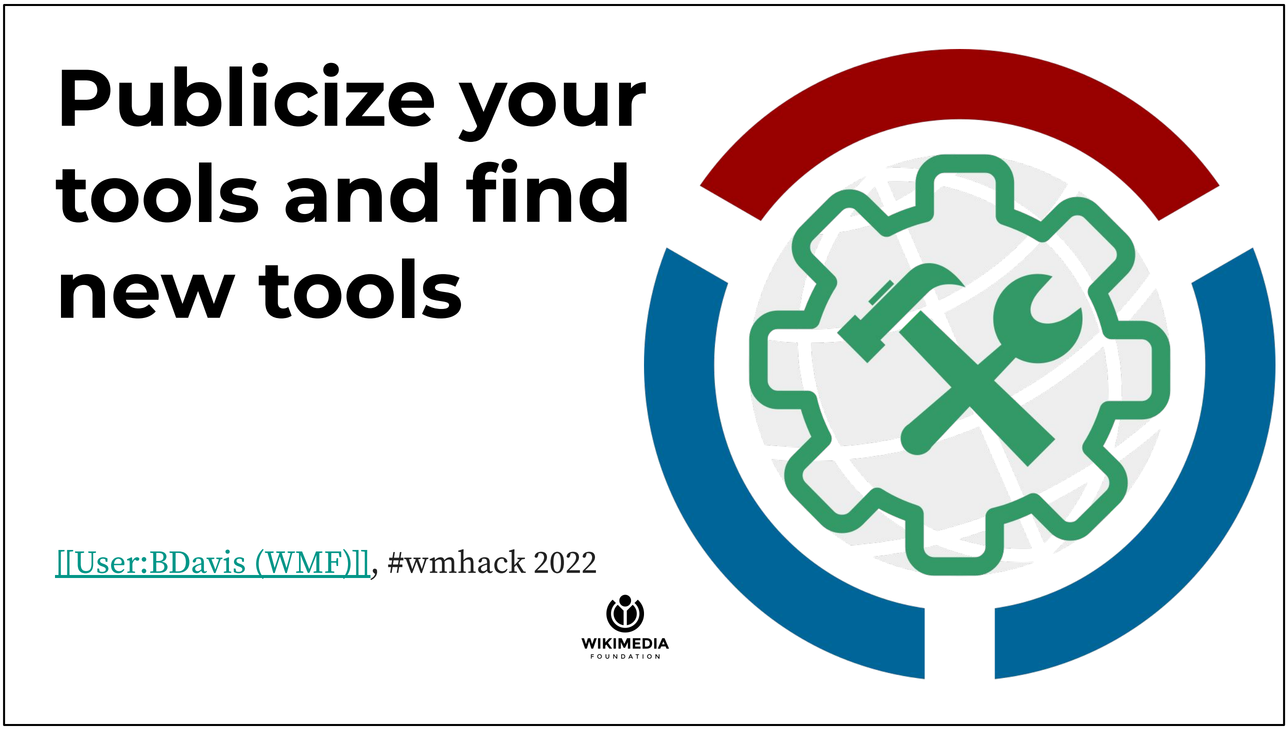
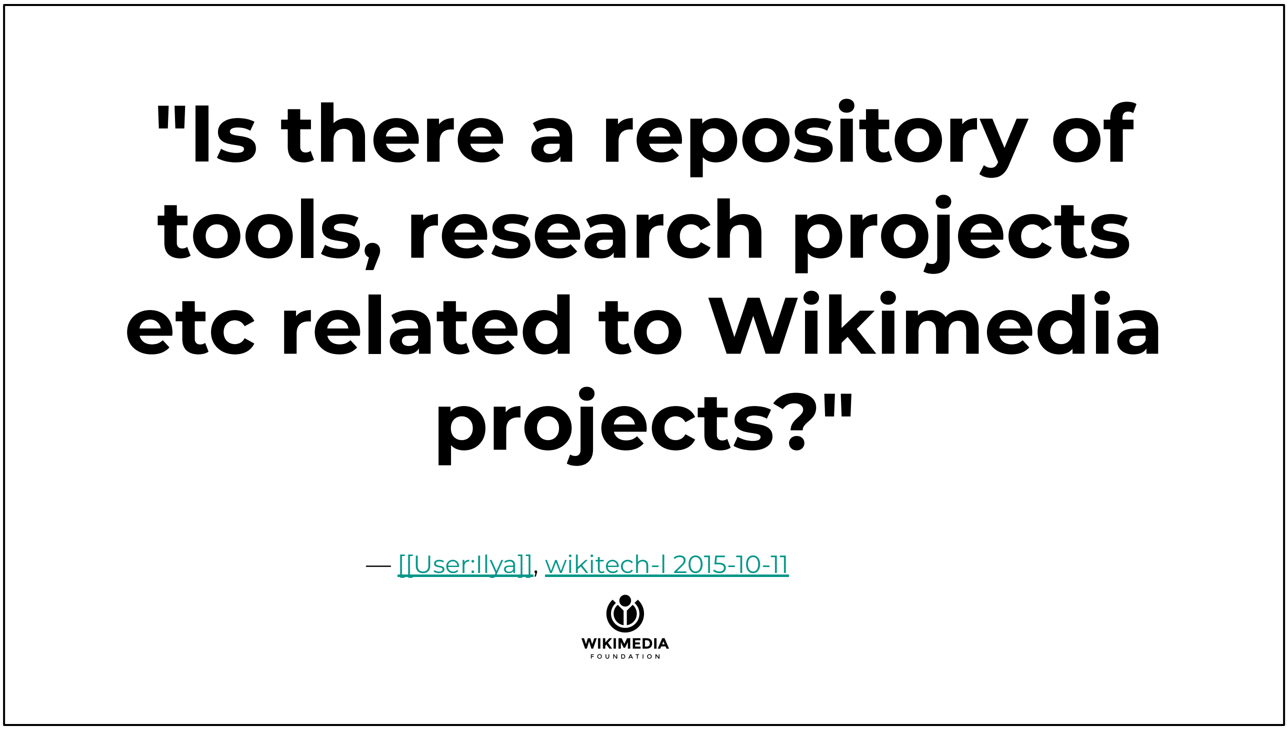


# Publicize your tools and find new tools

[[User:BDavis (WMF)]], #wmhack 2022

Hello everyone. My name is Bryan Davis. I am a software engineer at the Wikimedia Foundation working on the Technical Engagement team. My preferred pronouns are he/him.

