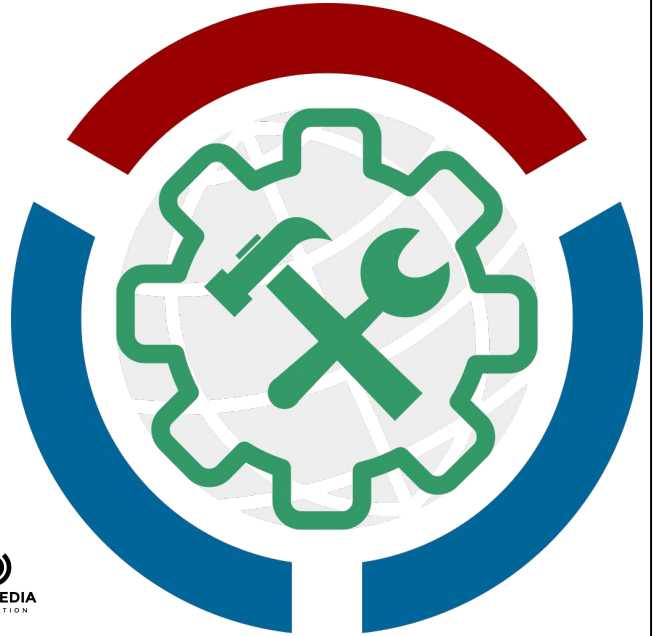


Introduction to Toolhub for tool maintainers

[\[\[User:BDavis \(WMF\)\]\]](#), wmhack 2021



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

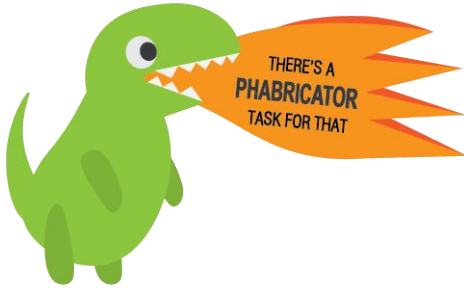
I'm here today to talk to you about the Toolhub project that Srishti Sethi and I have been working on for the last 5 months. Folks should leave the session with a general understanding of what gaps Toolhub is hoping to fill for the movement, how to get information about their own tools into the catalog, and where to look for discussion of future features and how to get involved in the project.

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

— [\[\[User:Ilya\]\], wikitech-l 2015-10-11](#)



There is a rich ecosystem of "tools" build 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?



T115650: Create an authoritative and well promoted catalog of Wikimedia tools

— [\[\[User:Ricordisamoal\], Phabricator 2015-10-15](#)



The wikitech-l discussion that I took the pull quote on the previous slide from inspired Ricordi Samoa to create a phabricator task that collected links to existing partial solutions. I discovered this task early in 2016 while researching ways to help the volunteer developers working in what we now call Toolforge. It was, and still is, a brilliant idea and a thing that many, many people have asked for over the years.

Toolhub 1.0 goals

- 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
- Community editable additions to core metadata
- Controlled vocabularies for classifying tools
- Faceted search
- Curated collections of tools
- Solid base on which to build future features



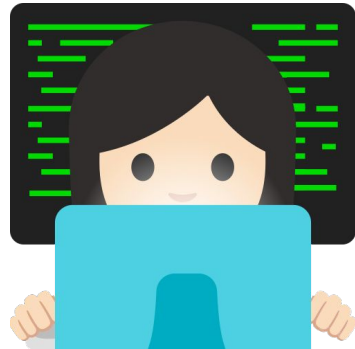
Our minimum viable product for Toolhub is focused on this set of goals. We are working towards a core product that makes collecting and reusing information about tools as open as we can. Rather than yet another one time list of tools, we want a platform that makes it possible to extend and remix the 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.

Personas & use cases

1. Editors

- Search for templates, modules, gadgets or tools to help with specific editing-related tasks on-wiki.
- Make or view public lists of tool categories to learn about tools for specific tasks.
- Contribute tools information.
- Write Lua module that queries Toolhub for certain types of tools and use the information returned on wikis.



