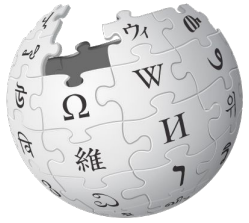
FAIR and open multilingual clinical trials in Wikidata and Wikipedia



WIKIPEDIA

WIKIDATA

Lane Rasberry

Wikimedian in Residence

School of Data Science

University of Virginia

Cherrie Kwok

PhD Candidate & Jefferson Fellow

Department of English

University of Virginia

Thanks to contributors

- Wiki Project Med Foundation
- all Wikidata volunteers
- Duke University, Sheri Tibbs
- ClinWiki
  - William Hoos
  - Amy Westermann
- Vanderbilt University
  - Steven James Baskauf
  - Clifford Anderson
  - Philip Walker
- PCORI, Jeffrey Keefer
- Scholia
  - Daniel Mietchen
  - Finn Årup Nielsen
  - Egon Willighagen
- Wikimedia Bangladesh, Tanvir Rahman
- Wellcome Trust

Notes about talk

- This is for wiki-beginners
- Documentation and results on-wiki
- Main challenge of translation:
  collaborate with stakeholders
  to promote diversity, equity, inclusion

Outline of talk

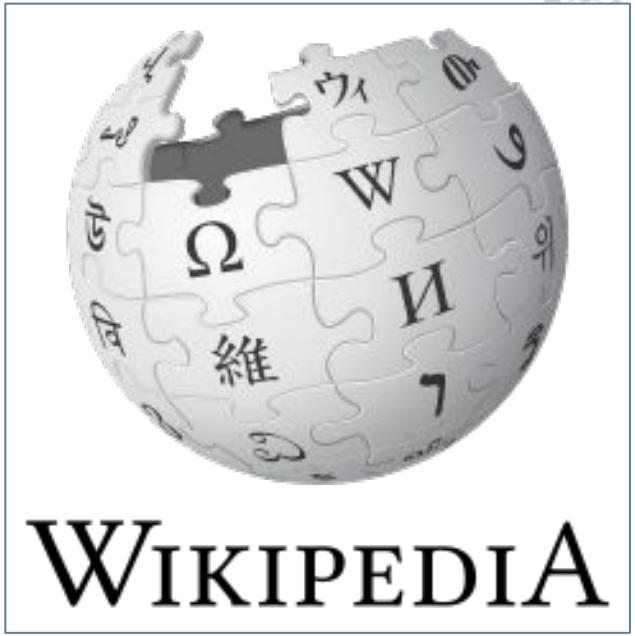
1. Why use Wikipedia?
2. WikiProject Clinical Trials
3. Wikipedia and translation
4. Social and ethical issues

Use Wikipedia to achieve communication goals

Search engines 💙 wiki

Google

Microsoft Bing



Traffic metrics prove popularity

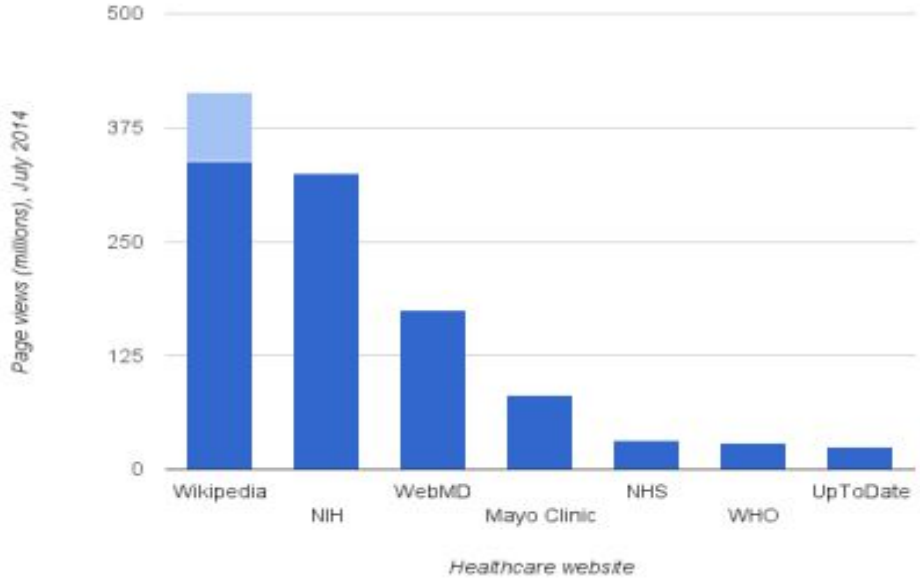
Traffic data is public for all articles in all languages

For example in medicine -

James Heilman, Andrew G. West. "Wikipedia and medicine: quantifying readership, editors, and the significance of natural language". *Journal of Medicine Internet Research*, March 2015.