I'm here today to talk to you about the Toolhub project and how you can use it to publicize the tools you have built within the Wikimedia community and also find new tools. Folks should leave the session with a general understanding of how to add tools to Toolhub, ways to find tools in the catalog, and how you can help improve the data quality of Toolhub.



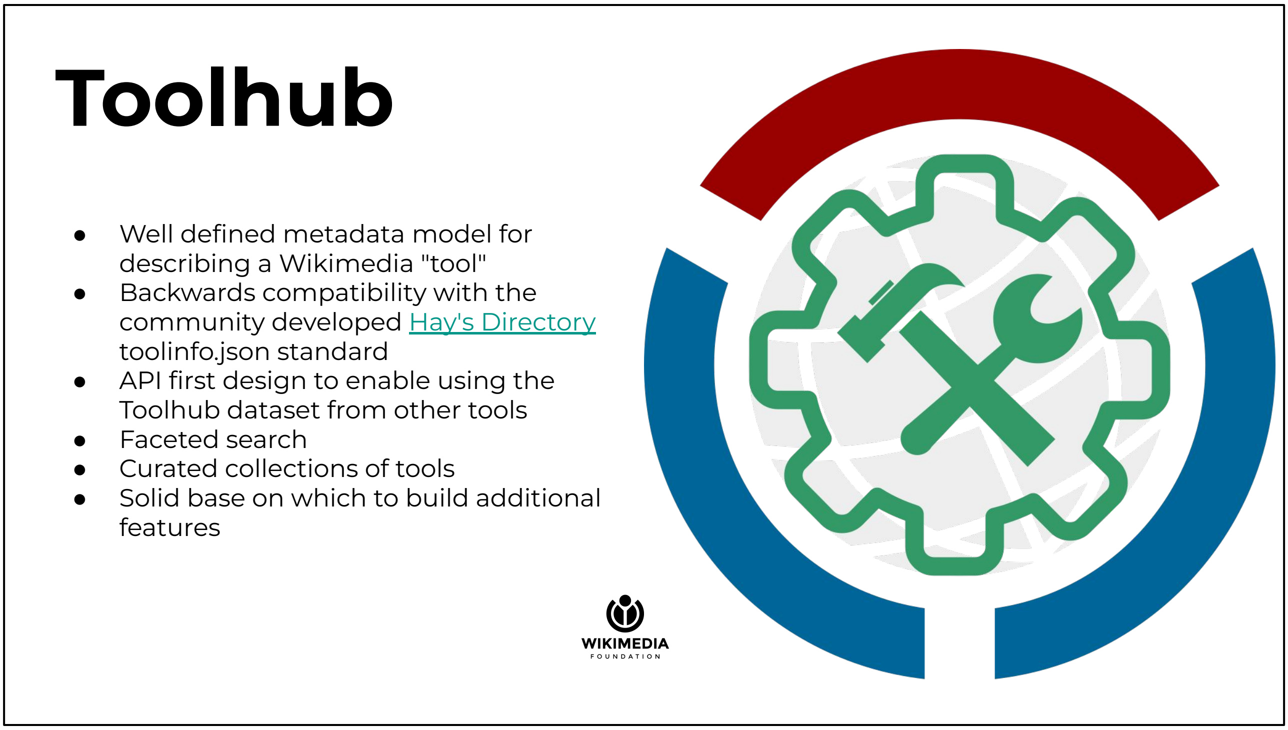
# "Is there a repository of tools, research projects etc related to Wikimedia projects?"

— [[User:Ilya]], wikitech-l 2015-10-11

WIKIMEDIA
FOUNDATION

There is a rich ecosystem of "tools" built by volunteers and staff to help fill in workflow gaps in the Wikimedia movement. There are thousands of bots, user scripts, web services, gadgets, desktop apps, and phone apps out there. Maybe even one that makes the exact thing you are trying to do easier (or possible). But how do you find them?

This is a problem that individuals and projects within the community have struggled with for years. There are many, many lists of tools on wiki pages across the movement. There are also lists in spreadsheets, mailing list messages, chat archives, and very likely even physical paper sitting on someone's desk. These lists are great, and I think they could be even better with some help.



