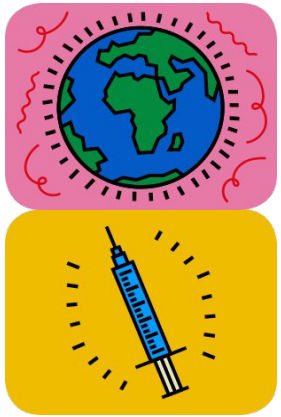
Virtual workshop attendees with different skill levels are harder to address than in real life

Societal unrest

Personnel outages, departures, hiring can be slow

COVID-19 / Vaccines

2021...so far



Supporting Wins

Service Operations

Mediawiki on Kubernetes and Buster - now in production

Data Persistence

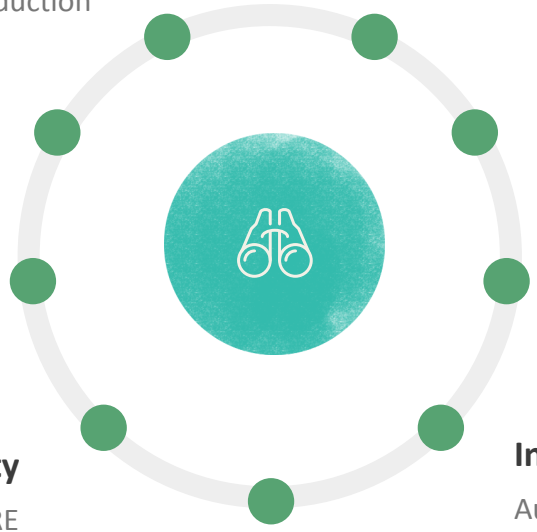
Final migration of wiki replicas nearing, hardware specs for backups and Swift are aligned

Data Center Operations

Finalized new data center site location - Marseille, FR, new Dell discount program

Security

Collaboration with Platform Eng and SRE teams; made improvements to services



Search Platform

Khmer search now has better recall *(more relevant results)*, SDoC data is joined with Wikidata graph

Machine Learning Platform

Lift Wing MVP is nearly done, with new models to deploy quickly, with a goal of having governing processes for all models

Traffic

SLO being defined for Varnish, new team member starting in April

Infrastructure Foundations

Automation tools to be available to developers *(without needing root access)*, fewer SPoFs of personnel and infrastructure

Observability

Old paging system retired, new reproducible SLO dashboard created

even moar Supporting Wins

Platform Engineering

MediaWiki kubernetes configuration ready *(Shellbox work done)*, paying down tech debt with Expedition work

Performance

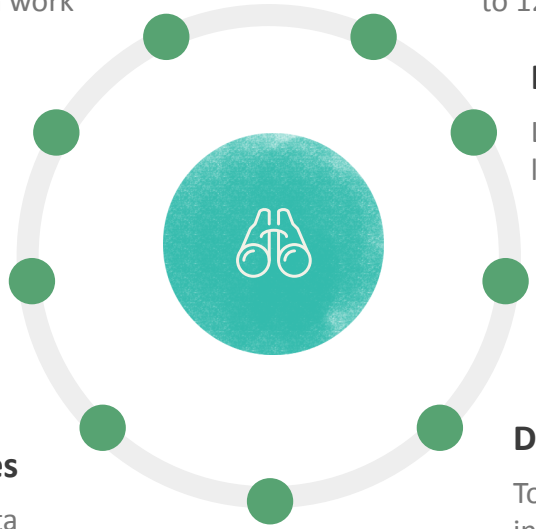
High attendance at FOSDEM virtual devroom, alerts have more configurable infrastructure

Quality & Test Engineering

Assisted Product teams with key releases *(MediaSearch, Discussion tool, New Topic and Section Translation)*

Wikimedia Cloud Services

Created Ceph performance baseline, can quota storage and updated VMs with attachable storage



Release Engineering

Code coverage for Mediawiki went from >4 hours to 12 minutes; Phatality bug reporting is back up

Fundraising Tech

LATAM countries has customized translations/UX, learning to securely serve donor data *(to donors)*

Research

Testing 3 language agnostic list building models tested, finalizing metrics for knowledge gaps, prepping Wiki Workshop

Developer Advocacy

Toolforge/Cloud user survey contained helpful insights, Outreachy / GSOC: 13 project promoted

Architecture

Released proof of value / artifact on structuring Wikimedia content, better understanding on use of events in systems

Questions?

Appendix

WIKIMEDIA
FOUNDATION

Questions about this deck?

Content Integrity: Leila Zia
Evolutionary Architecture: Kate Chapman
Front Line Defenses: Faidon Liambotis / John Bennett
Technical Community Building: Birgit Müller
Production Quality: Greg Grossmeier
Department Slides: Deb Tankersley