Wikipedia more popular

than other sources



Clinical trial

From Wikipedia, the free encyclopedia

"Clinical Trials" redirects here. For the journal, see Clinical Trials (journal).

Clinical trials are experiments or observations done in clinical research. Such prospective biomedical or behavioral research studies on human participants are designed to answer specific questions about biomedical or behavioral interventions, including new treatments (such as novel vaccines, drugs, dietary choices

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute
Help
Learn to edit
Community portal

Clinical trial

From Wikipedia, the free encyclopedia

"Clinical Trials" redirects here. For the journal, see Clinical Trials (journal).

Clinical trials are experiments or observations done in clinical research. Such prospective biomedical or behavioral research studies on human participants are designed to answer specific questions about biomedical or behavioral interventions, including new treatments (such as novel vaccines, drugs, dietary choices

WIKIPEDIA
The Free Encyclopedia

Clinical trial: Revision history

View logs for this page (view filter log)

Filter revisions

External tools: Find addition/removal (Alternate) · Find edits by user (Alternate) · Page statistics · Pageviews links

For any version listed below, click on its date to view it.

For more help, see Help:Page history and Help:Edit summary. (cur) = difference from current version, (prev) = difference from preceding version, **m** = minor edit, → = section edit, ← = automatic edit summa

(newest | oldest) View (newer 50 | older 50) (20 | 50 | 100 | 250 | 500)

Compare selected revisions

- (cur | prev) ◉ 18:04, 20 November 2021 Jukeboksi (talk | contribs) **m** . . (95,347 bytes) (0) . . *(depluralizing a link)* (undo)

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute
Help
Learn to edit
Community portal
Recent changes
Upload file

WIKIPEDIA
The Free Encyclopedia

Not logged in Talk Contributions Create acc

Article Talk Read Edit View history Search Wikipedia

Clinical trial: Revision history

View logs for this page (view filter log)

Filter revisions

External tools: Find addition/removal (Alternate) · Find edits by user (Alternate) · Page statistics · Pageviews links

For any version listed below, click on its date to view it.
For more help, see Help:Page history and Help:Edit summary. (cur) = difference from current version, (prev) = difference from preceding version, **m** = minor edit, → = section edit, ← = automatic edit summa
(newest | oldest) View (newer 50 | older 50) (20 | 50 | 100 | 250 | 500)