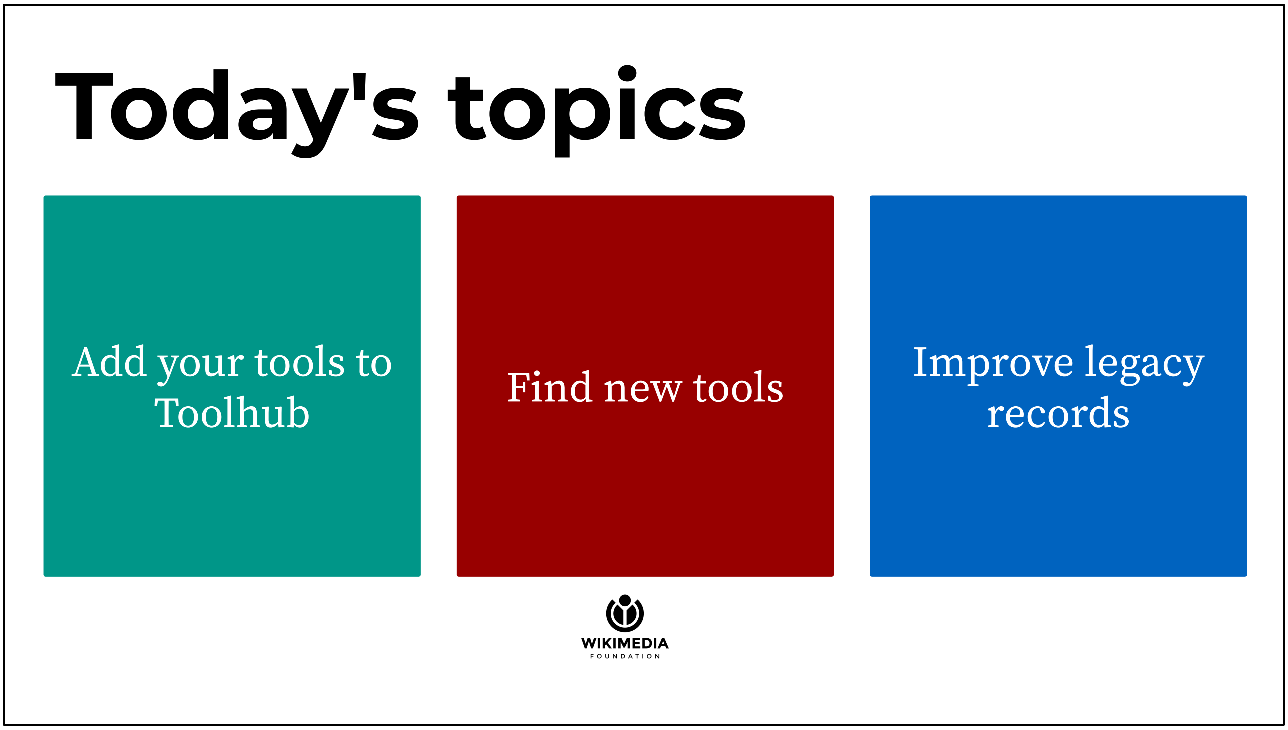
# Toolhub

- Well defined metadata model for describing a Wikimedia "tool"
- Backwards compatibility with the community developed Hay's Directory toolinfo.json standard
- API first design to enable using the Toolhub dataset from other tools
- Faceted search
- Curated collections of tools
- Solid base on which to build additional features

WIKIMEDIA FOUNDATION

Toolhub is a community-authored catalog of Wikimedia tools. We have tried to build a core product that makes collecting and reusing information about tools as open as we can. Rather than yet another one time list of tools, we want a platform that makes it possible to extend and remix the catalog.

Critical to this openness is our "API first" design. A "web API" is a fancy way of saying that the web application has features that can be used by other software, rather than just humans. And, in our case, Wikimedia volunteers should be able to build tools that interact with the data stored in Toolhub in many ways.



# Today's topics

Add your tools to Toolhub

Find new tools

Improve legacy records

WIKIMEDIA
FOUNDATION

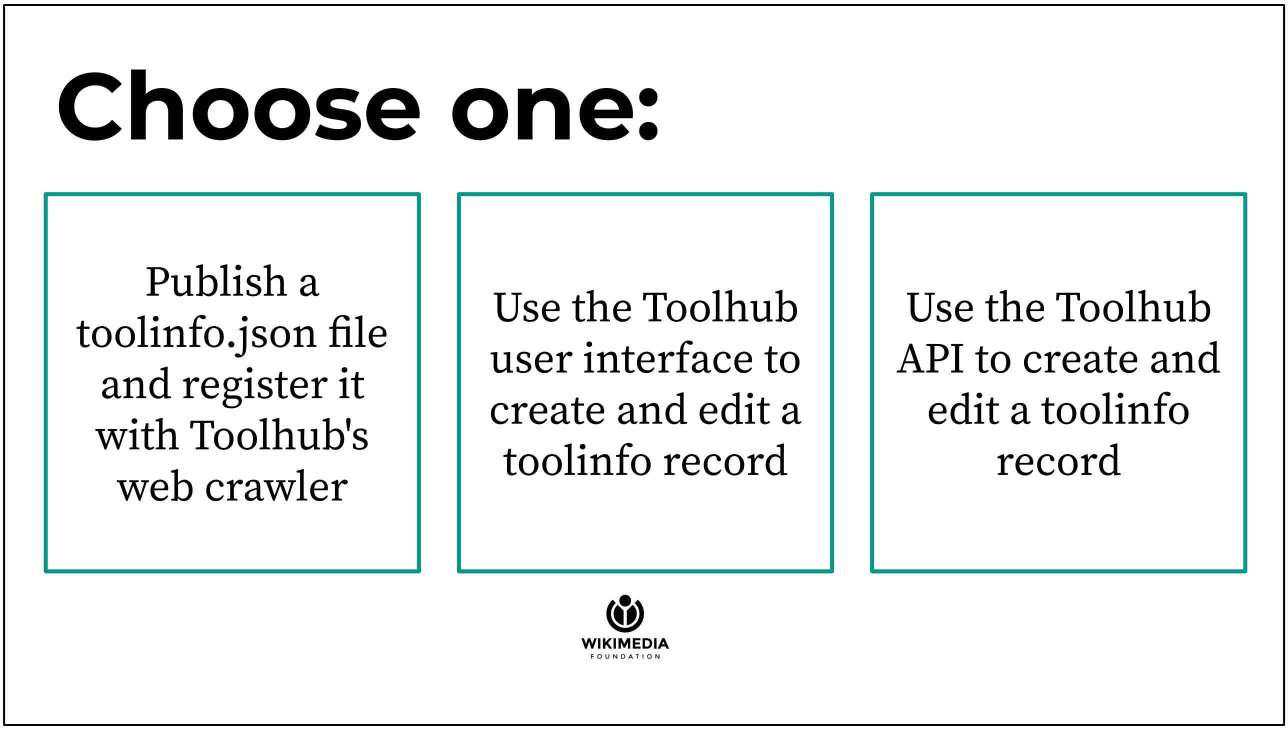
I have given several talks on Toolhub in the past that walked though details like the deep history of the idea; our experiments in using an advisory board, decision record, and weekly status posts; and details about the tech stack used to build the application. I have also give talks which were basically lists of the screens in the user interface with some information about what can be done on each of them.