For the first bullet point—think of the tools that allow for bulk editing, or adding geo data to commons files etc.

- Images:
https://commons.wikimedia.org/wiki/File:Emoji_u1f469_1f3fb_200d_1f4bb.svg

Personas & use cases

2. Developers

- Build catalog subsets via the Toolhub's API for personal or community use.
- Develop a gadget or user script for registering a tool in Toolhub.
- Learn about the tools available in a wide variety of programming languages they can contribute to (particularly helpful for new developers).
- Connect with users, each other, and resources such as documentation.



Images:

- https://commons.wikimedia.org/wiki/File:Emoji_u1f469_1f3fe_200d_1f3ed.svg

Personas & use cases

3. Researchers

Entry point to learn about specific tools (e.g., bots) available in the Wikimedia ecosystem.



Images:

- https://commons.wikimedia.org/wiki/File:Noto_Emoji_Oreo_1f469_1f3fb_200d_1f3eb.svg

Personas & use cases

4. Movement organizers

Search for a list of tools that help with organizing programs and events (for example Wiki Loves Monuments).



Images:

- https://commons.wikimedia.org/wiki/File:Noto_Emoji_Oreo_1f9d5_1f3fd.svg
- https://commons.wikimedia.org/wiki/File:Noto_Emoji_Oreo_1f9d1.svg
- https://commons.wikimedia.org/wiki/File:Noto_Emoji_Oreo_1f475_1f3fb.svg
- https://commons.wikimedia.org/wiki/File:Noto_Emoji_Oreo_1f472_1f3ff.svg

Personas & use cases

5. Readers

Find tools that recommend articles and new ways to experience Wikimedia content.



Images:

- https://commons.wikimedia.org/wiki/File:Emoji_u1f939_1f3fd.svg

Team and tech stack

- 1 backend + tooling developer
- 1 frontend developer
- 5 member advisory board
 - 1 SRE, 1 Security, 1 Community Relations, 1 wiki power user, 1 RTL native speaker and tech volunteer
- [#Toolhub](#) on Phabricator
- toolhub-dev@lists.wikimedia.org
- [Decision record](#)
- [Weekly progress reports](#)
- Docker-compose dev environment with few local system requirements
- [Blubber](#) + [PipelineLib](#)
- Django (Python) backend
- Vue.js + Vuetify frontend
- Material Design look and feel
- Cloud VPS [demo server](#)
- Planned Kubernetes production hosting



We are running some social and technical experiments in this project. We formed an "advisory board" to help get input from key roles in the movement during development. Our advisors currently include Foundation staff as well as community volunteers. These folks provide us with feedback on design and implementation ideas. That feedback helps us iterate on our collective thinking about the project from a broader set of perspectives.

We are also trying to leave behind documentation about why certain decisions were made in the form of a decision record. We want Toolhub to have a life beyond the contributions of any single member of the team. We hope that documenting why we have made certain technical choices will help future maintainers when they have to make similar editorial decisions.

To keep the advisors, and anyone else who is interested in following along, up to date with what Srishti and I have been working on we are also producing a progress report at the end of each week on meta and posting a summary to the project mailing list.

On the tech side, we are trying to keep the development environment requirements simple. We are also using newer technologies that are being adopted in other areas of the movement like Vue.js and container based deployment tools.

