

wikipedia

4/4 INTERFACE

Wendy Ho '21 & Hien Pham '23

WHERE CITATIONS COME FROM:

CITOGENESIS STEP #1:
THROUGH A CONVOLUTED PROCESS,
A USER'S BRAIN GENERATES FACTS.
THESE ARE TYPED INTO WIKIPEDIA.

THE "SCROLL LOCK" KEY WAS
DESIGNED BY FUTURE
ENERGY SECRETARY STEVEN
CHU IN A COLLEGE PROJECT.



A RUSHED WRITER CHECKS WIKIPEDIA
FOR A SUMMARY OF THEIR SUBJECT.

US ENERGY SECRETARY STEVEN CHU,
(NOBEL PRIZEWINNER AND CREATOR OF
THE UBIQUITOUS "SCROLL LOCK" KEY)
TESTIFIED BEFORE CONGRESS TODAY...



STEP #2

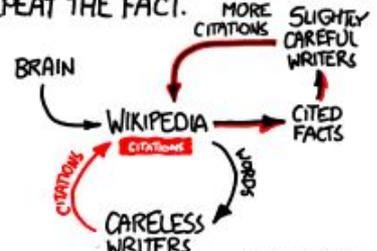
SURPRISED READERS CHECK WIKIPEDIA,
SEE THE CLAIM, AND FLAG IT FOR REVIEW.
A PASSING EDITOR FINDS THE
PIECE AND ADDS IT AS A CITATION.

GOOGLE IS YOUR
FRIEND, PEOPLE.
<REF>{{CITE WEB|URL=



STEP #3

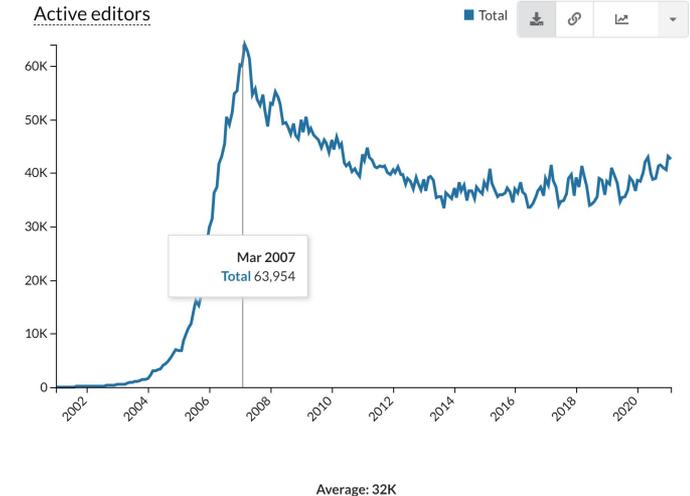
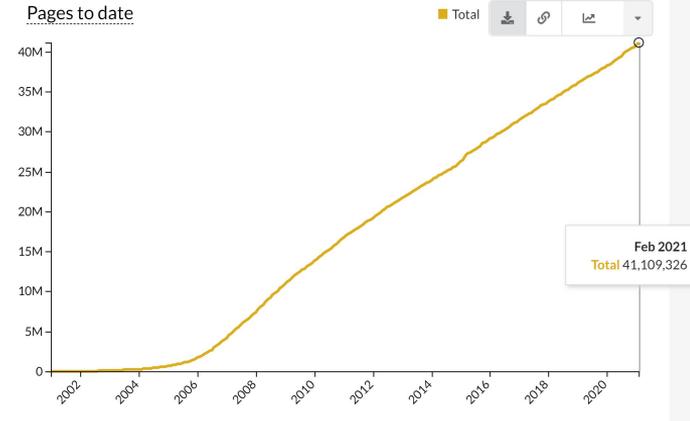
STEP #4
NOW THAT OTHER WRITERS
HAVE A REAL SOURCE, THEY
REPEAT THE FACT.



REFERENCES
PROLIFERATE, COMPLETING
THE CITOGENESIS PROCESS.

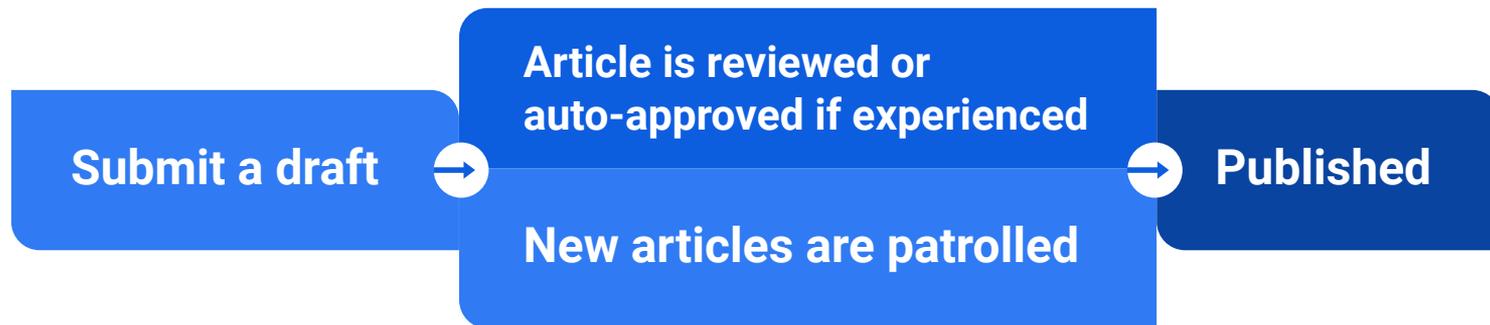
History of Wikipedia

- Launched in 2001, under non-profit WikiMedia
- An encyclopedia using the wiki system:
 - “A wiki **invites all users**—not just experts—to edit any page or to create new pages within the wiki Web site.
 - Wiki promotes meaningful topic associations between different pages by making page link creation intuitively easy and showing whether an intended target page exists or not.”
- In 2009 there was a significant drop of editors (due to stricter “rules” & less low-hanging fruits), and it’s been stable since 2012



How is a Wikipedia page made?