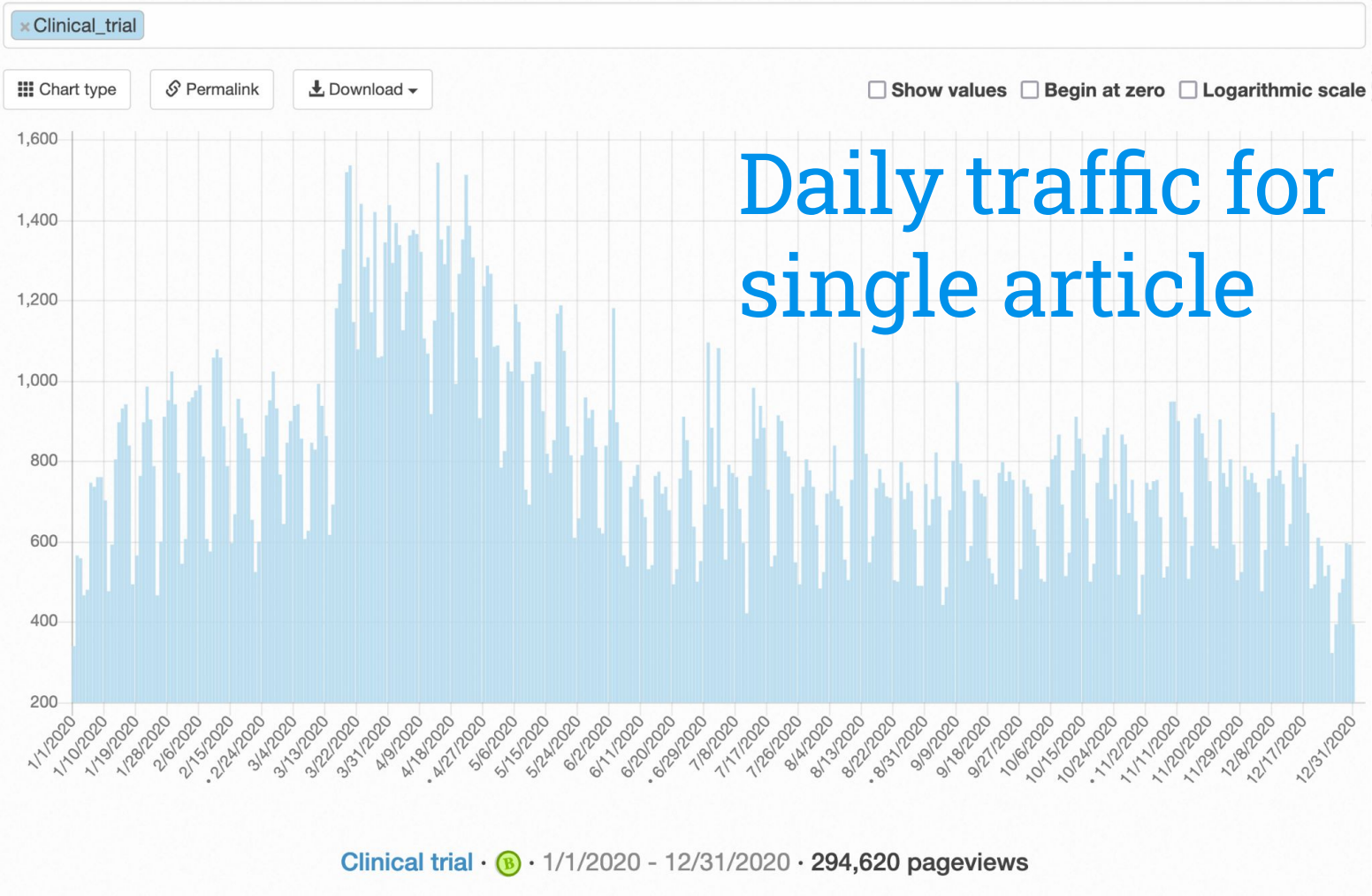
Compare selected revisions

- (cur | prev) 18:04, 20 November 2021 Jukeboksi (talk | contribs) **m** . . (95,347 bytes) (0) . . *(depluralizing a link)* (undo)