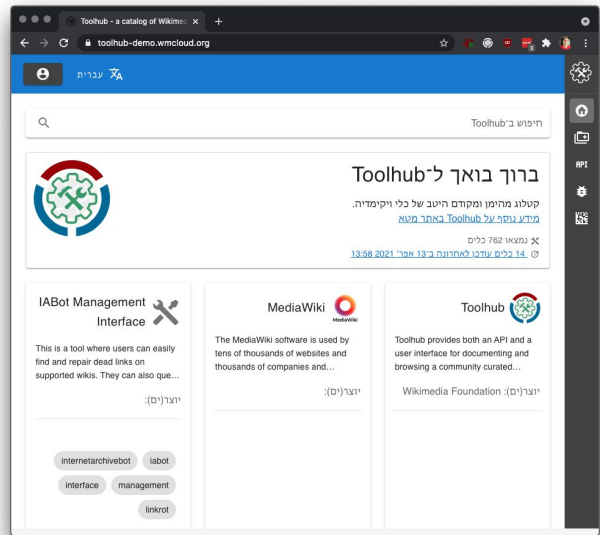
The dev environment uses docker-compose run containers for the Django backend, Vue frontend, MariaDB database, and Elasticsearch full text search. This environment builds on top of the Blubber and PipelineLib configuration that is used in CI and will be

used for the production deployment. It also encapsulates as much as possible in the Docker layer. On your local machine you only need Docker, git, and GNU Make. Everything else that is needed is included in Docker containers with Makefile targets to automate things like running tests and generating localization files. All coding standards are enforced with linters which can be run locally and are voting in CI. Srishti and I don't need to quibble during code review on trivial things like formatting. If it passes the linter then it must be ok--or a new linter is needed!

Features

00. Home page

- Search box
- Paginated display of tool info cards
- Over 700 tools available already!

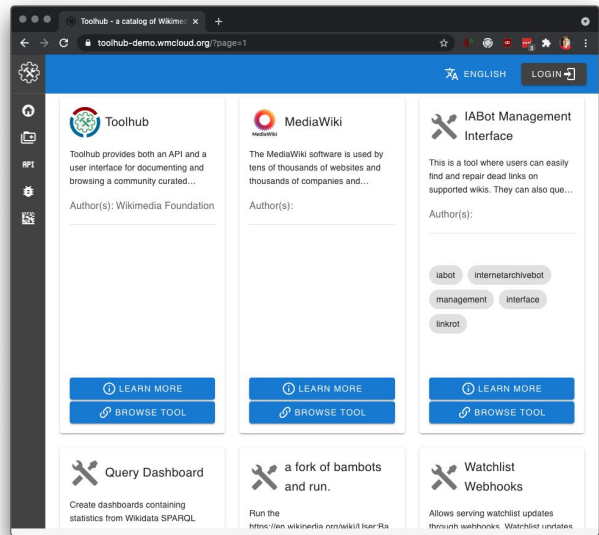


762 tools are currently available on the demo server thanks to our compatibility with Hay's Directory and it's toolinfo.json standard.

Features

01. Tool info cards

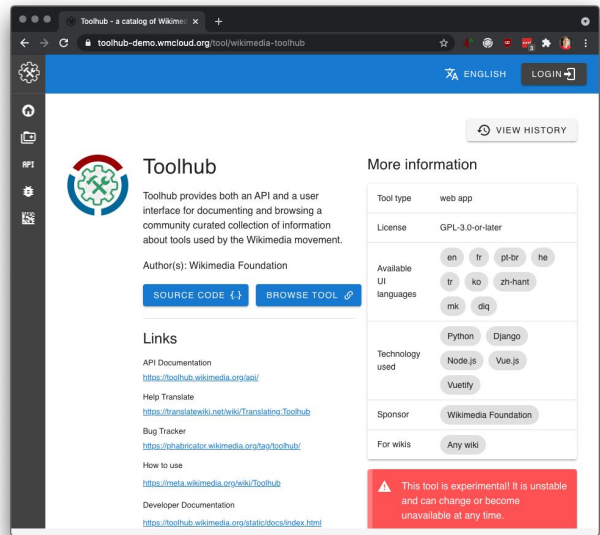
- Image, title, description, authors, and keywords from tool info shown
- "Learn more" drills down into tool info details
- "Browse tool" links to the tool info provided URL for the tool



Features

02. Tool info detail

- See detailed information about a tool (for example: the tool's type, wikis it is designed for, technology used, etc.)