For today's talk I decided to be a bit more task oriented. I am going to try to cover 3 tasks: adding your tools to toolhub, using toolhub to find new tools, and improving legacy records in Toolhub via community editing.

# Add your tools to Toolhub

WIKIMEDIA
FOUNDATION

Most of us here at the hackathon are folks who build tools. Sometimes we do that just to learn a bit about a new API or programming language or hosting platform. More often though we build tools to help us and others contribute to the Wikimedia projects. Those tools are usually also ones that we want others to find and use to make their contributions better or easier as well. One way to share your tools with others is to make sure they are added to Toolhub.



# Choose one:

Publish a toolinfo.json file and register it with Toolhub's web crawler

Use the Toolhub user interface to create and edit a toolinfo record

Use the Toolhub API to create and edit a toolinfo record

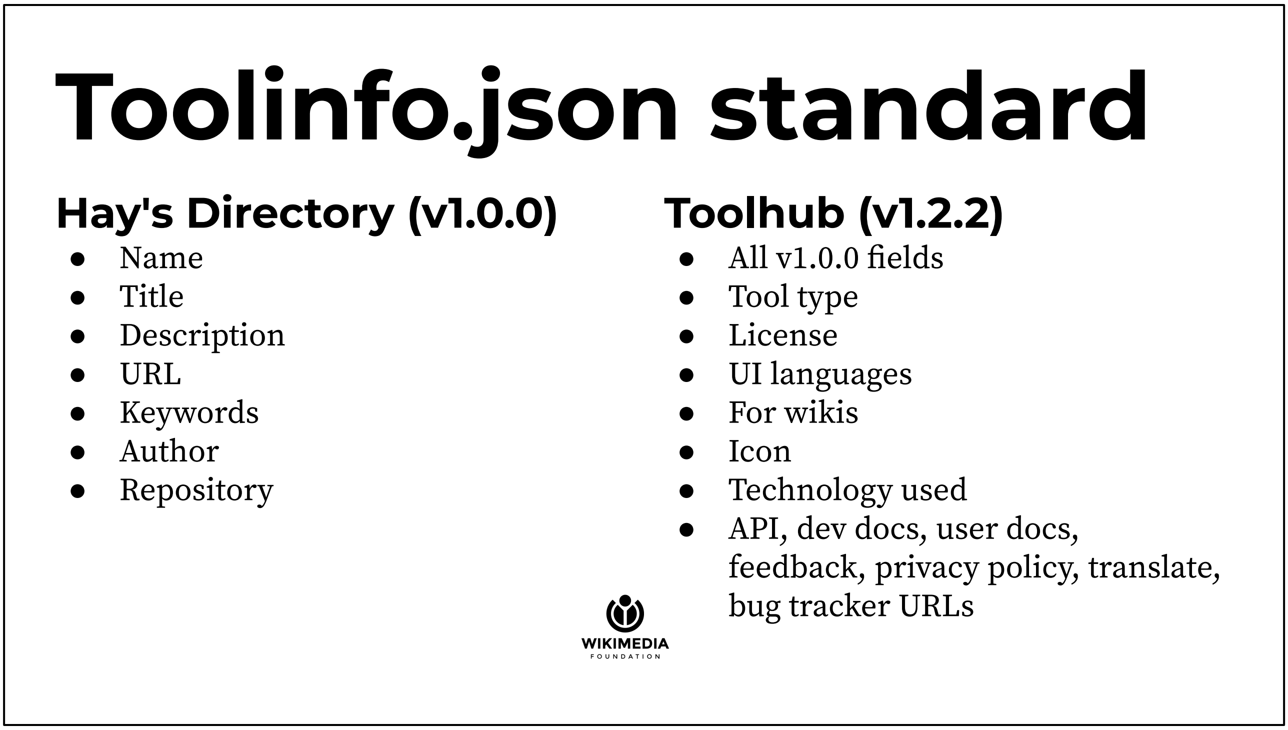
WIKIMEDIA
FOUNDATION

There are three ways to add a tool to Toolhub.

Publishing a toolinfo.json file is how most tools currently in the catalog got there. This is how tools were added to Hay's Directory, and as I mentioned earlier being backwards compatible with Hay's Directory was one of the design goals of Toolhub.

Toolhub also has a form based web interface that you can use to create and later update a toolinfo record about your tool. Depending on what type of tool you have built and how it is hosted, you might like using this method more than a toolinfo.json file.

Finally Toolhub has an API that you can use to submit information about your tool directly into the catalog. This API is used under the hood when you edit with the Toolhub UI.