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute

Help
Learn to edit
Community portal
Recent changes
Upload file



Daily traffic for single article

Yearly traffic, single article, all languages

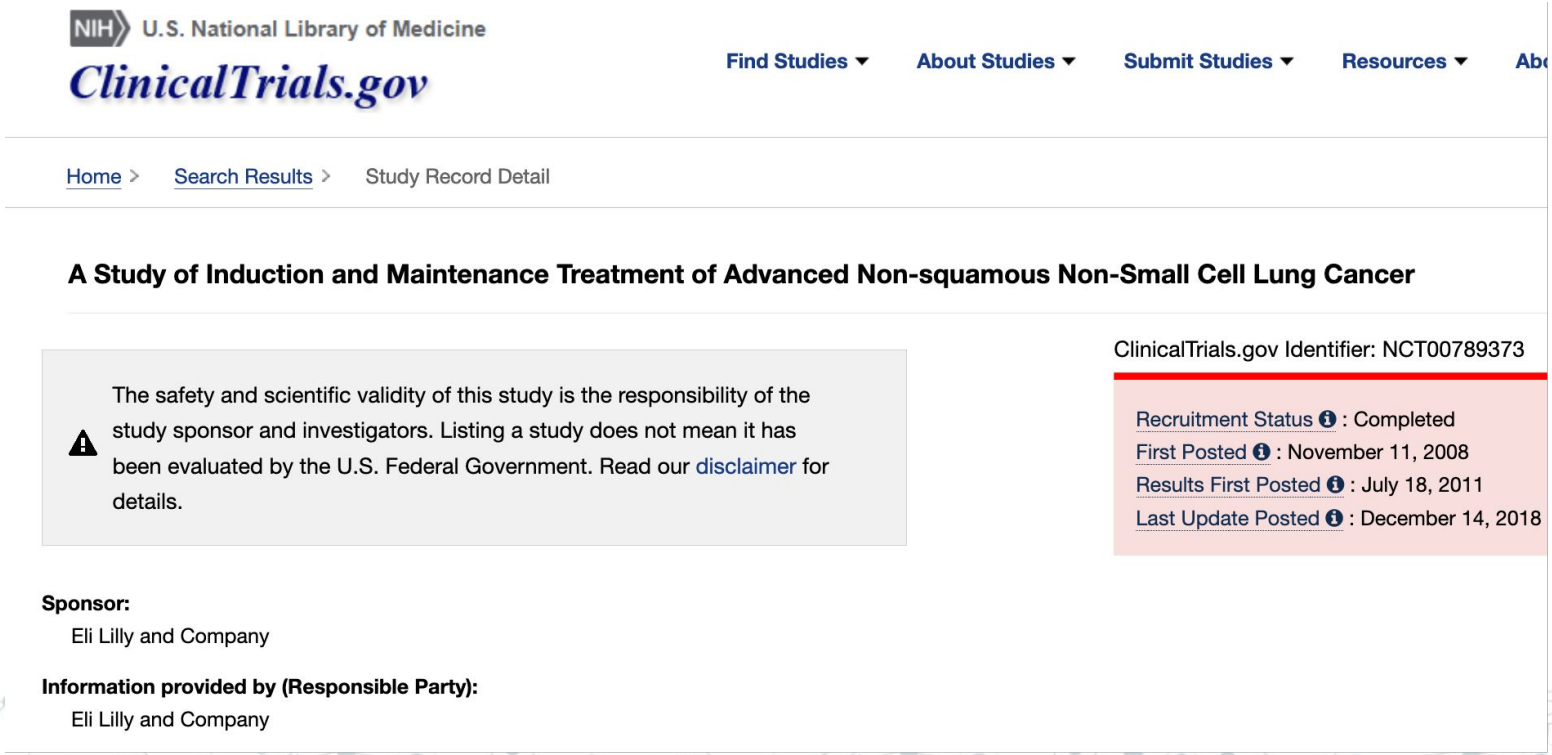
Clinical trial 1/1/2020 - 12/31/2020

Permalink | Download

#	Language	Page title	Badges	Pageviews
Totals	**40 languages**	**37 unique titles**		**909,401**
1	en	Clinical trial		294,620
2	ru	Клиническое исследование		91,552
3	de	Klinische Studie		75,057
4	ja	治験		71,004
5	fr	Essai clinique		65,158
6	es	Ensayo clínico		56,040
7	it	Studio clinico		49,279
8	zh	臨床試験		33,367
9	ar	تجربة سريرية		27,308
10	ko	임상시험		23,722

We imported ClinicalTrials.gov to Wikidata and curated it wiki style

Sample clinical trial



NIH U.S. National Library of Medicine

ClinicalTrials.gov

Find Studies ▾ About Studies ▾ Submit Studies ▾ Resources ▾ Abo

Home > Search Results > Study Record Detail

A Study of Induction and Maintenance Treatment of Advanced Non-squamous Non-Small Cell Lung Cancer

ClinicalTrials.gov Identifier: NCT00789373

⚠ The safety and scientific validity of this study is the responsibility of the study sponsor and investigators. Listing a study does not mean it has been evaluated by the U.S. Federal Government. Read our disclaimer for details.

Recruitment Status ⓘ : Completed
First Posted ⓘ : November 11, 2008
Results First Posted ⓘ : July 18, 2011
Last Update Posted ⓘ : December 14, 2018

