

An Introduction to SuggestBot

Morten Warncke-Wang – User:Nettrom – 2022-05-21

User:Nettrom

- Staff Data Scientist, Product Analytics, WMF
 - Supporting the Growth Team
- Wikipedia researcher
 - Content quality
 - Machine learning
- SuggestBot maintainer since 2010



We will talk about...

- What different features does the bot offer
 - Updating open tasks
 - Article recommendations
- How to enable it on a wiki
 - Hosting on Toolforge & Cloud VPS
- What will it take to monitor the bot and ensure it is working as expected?
- What does the bot's emergency response procedure look like?

Part 1: Open Tasks

Википедия:Сообщество - Вики

← → ↻ <https://ru.wikipedia.org/wiki/Википедия:Сообщество> 📄 ☆ 🏠 ☰

Обустройстваем русскую Википедию

Википедия каждый день нуждается в выполнении рутинных задач, наполнении новыми материалами и исправлении ошибок. Выберите задачу себе по вкусу и подключайтесь к работе. Если вас больше всего интересует создание статей, но у вас нет опыта, — попробуйте начать статью в [Инкубаторе](#), там вам окажут необходимую помощь.

Чем помочь

Это список первоочередных задач, на исправление которых у сообщества уходит много времени. Он регулярно обновляется — возвращайтесь, чтобы посмотреть новые задачи. Чтобы узнать детали выполнения каждой задачи, нажмите ? рядом с соответствующим заголовком.

<h4>Проставить категории ?</h4> <ul style="list-style-type: none">Proton Saga (третье поколение)Асымбекова, Гульнара УметовнаСорочкин, Владимир ЕвгеньевичЭрик ЧантурияВторичный гиперальдостеронизм <p>и ещё 172 статьи</p>	<h4>Уточнить шаблон Викификации ?</h4> <ul style="list-style-type: none">Flush (песня)Anthology 2Уркиса, Хусто Хосе деНикотиновый ацетилхолиновый рецепторКузева, София <p>и ещё 15 192 статей</p>	<h4>Обновить информацию ?</h4> <ul style="list-style-type: none">Ветроэлектростанция УфимцеваЭстарВолочаевская (станция метро)Эль-КуфраBA 330 <p>и ещё 6477 статей</p>
<h4>Расширить короткие статьи ?</h4> <ul style="list-style-type: none">Нижний БараАнтоновское сельское поселение (Волгоградская область)Delivered Ex QuayБжовский, ЮзефМинамото-но Ёсиясу <p>и ещё 13 666 статей</p>	<h4>Проверить и добавить источники^[1] ?</h4> <ul style="list-style-type: none">Викинг (ракета)Окубо, РумиНазаровский районHenney KilowattЕрмаков, Юрий Александрович <p>и ещё 6432 статьи</p>	<h4>Проверить на наличие оригинального исследования ?</h4> <ul style="list-style-type: none">Евразийские кочевникиДхармендраСмартбукРайонное интернет-сообществоСмысл жизни <p>и ещё 167 статей</p>
<h4>Улучшить преамбулу ?</h4> <ul style="list-style-type: none">Йессурун де Мескита, СамуэльАрнёмНеймегенАвтоматическое получение размеченного корпуса	<h4>Добавить вики-ссылки на статью ?</h4> <ul style="list-style-type: none">Кондрацкая, Ольга ИгоревнаТаранов, Иван ИльичФейзиев Гасан Кулы оглыБромид трис(этилендиамин)кобальта(III)	<h4>Перевести и почистить ?</h4> <ul style="list-style-type: none">Британская система наградБольшой вывернутый обратноплосконосый икосододекаэдрTata HexaКривая Госпера

How does that work?



WIKIPEDIA
The Free Encyclopedia

[Nettrom](#)
[Talk](#)
[Sandbox](#)
[Preferences](#)
[Beta](#)
[Watchlist](#)
[Contributions](#)
[Log out](#)

[Category](#)
[Talk](#)

[Read](#)
[Edit](#)
[Edit source](#)
[View history](#)

[Help](#)

Category:Wikipedia articles needing copy edit

From Wikipedia, the free encyclopedia


 The May backlog-reduction drive, focusing on our oldest backlog months, and February and March requests, is underway. [Sign up now!](#)

Our [one-week April copyediting blitz](#) has ended, and barnstars have been distributed. Thanks to all who participated!