# Toolinfo.json standard

## Hay's Directory (v1.0.0)

- Name
- Title
- Description
- URL
- Keywords
- Author
- Repository

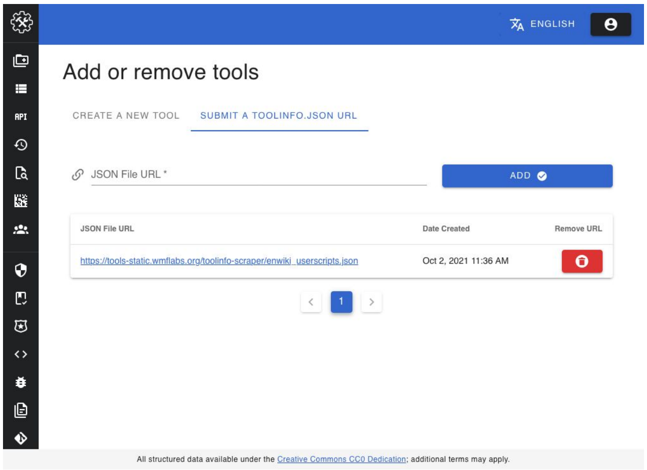
## Toolhub (v1.2.2)

- All v1.0.0 fields
- Tool type
- License
- UI languages
- For wikis
- Icon
- Technology used
- API, dev docs, user docs, feedback, privacy policy, translate, bug tracker URLs

Using toolinfo.json files to describe your tool is a standard started by Hay (Huskey) with his Tools Directory (https://hay.toolforge.org/directory/). This was an awesome innovation for the Wikimedia tools community that came out of discussions at Wikimania in 2014. We made a very deliberate decision when designing Toolhub to start from this standard and build on it in a backwards compatible way.

# Register toolinfo.json

1. Create a JSON file conforming to the schema
2. Publish that file on a website
3. Register the URL of the file with Toolhub's web crawler



Add or remove tools

JSON File URL *

| JSON File URL | Date Created | Remove URL |
|---|---|---|
| https://tools-static.wmflabs.org/toolinfo-scraper/enwiki_userscripts.json | Oct 2, 2021 11:36 AM | |

All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

Tool maintainers can submit URLs to JSON files which use the toolinfo schema to describe their tools.

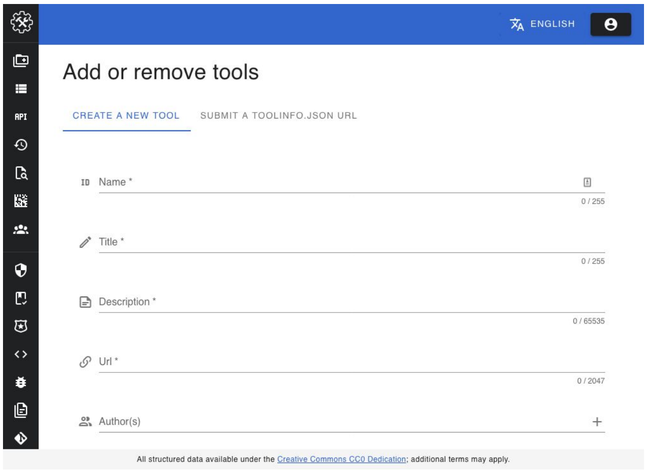
Toolhub will periodically visit the URLs, read the tool records from them, and update the catalog.

Toolinfo collected from a URL cannot currently be edited via the Toolhub API or user interface. Changes must be made in the external file and then imported by the crawler.

Tools hosted on Toolforge can use toolsadmin.wikimedia.org to create a toolinfo record which will automatically be imported by both Toolhub and Hay's Directory. The data you can add in toolsadmin is limited compared to the modern toolinfo.json standard, but we will cover a way to supplement this basic information later when we talk about improving legacy records.

# Create via UI

- *`Add or remove tools > Create a new tool`*
- Use a form to create your initial toolinfo record.
- The creating user "owns" the record.
- The record owner can edit it via the UI and API.
- Other users can add limited data called "annotations" via the UI and API.



Add or remove tools

Name *
0 / 255

Title *
0 / 255

Description *
0 / 65535

Url *
0 / 2047

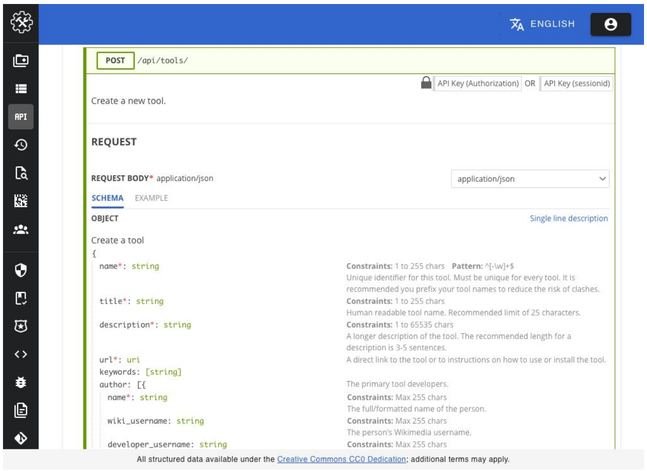
Author(s)

All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

Users can create new toolinfo records using the "add or remove tools" screen.

# Create via API

- `POST /api/tools/` endpoint
- The creating user "owns" the record.
- The record owner can edit it via the UI and API.
- Other users can add limited data called "annotations" via the UI and API.



POST /api/tools/

Create a new tool.

REQUEST

REQUEST BODY* application/json

OBJECT

Create a tool

```
{
  name*: string                      Constraints: 1 to 255 chars   Pattern: ^[-\w]+$
                                     Unique identifier for this tool. Must be unique for every tool. It is
                                     recommended you prefix your tool names to reduce the risk of clashes.
  title*: string                     Constraints: 1 to 255 chars
                                     Human readable tool name. Recommended limit of 25 characters.
  description*: string               Constraints: 1 to 65535 chars
                                     A longer description of the tool. The recommended length for a
                                     description is 3-5 sentences.
  url*: uri                          A direct link to the tool or to instructions on how to use or install the tool.
  keywords: [string]
  author: [{                         The primary tool developers.
    name*: string                    Constraints: Max 255 chars
                                     The full/formatted name of the person.
    wiki_username: string            Constraints: Max 255 chars
                                     The person's Wikimedia username.
    developer_username: string       Constraints: Max 255 chars
```

All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

The POST /api/tools/ endpoint of Toolhub's API can be used to create a new toolinfo record directly in Toolhub's catalog. This API is used behind the scenes by the Toolhub user interface. See Toolhub's API documentation for detailed information about the expected inputs and outputs of this API endpoint.