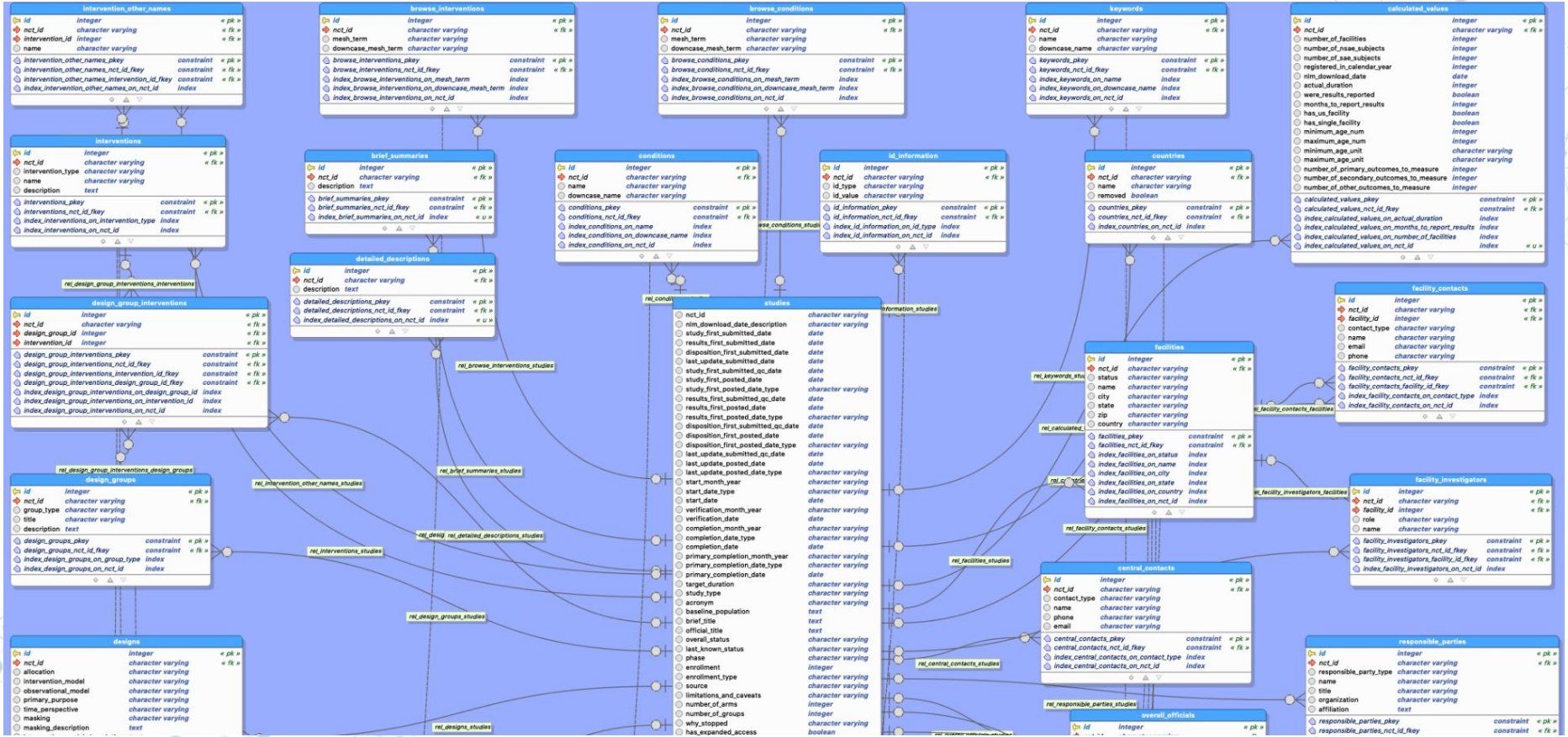
Sponsor:
Eli Lilly and Company

Information provided by (Responsible Party):
Eli Lilly and Company

Basic information

- disease treated
- drug used
- who managed
- sponsors
- research sites
- when
- how many people
- phase
- status
- catalog numbers

Data schema


intervention_other_names
browse_interventions
browse_conditions
keywords
calculated_values
interventions
brief_summaries
conditions
id_information
countries
detailed_descriptions
studies
facility_contacts
design_group_interventions
facilities
design_groups
facility_investigators
central_contacts
designs
responsible_parties
overall_officials

Project goals

summary: put data in wiki and make it useful

1. Import to Wikidata
2. Design use cases
3. Query Wikidata
4. Curate query results
5. Showcase output

Project goals

summary: put data in wiki and make it useful

1. Import to Wikidata - COVID research
2. Design use cases - university profile
3. Query Wikidata - "show my school's research"
4. Curate query results - cleanup, translation
5. Showcase output - data visualization

Project outcomes

Did Wikidata import and curation

- 356,831 clinical trial records
- 356,536 - links to ClinicalTrials.gov
- 177,420 - list phase of research
- applied other labels for query

Modelled our process

- set up WikiProject
- organized community
- documented method
- organized translation

Wikipedia has tools for translation

Various tools support translation

WikiProject Medicine Translation Task Force

Social group with documentation leaderboard

Programs & Events Dashboard

group collaboration tool with multilingual tracking

Wikipedia Content Translation Tool

interface providing access to Google translations

TABernacle

multilingual Wikidata Query Service frontend

WikiProject Medicine

Translation Task Force model:

- make a list of priority articles
- gather people to translate
- track progress

Leaderboard

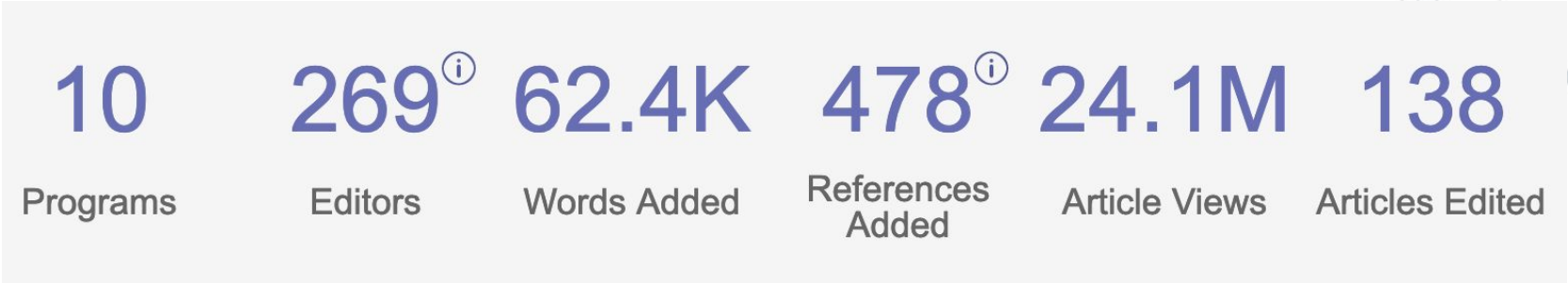
Top users by number of translations

Top languages by number of Articles

List of essential medicines

- Acetylsalicylic acid
- Ibuprofen
- Paracetamol
- Codeine
- Fentanyl
- Morphine
- Methadone
- Amitriptyline
- Cyclizine
- Dexamethasone
- Diazepam
- Docusate sodium
- Fluoxetine
- Amphotericin B
- Miltefosine
- Paromomycin
- Amikacin
- Amoxicillin
- Amoxicillin/clavulanic acid
- Ampicillin
- Benzathine benzylpenicillin
- Benzylpenicillin
- Cefalexin
- Cefazolin
- Chloramphenicol
- Clindamycin
- Cloxacillin
- Doxycycline
- Gentamicin
- Metronidazole
- Nitrofurantoin
- Atazanavir/ritonavir
- Darunavir
- Lopinavir/ritonavir
- Ritonavir
- Bedaquiline
- Clofaziminеα
- Cycloserine
- Delamanid
- Ethionamide
- Levofloxacin
- Linezolid
- Meropenem
- Moxifloxacin
- P-aminosalicylic acid
- Streptomycin
- Antifungal medicines
- Amphotericin B
- Clotrimazole
- Fluconazole
- Flucytosine
- Griseofulvin
- Doxycycline
- Mefloquine

Programs and Events Dashboard



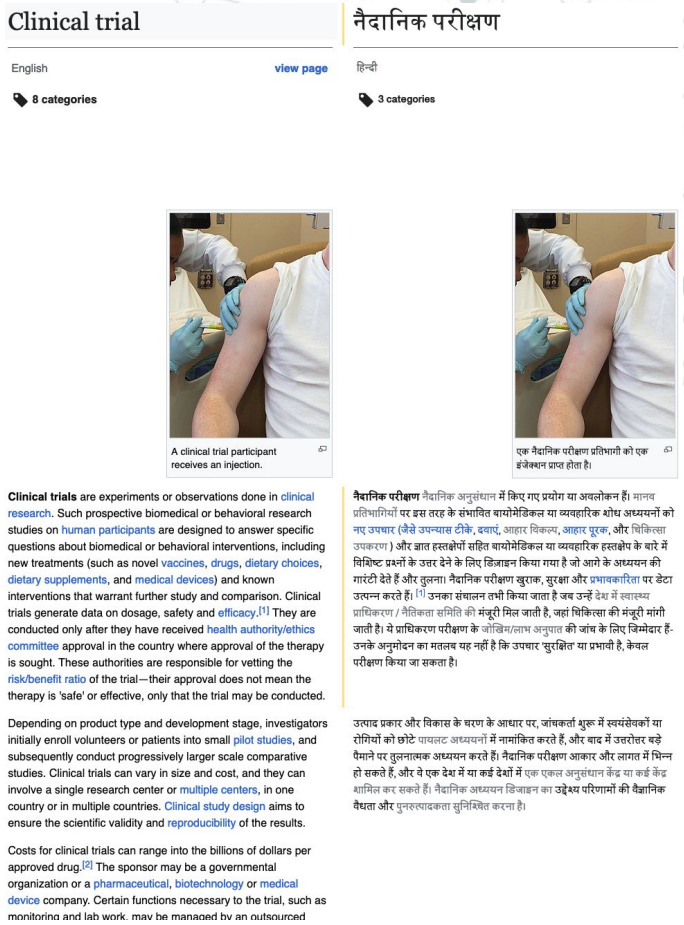
Tool reports wiki edits of a group of people