 This page is currently receiving attention from the [Guild of Copy Editors](#). Information and assistance can be found at the [project page](#) and our [talk page](#). Anyone who copy edits articles, or wishes to start, is welcome to join the project – all help is appreciated!



[Update this page](#)


 This is a **maintenance category**, used for [maintenance of the Wikipedia project](#). It is not part of the encyclopedia and contains [non-article pages](#), or groups articles by status rather than subject. Do not include this category in content categories.

This is a **hidden category**. It is not shown on its member pages, unless the corresponding [user preference](#) (appearance → show hidden categories) is set.

!

Administrators: Please do not delete this category as empty!

This category may be empty occasionally or even most of the time.

Wikipedia articles needing copy edit

(refresh)

Subtotals	[show]
All articles	1,283

[Random subcategory](#)

For more information, see [Wikipedia:Basic copyediting](#).

This **category** is for articles tagged with `{{copy edit}}`, as well as `{{awkward}}`, `{{copy edit inline}}`, `{{copy edit section}}`, `{{inappropriate person}}`, `{{cleanup tense}}`, and their redirects. To add an article to this category, place `{{copy edit|for=}}` at the top of it. A request for [copy editing](#) is most effective when the article is more-or-less complete, and fully cited.

This category, or its monthly subcategories, should not be added to an article directly; use one of the above templates instead.

You can help in this process either by copy editing articles from the months listed below, by choosing an article from the list at [Opentask](#), or by picking a random article of interest and copy editing it right now:

[Pages in this category sorted by number of page views last month](#)

[Random page in category All articles needing copy edit](#)

Shortcuts

[CAT:COPY](#)

[CAT:CE](#)

[Print/export](#)
[Download as PDF](#)
[Printable version](#)

[Languages](#)

[العربية](#)
[বাংলা](#)
[Башҡортса](#)
[भोजपुरी](#)
[Català](#)
[Cebuano](#)
[Español](#)

Open Tasks

- The wiki has categories of articles needing work
- The bot grabs a few articles from each category
- Puts them in lists on a specific page
 - English, Russian: sub-page of the Community Portal, which is transcluded as a template
 - Polish: a specific page in the Wikipedia namespace, used as needed

Source code:

<https://github.com/nettrom/suggestbot/blob/master/opentasks/opentasks.py>

Per-wiki configuration, for example:

<https://github.com/nettrom/suggestbot/blob/master/opentasks/enwiki.yaml>

Wikipedia:Community portal/Open tasks

Project page **Talk** Read Edit source View history More Search Wikipedia

Wikipedia:Community portal/Open tasks

From Wikipedia, the free encyclopedia

Wikipedia maintenance
Task Center (open tasks) · Backlog · Dashboard · Database reports
Special pages (new pages) · Category tracker · WikiProjects (resources)

This page is automatically updated by SuggestBot. Feel free to modify the page layout, if you have consensus, but do not delete or rename the <div id="></div> tags, as they assist the bot in placing the recalculated open task lists. Get in touch with the bot's owner if you have any questions.

You can help improve the articles listed below! This list updates frequently, so check back here for more tasks to try. (See [Wikipedia:Task Center](#) for brief guides.)

Fix spelling and grammar	Fix wikilinks	Update with new information
<ul style="list-style-type: none">List of Kamen Rider Ex-Aid charactersJames Webb Space Telescope sunshield	<ul style="list-style-type: none">Notre Dame Fighting Irish footballEuropean Dressage Championships	<ul style="list-style-type: none">Daewon C.I.Johanna Moore

Editing Wikipedia:Community po X

https://en.wikipedia.org/w/index.php?title=Wikipedia:Community_portal/Open_tasks&action=edit

Paragraph A Cite Insert ? Publish changes...

```
<div id="copyedit">
* [[List of Kamen Rider Ex-Aid characters]]
* [[James Webb Space Telescope sunshield]]
* [[Forensic Heroes]]
* [[Salman Raduyev]]
* [[Las Vegas Sands]]</div>

<div style="margin: .2em 1.6em .6em; font-weight: bold;"><!--
-->'[[Category:Wikipedia articles needing copy edit|More...]]' • [[Wikipedia:Basic copyediting|Learn how]]
</div>
</div>
```

W Editing Wikipedia:Community po X +

https://en.wikipedia.org/w/index.php?title=Wikipedia:Community_portal/Open_tasks&action=edit

Paragraph A Cite Insert ? !