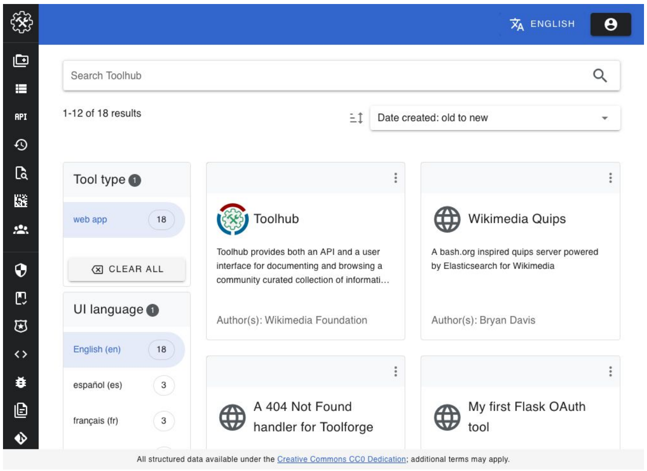
# Find new tools

WIKIMEDIA
FOUNDATION

Now that you have your tools in the catalog you might want to use it to look for some new tools that others have made that could help you contribute to the wikis and the movement.

# Faceted search

- Full text search using technology similar to the on-wiki search
- Searches can be refined using filters based on faceted classification of various tool info fields like tool type, UI language, keywords, and target wiki



1-12 of 18 results

Toolhub

Toolhub provides both an API and a user interface for documenting and browsing a community curated collection of informati…

Author(s): Wikimedia Foundation

Wikimedia Quips

A bash.org inspired quips server powered by Elasticsearch for Wikimedia

Author(s): Bryan Davis

A 404 Not Found handler for Toolforge

My first Flask OAuth tool

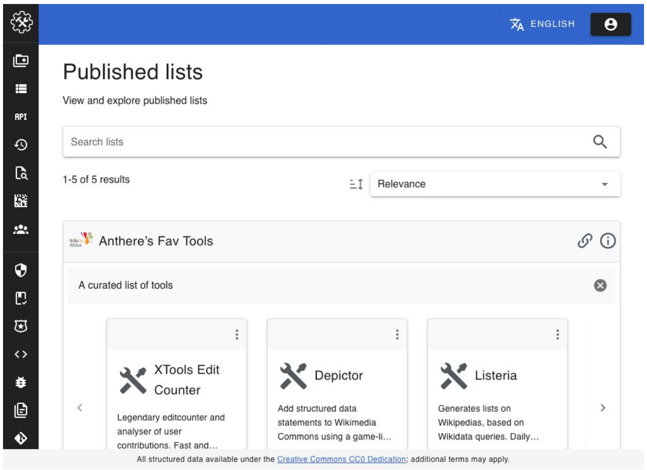
All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

You can search through the tools and refine those searches by selecting common values from the matched documents.

These "facets" are the sort of search navigation you have probably seen on ecommerce sites where a list of departments, or sizes, or colors is shown along with the results and you can click them to add that attribute and value to the search constraints.

# Browse lists

- Anyone can create lists of tools
- Lists can be private or public
- Public lists can be searched and browsed



Published lists

View and explore published lists

Search lists

1-5 of 5 results

Relevance

Anthere's Fav Tools

A curated list of tools

XTools Edit Counter

Legendary editcounter and analyser of user contributions. Fast and…

Depictor

Add structured data statements to Wikimedia Commons using a game-li…

Listeria

Generates lists on Wikipedias, based on Wikidata queries. Daily…

All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

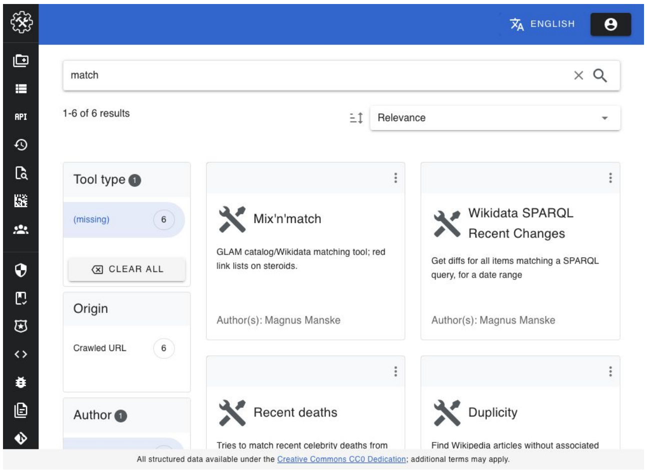
# Improve legacy records

The recently added feature of annotations editing allows everyone in the Wikimedia community to add data to existing toolinfo records. This can be used to fill in details that are missing from the toolinfo.json data submitted to the Toolhub web crawler or added directly to the catalog by the tool's maintainers.

# Find legacy records

Tool type is a new field that was not supported by Hay's Directory.

We can filter search results to show tools without a tool type which roughly equals searching for toolinfo.json v1.0 records.



match

1-6 of 6 results

Relevance

Mix'n'match

GLAM catalog/Wikidata matching tool; red link lists on steroids.