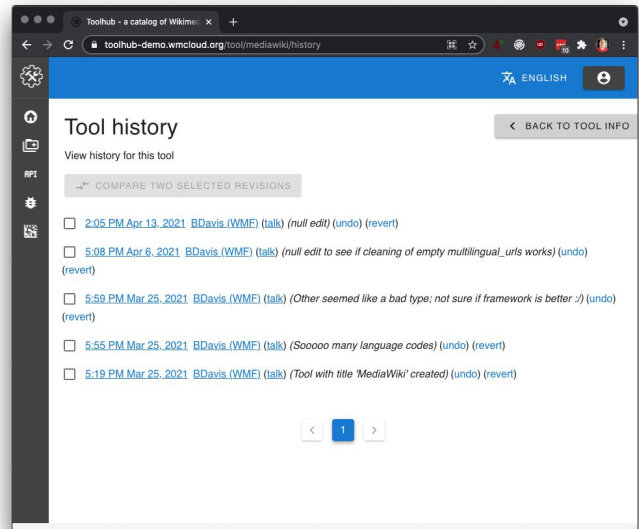


Each tool info card links to a more detailed page listing all the information that Toolhub has about the tool.

Features

03. Edit history

- View edit history for a tool info record
- Compare changes across between two revisions
- Undo the changes from a revision
- Revert to an older revision

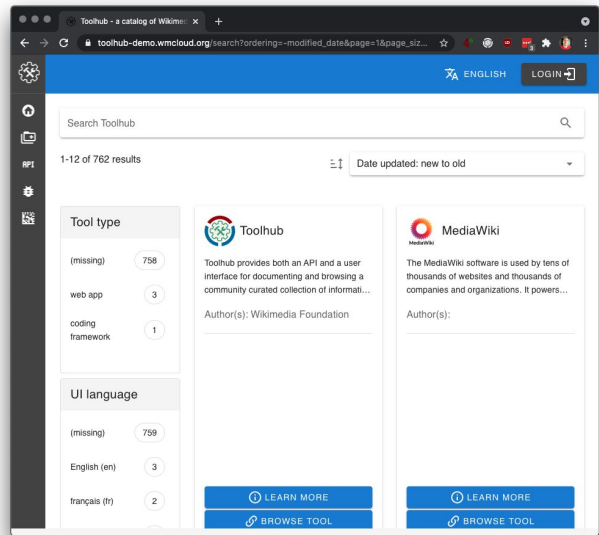


Each tool info card links to a more detailed page listing all the information that Toolhub has about the tool.

Features

04. 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

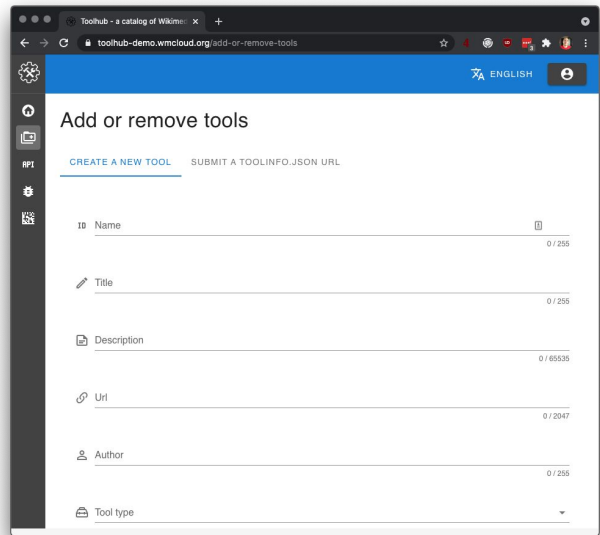


Users can search through the tools and refine those searches by selecting common values from the matched documents. These "facets" are the sort of search navigation you have probably seen on ecommerce sites where a list of departments, or sizes, or colors is shown along with the results and you can click them to add that attribute and value to the search constraints.