Publish changes...

```
<div id="copyedit">
* [[List of Kamen Ri
* [[James Webb Space
* [[Forensic Heroes]]
* [[Salman Raduyev]]
* [[Las Vegas Sands]]
<div style="margin:
-->'[[[:Category:Wi
</div>
</div>
```

suggestbot/enwiki.yaml at mas x +

github.com/nettrom/suggestbot/blob/master/opentasks/enwiki.yaml#L53

```
37 #
38 # For items 2, 3, and 4 above, articles in the categories are combined into
39 # a set before articles are randomly picked.
40 #
41 # If the key is "stub", then the code follows item 3 above, but sub-categories
42 # are picked at random until we have enough articles. This is done to support
43 # the English Wikipedia's large number of stubs (and stub categories).
44
45 tasks:
46   wikify:
47     - "All dead-end pages"
48     - "All articles with too few wikilinks"
49     - "All articles with too many wikilinks"
50   leadcleanup:
51     - "use-subs"
52     - "Wikipedia introduction cleanup"
53   copyedit: "All articles needing copy edit"
54   update: "All Wikipedia articles in need of updating"
55   translate: "Wikipedia articles needing cleanup after translation"
56   verify: "All pages needing factual verification"
57   or: "All articles that may contain original research"
58   stub: "Stub categories"
```

Wikipedia:Community portal - v. x +

https://en.wikipedia.org/wiki/Wikipedia:Community_portal

Wikivoyage
Wiktionary

Languages

Aquraḥoo
Afrikaans
Alemannisch
Amlaif Yau
English
العربية
Aragonés
Armáneashtí
অসমীয়া
Asturianu
Атқамекиw
Аварга
Aaap
Azərbaycanca
Basa Bali
Беларуская
Bân-lâm-gú
Basa Banyuwasan
Башҡортса
Беларуская
Беларуская
(тарашкевіца)
বাংলা
Български
Boarisch
Bosanski
Brezhoneg
Català

Help out

You can help improve the articles listed below! This list updates frequently, so check back here for more tasks to try. (See [Wikipedia:Task Center](#) for brief guides.)

Fix spelling and grammar	Fix wikilinks	Update with new information
<ul style="list-style-type: none"> List of Kamen Rider Ex-Aid characters James Webb Space Telescope sunshield Forensic Heroes Salman Raduyev Las Vegas Sands <p>More... · Learn how</p>	<ul style="list-style-type: none"> Notre Dame Fighting Irish football European Dressage Championships 2008–09 in English football Iran national futsal team List of Morphoses productions <p>More... · Learn how</p>	<ul style="list-style-type: none"> Daewon C.I. Johanna Moore Hikvision Crotalus stejnegeri North American integration <p>More... · Learn how</p>
Expand short articles	Check and add references	Fix original research issues
<ul style="list-style-type: none"> The Advisory Circle Rozzi Plain Ray St. John Henry Dagg Robin Richmond <p>More... · Learn how</p>	<ul style="list-style-type: none"> War of 1812 Folktronica Steph Davis Tokyo Metro 07 series ConsumerLab.com <p>More... · Learn how</p>	<ul style="list-style-type: none"> Toter Theaker Wilder Snohomish High School -30- (The Wire) Queen + Adam Lambert Tour 2014–2015 <p>More... · Learn how</p>
Improve lead sections	Add an image	Translate and clean up
<ul style="list-style-type: none"> Mabel Sharman Crawford Granary Padua, Italy 	<ul style="list-style-type: none"> Ashwood, Victoria Al-Quds Mosque Hamburg Al-Quds Mosque, Hall of Fame 	<ul style="list-style-type: none"> Klootschieten FRÖLINAT AG Negeri Sembilan

Wikipedia:Community portal/Open tasks

https://en.wikipedia.org/wiki/Wikipedia:Community_portal/Open_tasks

Netrom

Talk Sandbox Preferences Beta Watchlist Contributions Log out

Project page [Talk](#) [Read](#) [Edit source](#) [View history](#) [More](#) Search Wikipedia

Wikipedia:Community portal/Open tasks

From Wikipedia, the free encyclopedia
< Wikipedia:Community portal

Wikipedia maintenance

[Task Center \(open tasks\)](#) · [Backlog](#) · [Dashboard](#) · [Database reports](#)
[Special pages \(new pages\)](#) · [Category tracker](#) · [WikiProjects \(resources\)](#)

This page is automatically updated by [SuggestBot](#). Feel free to modify the page layout, if you have consensus, but do not delete or rename the <div id=""></div> tags, as they assist the bot in placing the recalculated open task lists. Get in touch with [the bot's owner](#) if you have any questions.