Author(s): Magnus Manske

Wikidata SPARQL Recent Changes

Get diffs for all items matching a SPARQL query, for a date range

Author(s): Magnus Manske

Recent deaths

Tries to match recent celebrity deaths from

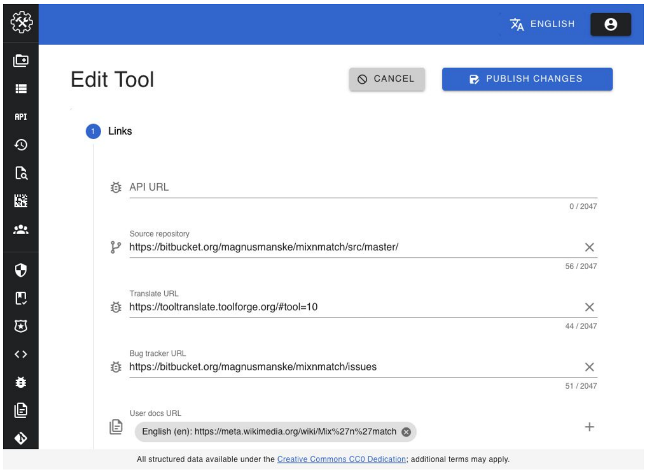
Duplicity

Find Wikipedia articles without associated

All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

# Add links

- API
- Source repository
- Translate
- Bug tracker
- User docs
- Developer docs
- Feedback
- Privacy policy



Edit Tool

1 Links

API URL
0 / 2047

Source repository
https://bitbucket.org/magnusmanske/mixnmatch/src/master/
56 / 2047

Translate URL
https://tooltranslate.toolforge.org/#tool=10
44 / 2047

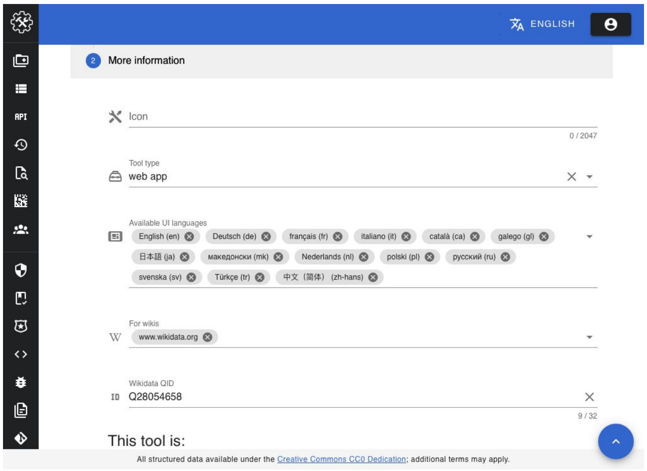
Bug tracker URL
https://bitbucket.org/magnusmanske/mixnmatch/issues
51 / 2047

User docs URL
English (en): https://meta.wikimedia.org/wiki/Mix%27n%27match

All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

# Add more information

- Icon
- Tool type
- UI languages
- Wikis
- Wikidata QID
- Deprecated
- Experimental
- Replaced by



2 More information

Icon

0 / 2047

Tool type
web app

Available UI languages
English (en) · Deutsch (de) · français (fr) · italiano (it) · català (ca) · galego (gl) · 日本語 (ja) · македонски (mk) · Nederlands (nl) · polski (pl) · русский (ru) · svenska (sv) · Türkçe (tr) · 中文（简体）(zh-hans)

For wikis
www.wikidata.org

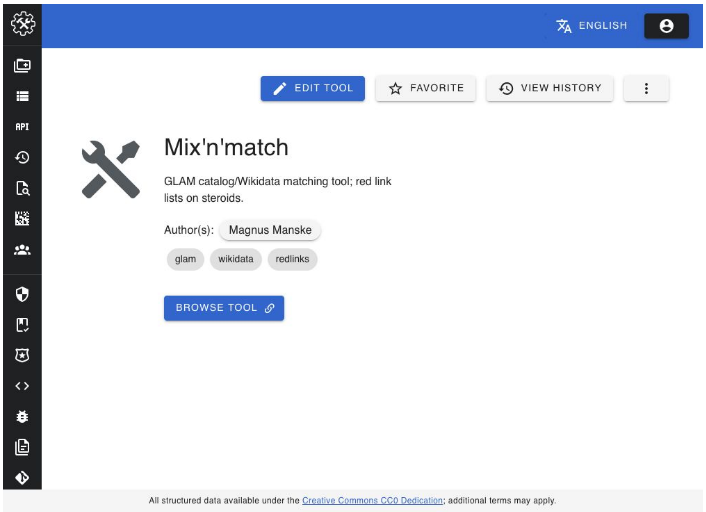
Wikidata QID
Q28054658

9 / 32

This tool is:

All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

# Mix'n'match before:



Mix'n'match

GLAM catalog/Wikidata matching tool; red link lists on steroids.

Author(s): Magnus Manske

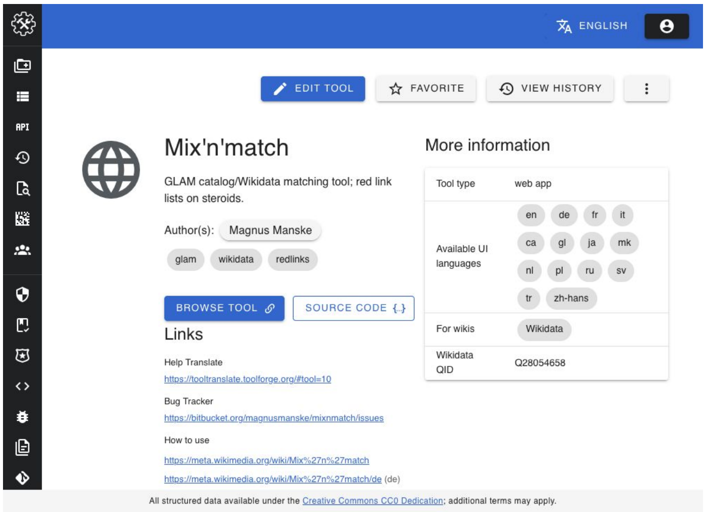
glam wikidata redlinks

All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

If you were looking closely at the last three slides, the screen shots were taken as I followed the steps of searching, adding links, and adding other information to the record for the Mix'n'match tool. Mix'n'match is a tool that Magnus Manske built to help importing various data sets into Wikidata.