Tracks all language Wikipedia versions

Wikipedia Content Translation Tool

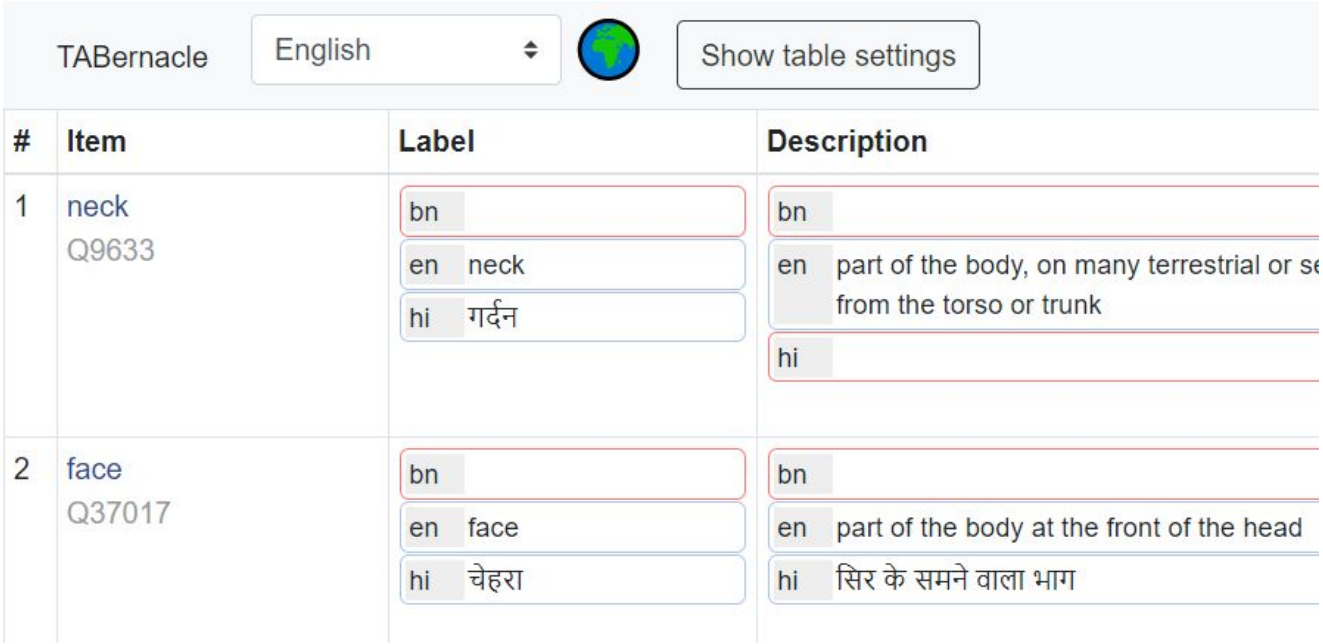
Start with any Wikipedia article

Get interface for translating text side by side



TABernacle

Translate Wikidata terms



Wikipedia has social processes for translation

New York City translation

Multilingual community sharing across cultures

WIKIMEDIA NEW YORK CITY

CUNY THE CITY UNIVERSITY OF NEW YORK

public schools, especially Laguardia Community College

community organizations, including Sure We Can recycling center

art museums, including The Metropolitan Museum of Art

Social and ethical issues

General Wikipedia ethical issues

- Should I use Wikipedia?
- Is my content good enough?
- Who are these other people?
- What if other people edit?

Social and ethical issues

General Wikipedia ethical issues

- Should I use Wikipedia?
- Is my content good enough?
- Who are these other people?
- What if other people edit?

We have to focus on translation

Social and ethical issues for translating in Wikipedia

- practicing transparency
- diversity, equity, inclusion

Documentation for transparency

Wikidata:WikiProject Clinical Trials

Home | Models | Queries | Members | About | Discuss

WikiProject Clinical Trials is a Wikidata community project to curate Wikidata's collection of information on clinical trials.

Contents [hide]

1 Typical activities

2 Objectives

Most popular WikiProjects

1. Contents
2. Military history
3. Spam
4. Medicine
5. Video games
6. Mathematics
7. LGBT studies
8. Football
9. Articles for creation
10. Film