Features

05. Tool registration

- Create and maintain a tool info record directly in Toolhub



The screenshot shows a web browser window with the URL `toolhub-demo.wmcloud.org/add-or-remove-tools`. The page title is "Add or remove tools". There are two tabs: "CREATE A NEW TOOL" (active) and "SUBMIT A TOOLINFO.JSON URL". The form contains the following fields:

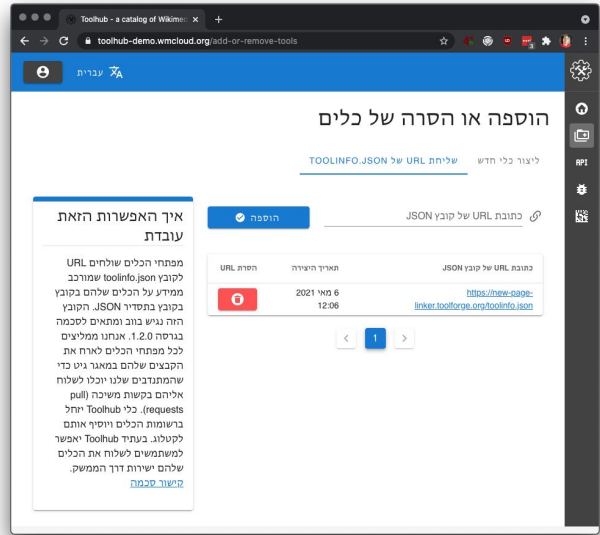
- Name**: Text input field with a character count of 0 / 255.
- Title**: Text input field with a character count of 0 / 255.
- Description**: Text input field with a character count of 0 / 65535.
- Uri**: Text input field with a character count of 0 / 2047.
- Author**: Text input field with a character count of 0 / 255.
- Tool type**: A dropdown menu.

Next is the adding or removing tools feature. To use this feature, you will need to login to Toolhub with your Wikimedia account. In the future, Toolhub will also allow its users to submit their tools directly via the interface.

Features

06. toolinfo.json URLs

- Register an externally hosted toolinfo.json to be indexed by Toolhub's web crawler
- Demo server seeded with all URLs registered with Hay's Directory



Tool maintainers can submit a URL to a JSON file consisting of information about their tools and adhering to a schema format. Toolhub will then periodically crawl the URL, read the tool records from it, and add them to the catalog.

Toolinfo.json standard

Hay's Directory (v1.0.0)

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

Toolhub (v1.2.0)

- 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

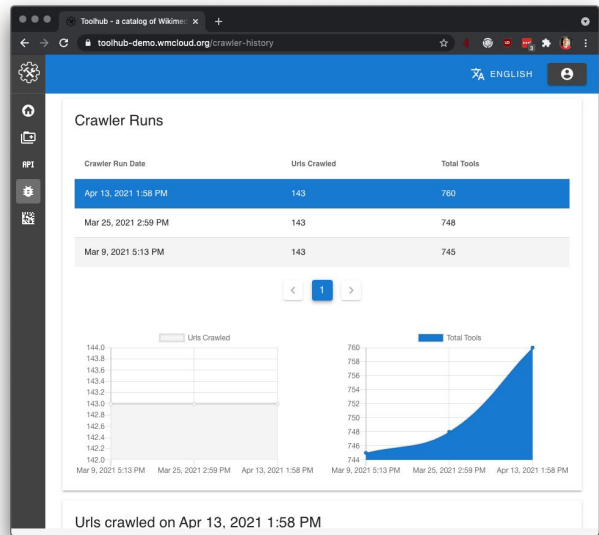


Using toolinfo.json files to describe your tool is a standard [started by Hay \(Huskey\)](#) with his Tools Directory (<https://hay.toolforge.org/directory/>). This was an awesome innovation for the Wikimedia tools community that came out of discussions at Wikimania in 2014. James Hare and I made a very deliberate decision when designing Toolhub to start from this standard and build on it in a backwards compatible way. This helps Toolhub by providing useful data even before we have launched the project for public use.