Before my edits, mix'n'match had a pretty typical toolinfo.json v1.0 record. Name, description, author, keywords, and a link to the tool.

# Mix'n'match after:



## Mix'n'match

GLAM catalog/Wikidata matching tool; red link lists on steroids.

Author(s): Magnus Manske

glam wikidata redlinks

### Links

Help Translate

https://tooltranslate.toolforge.org/#tool=10

Bug Tracker

https://bitbucket.org/magnusmanske/mixnmatch/issues

How to use

https://meta.wikimedia.org/wiki/Mix%27n%27match

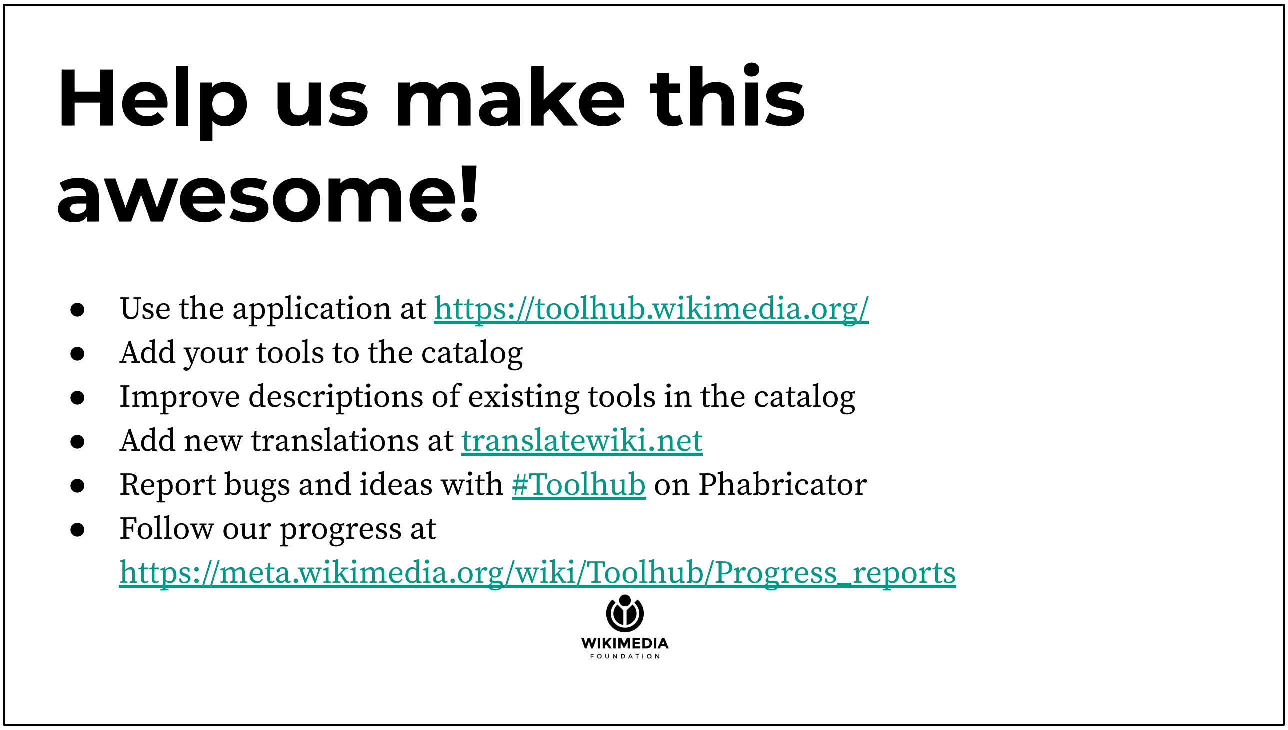
https://meta.wikimedia.org/wiki/Mix%27n%27match/de (de)

### More information

| | |
|---|---|
| Tool type | web app |
| Available UI languages | en de fr it ca gl ja mk nl pl ru sv tr zh-hans |
| For wikis | Wikidata |
| Wikidata QID | Q28054658 |

All structured data available under the Creative Commons CC0 Dedication; additional terms may apply.

By looking at the tool's user interface and the help docs it linked to I was able to find quite a few things I could fill using annotations. I think similar updates could be made for many of Magnus' 144 !!! tools that are currently listed in the Toolhub catalog.



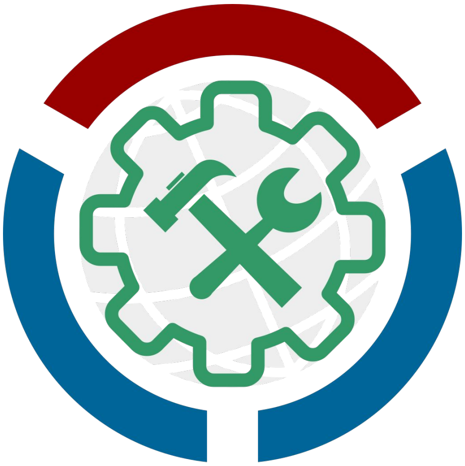
# Help us make this awesome!

- Use the application at https://toolhub.wikimedia.org/
- Add your tools to the catalog
- Improve descriptions of existing tools in the catalog
- Add new translations at translatewiki.net
- Report bugs and ideas with #Toolhub on Phabricator
- Follow our progress at https://meta.wikimedia.org/wiki/Toolhub/Progress_reports

WIKIMEDIA FOUNDATION

Hopefully you have heard something in this presentation that gets you a bit excited about Toolhub. We would love to have any and all of you join us in making this project the success that the movement deserves. You can use the application, add your tools, improve the records for other tools, help translate the user interface, and report bugs and feature ideas. You can even follow our progress each week on meta!

# Credits



- Toolhub community logo.svg By BDavis (WMF), CC0
- Toolhub screenshots By BDavis (WMF), CC BY-SA 4.0