You can help improve the articles listed below! This list updates frequently, so check back here for more tasks to try. (See [Wikipedia:Task Center](#) for brief guides.)

Fix spelling and grammar	Fix wikilinks	Update with new information
<ul style="list-style-type: none"> List of Kamen Rider Ex-Aid characters James Webb Space Telescope sunshield Forensic Heroes 	<ul style="list-style-type: none"> Notre Dame Fighting Irish football European Dressage Championships 2008–09 in English football 	<ul style="list-style-type: none"> Daewon C.I. Johanna Moore Hikvision

https://en.wikipedia.org/wiki/Wikipedia:Community_portal?veaction=editsource

Paragraph **A** Cite ?

```
<!-- ----- HELP OUT ----- -->
<div class="comportal-section">
<h2>
Help out
</h2> <!-- 3 line spacing of h2 is important for subsection edit links to work correctly, PLEASE DO NOT FIX -->
-->{{Wikipedia:Community portal/Open task}}
</div>
```

How do I get it?

Does your wiki already have it?

- Wikipedia in English, Polish, Russian, and Thai do

Does your wiki have:

- Categories of articles you'd like to list somewhere?
- A bot approval process? (SuggestBot is not a global bot)

Open tasks is run from Toolforge:

- Set it up yourself – Python, Pywikibot, PyMySQL, PyYAML
- Help set up the configuration file and have me run it

Monitoring

- Watchlist the open tasks page
 - Typically we update it once an hour, easy to notice if that's not happening
- Check the log files on Toolforge

If not run yourself, then:

- If it's making bad edits, block it and post on its English talk page
 - This has happened: [https://en.wikipedia.org/wiki/Wikipedia:Lamest_edit_wars#\(Ro\)bot_wars](https://en.wikipedia.org/wiki/Wikipedia:Lamest_edit_wars#(Ro)bot_wars)
- If it's not working, post on its English talk page and I'll follow up

Part 2: Article Recommendations

How do I get one of these?

The screenshot shows a web browser window displaying a Wikipedia page. The address bar shows the URL: <https://fr.wikipedia.org/wiki/Utilisateur:Nettrom/sandbox/suggestions>. The page title is "Utilisateur:Nettrom/sandbox/suggestions". The page content includes a navigation bar with "Page d'accueil", "Page d'utilisateur", "Discussion", "Lire", "Modifier", "Modifier le code", "Voir l'historique", and "Plus". The main heading is "Utilisateur:Nettrom/sandbox/suggestions". Below the heading, there is a section titled "SuggestBot vous propose..." with a link to "[modifier | modifier le code]". The text below reads: "SuggestBot a le sentiment que vous apprécieriez de contribuer aux articles listés ci-dessous. N'hésitez pas !". The page lists several categories of suggestions:

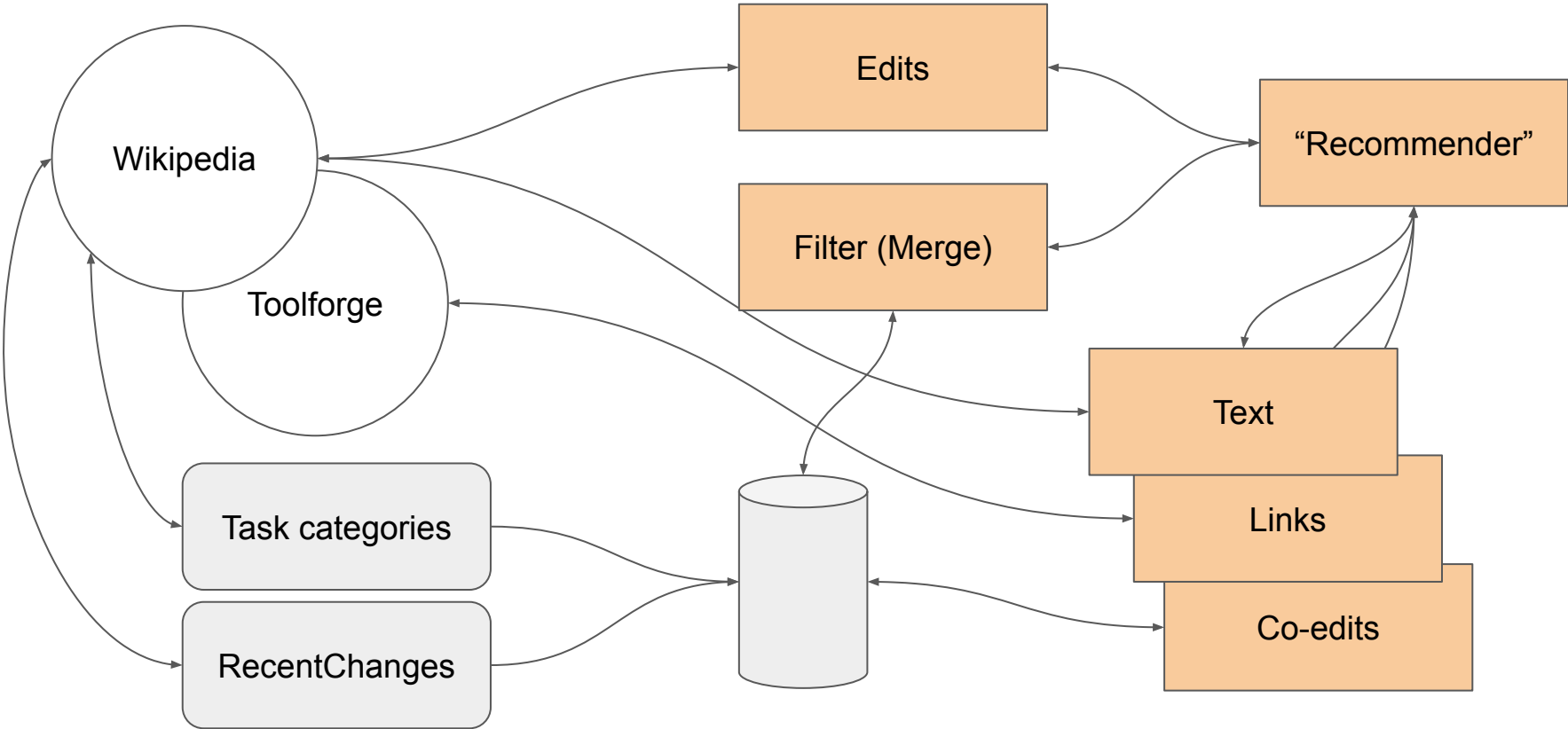
- Manque de sources**: Steve Howe, Guitare baryton, Sonny Burgess, New York Sessions '67, Chronologie du metal, Venom (groupe).
- À wikifier**: God and Guns, Real Live Roadrunning, Allison Wolfe.
- Articles orphelins**: Joy of Cooking, Marie Defrère, Opérateur de Householder.
- À vérifier**: World Team Tennis, The Navigator Company, Estella Blain.
- À mettre à jour**: Rock and Roll Hall of Fame, The Doobie Brothers, Joan Jett.
- Ton non encyclopédique**: Bobby and the Midnites, Bill Bottrell, Twisted Sister.
- Ébauches**: Reeves Gabrels, David Lee Roth, Panty Raid, Glenn Hughes, Ovation Guitar Company, Duff McKagan.

At the bottom of the page, there is a paragraph: "SuggestBot propose ses suggestions à partir des articles auxquels vous avez contribué, que ce soit en en comparant le contenu avec celui d'autres articles, en suivant leurs liens internes et en comparant vos contributions avec celles d'autres wikipédiens. SuggestBot tend à ne proposer que des articles identifiés comme nécessitant des améliorations. Vos contributions sont utiles ; merci pour votre travail au sein de Wikipédia."

SuggestBot's architecture

Cron jobs

XML-RPC daemons



Requests demo

Subscribers

- Daily cron job
- Iterates through subscription & userbox templates
 - Updates database with subscribers
- Identifies all users due for suggestions
- *Always* uses the user's contributions as seeds

How do I get it?

Does your wiki already have it?

- Wikipedia in English, Norwegian (bokmål), Swedish, French, Persian, and Portuguese already do

Does your wiki have:

- Categories of articles needing improvement?
- A bot approval process? (SuggestBot is not a global bot)
- Help set up and translate configuration files and pages