Features

07. Web crawler status

- See when the crawler has run, how many URLs it examined, and how many tools were found

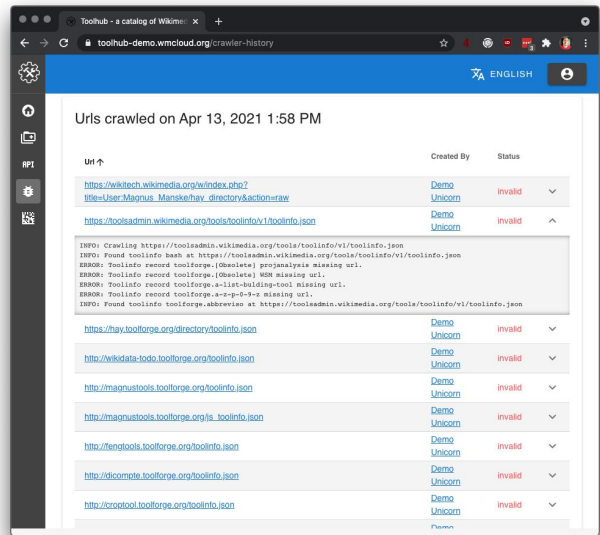


There is a little analytics feature that you can use to see the URLs crawled and tools added by Toolhub periodically.

Features

08. Crawler run details

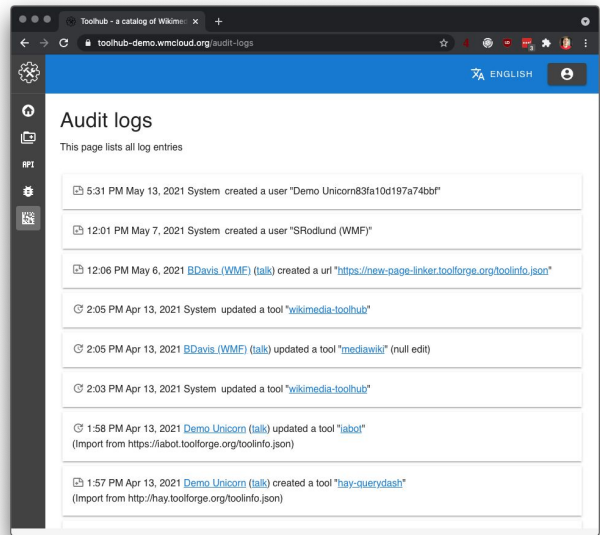
- View details about each URL crawled during a crawler run
- Drill down provides information about specific errors that the crawler encountered fetching or parsing a toolinfo.json URL



Features

09. Audit logs

- Similar to Special:Log and Special:RecentChanges combined into one list of actions taken
- This screen will eventually allow you to filter logs by data range, user, and action type

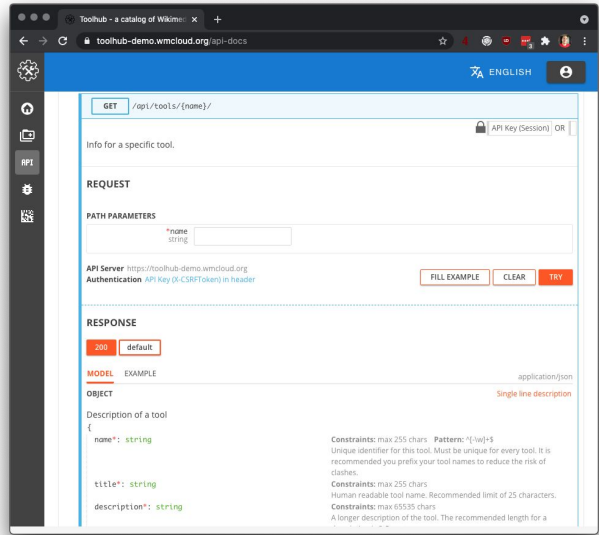


Then there is an audit logs feature that keeps track of all the actions ever taken on Toolhub (when a user joined, a tool was added, removed, etc.). You can think of it as the RecentChanges page on a wiki.