content-focused projects (underlined)
have translation programs

WikiProjects are collaboration hubs

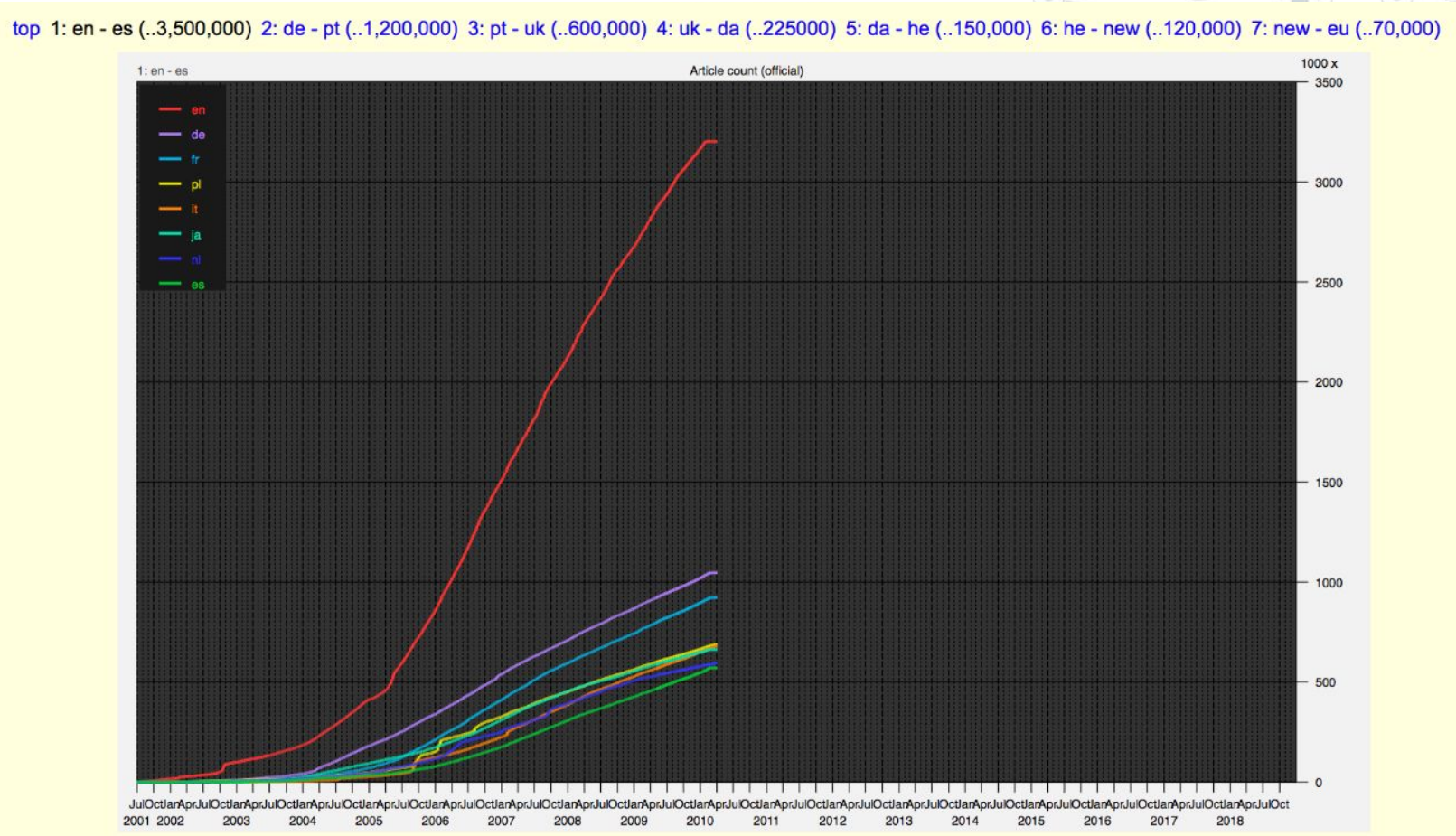
WikiProject Clinical Trials
is modelled after
WikiProject Medicine

- all discussion public
- reporting encouraged
- revisions tracked
- open copyright license
- permanent archiving

Diversity, equity, inclusion

Besides subject matter Wikipedia communities, there are also regional and language communities.

We translated clinical trials data into Bangla, Hindi, and Swahili.



Source: “Wikipedia Article Count by Language 2001-2018,” *Wikimedia.* (https://stats.wikimedia.org/EN/PlotsSvgArticlesTotal.htm)

English:	350 million native speakers
Mandarin Chinese:	1.3 billion native speakers
Spanish:	471 million native speakers
Hindi:	342 million native speakers
Arabic:	315 million native speakers
Bengali:	229 million native speakers

2010: The Case of “Makmende”

“Wikipedia admins then repeatedly deleted the [Makmende] page, initially on ‘criteria for speedy deletion’ G1 (‘Patent nonsense, meaningless, or incomprehensible’), then G12 (‘Unambiguous copyright infringement)’, and finally G3 (‘Pure Vandalism’). Wikipedia editors argued for deletion because there existed ‘no reliable sources, and no claims of notability’ . . . actually, it was quite well sourced, including a reference to a *Wall Street Journal* online publication and five weblogs . . . If Wikipedia wants to . . . address issues of systemic bias – the community needs to expand to include more Wikipedians from the developing world. Deleting three versions of an article important to Kenyans and trying to delete a fourth doesn’t send a strong message that Wikipedia is the open and welcoming community you and I both want it to be.” (259-260)

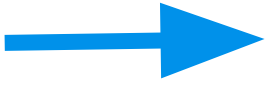
Source: “The Missing Wikipedians,” Heather Ford, *Critical Point of View: A Wikipedia Reader*

“Wikispace and the Politics of Exclusion”

“Wikipedia, therefore, is characterized by distinct directions in which information can be transmitted” (275)

Source: “Wikispace and the Politics of Exclusion,” Mark Graham, *Critical Point of View: A Wikipedia Reader*

বাংলা



- translation of English terms into Bangla
- For Wikidata, also migrates into Wikipedia
- access to information for everyone

Thanks

Tanvir Rahman

Ways to seek community participation

- document project
- give many notices
- make friends
- hire native speaker
- join meetings
- practice diversity
- use wiki tools
- partnership over years
- collaboration exchange
- give credit
- share funding
- many small projects
- equity in design planning
- use wiki social precedent

There is no one way to collaborate

Summary translation results

- We imported clinical trial data to Wikidata
- We collaborated with native speakers to translate it
- Wikidata now can return results in more languages

Thanks!

Lane Rasberry

Wikimedian in Residence

School of Data Science

University of Virginia

Cherrie Kwok

PhD Candidate & Jefferson Fellow

Department of English

University of Virginia