Features

10. API documentation

- OpenAPI Specification information is available for all API endpoints exposed by Toolhub
- RapiDoc UI allows using the APIs directly from your browser for testing and exploration



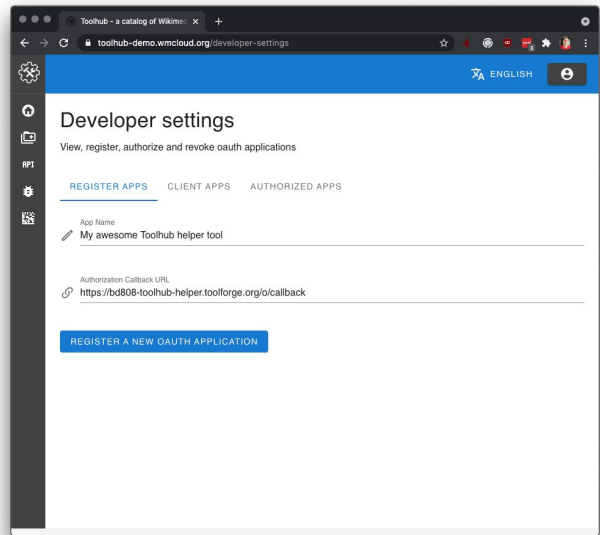
Toolhub has a web API using which you can have access to all of the features needed to power it's integrated UI, and there is nice documentation available for viewing on the platform itself.

Having it API first means it acts as a data repository that can be interacted with and other people can use to build upon. 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 can use to build tools that interact with the tools data stored in Toolhub in many ways.

Features

11. Developer settings

- Self-serve OAuth 2 application registration
- See all registered apps
- Manage your app authorizations



Toolhub's API first design allows you to write your own interface if you want. Our API requires that any create, update, and delete API actions be made by an authenticated user (we will not be allowing anonymous edits). Toolhub provides its own OAuth 2 service to allow you to register a tool that needs to authenticate users so that they can use the API to change content. Currently only confidential web flow clients are supported by Toolhub's OAuth 2 server.

Active work

- Content moderation support
- Curated lists of tools
- Community added information for tools



Content moderation support: This is what we are working on right now. We have edit history, diff, and undo for each individual toolinfo record working already. We are working on UI to show you all edits made by a particular user (like Special:Contributions in MediaWiki) and edits by all users (like Special:RecentChanges in MediaWiki). Suppressing vandal edits and marking edits as reviewed/patrolled are being worked on in the backend.

Curated lists of tools: We plan on adding several UI features based on curated lists. We want private "favorites" lists to let you easily keep track of your most often used tools. We also want to let you make themed lists of tools to share with everyone. And we want to have "featured" lists which will probably just be an extra flag that can be set on a themed list to make it show up in on the home screen. Think of this a something similar to featured articles on many of our wikis.

Community added information for tools: This feature is called "annotations" on our roadmap. The core data from a toolinfo

record has a single owner either the URL that the crawler collected it from, or the user who created it via the API. This is important to preserve the "authority" of things like this software license associated with the tool. We want community members to be able to add other information like categories, links to community specific documentation, tutorials, screen shots, and videos. We hope this provides a good balance between the developer control of information from the Hay's Directory standard and the desire for community to help improve documentation.

Help us make this awesome!

- Play with the demo server at <https://toolhub-demo.wmcloud.org/>
- 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



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 play with our demo server, help translate the user interface, and report bugs and feature ideas. Follow our progress each week on meta, and stay tuned for more direct calls for help as we near our target launch date in August 2021.

Credits

- [Toolhub community logo.svg](#) By [BDavis \(WMF\)](#), [CC0](#)
- [Phabricator Sticker 2.png](#) By [Waffles51](#), [CC BY-SA 4.0](#)
- Persona emojis by [Noto project](#), [Apache license 2.0](#)
- Toolhub screenshots By [BDavis \(WMF\)](#), [CC BY-SA 4.0](#)



Copyright © 2021, [Bryan Davis](#), [Srishti Sethi](#) and the [Wikimedia Foundation](#).
This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).