- Theory: Anyone can write new articles!! (true in 2007)
- Reality: Peer-reviewed journal style



How are Wikipedia pages maintained?

- Anyone can edit! But new edits are also heavily patrolled (& bot checked)
- That said, a 2015 experiment had 19/30 well-formatted vandalism undetected for 2 months - these were not on pop culture articles

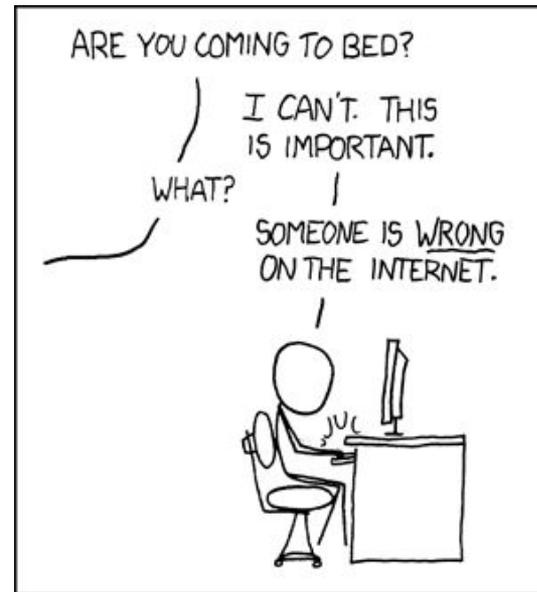
You can help improve the articles listed below! This list updates frequently, so check back here for more tasks to try. (See [Wikipedia:Task Center](#) for brief guides.)

<p>Fix spelling and grammar</p> <ul style="list-style-type: none"> • Zindagi Aa Raha Hoon Main • Sopra Steria • Alexander and Viktor Zhizhich • Ujjwala Majumdar • Kannukku Kannaga <p><i>More...</i> • Learn how</p>	<p>Fix wikilinks</p> <ul style="list-style-type: none"> • Social Reconciliation Reform and Development Party • Parliament of the Azerbaijan Democratic Republic • Minibond • International Missionary Society • Mahadebsinan Hill <p><i>More...</i> • Learn how</p>	<p>Update with new information</p> <ul style="list-style-type: none"> • Canadian Institute of Ecology and Evolution • Seppe Smits • List of Danish-language radio stations • Industrial Bank (China) • Communist Party of Peru (Marxist–Leninist) <p><i>More...</i> • Learn how</p>
<p>Expand short articles</p> <ul style="list-style-type: none"> • Mauriel Carty • Chesney Hughes • Shinelle Proctor • Eustace Proctor • Kieron Rogers <p><i>More...</i> • Learn how</p>	<p>Check and add references</p> <ul style="list-style-type: none"> • Controlled vocabulary • Jason Greene (Freckle) • Training to failure • Chris Johnson (presenter) • Mannerism <p><i>More...</i> • Learn how</p>	<p>Fix original research issues</p> <ul style="list-style-type: none"> • Herbert von Karajan • International Socialist Movement • Urban wilderness • Northern and southern China • List of Macross Delta characters <p><i>More...</i> • Learn how</p>
<p>Improve lead sections</p> <ul style="list-style-type: none"> • Diliman Commune • Fruit hat • Joseph Stroud • Codex Hermogenianus • Euroleague 2003–04 Top 16 Group C <p><i>More...</i> • Learn how</p>	<p>Add an image</p> <ul style="list-style-type: none"> • O'Shannassy River • Oyster Bay Regional Shoreline • Outer Ring Road System • Otaru University of Commerce • Ozark pudding <p><i>More...</i> • Learn how</p>	<p>Translate and clean up</p> <ul style="list-style-type: none"> • Evergreen International Airlines Flight 17 • Wolfgang Schleidt • Telkom Indonesia • Mutara II Rwigera • 2018 Colombian Census <p><i>More...</i> • Learn how</p>

Help counter systemic bias by creating new articles on important women.

Help improve popular pages, especially those of low quality.

Windows Ink Workspaces



How are Wikipedia pages maintained?

- Articles are also assessed and rated based on their *completeness, usefulness to reader, style and edits* (i.e. infrequently edited articles can be demoted)
- Edits should not be implemented/ reverted more than 3 times, or they go to an Arbitration Committee.
- Edit wars can have a Request for Comment, which means 3rd parties can vote/comment on the war

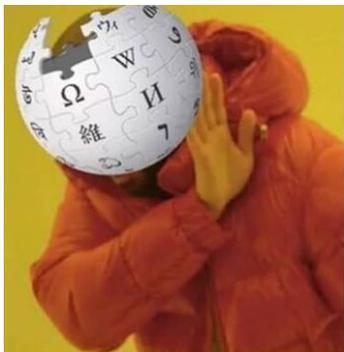
All rated articles by quality and importance						
Quality	Importance					Total
	Top	High	Mid	Low	???	
★ FA	1,373	2,172	2,119	1,481	169	7,314
★ FL	161	596	661	603	104	2,125
ⓘ A	293	603	749	485	87	2,217
⊕ GA	2,670	6,136	12,304	14,807	1,711	37,628
B	14,170	27,368	43,719	43,118	17,324	145,699
C	13,956	43,111	103,445	181,656	67,300	409,468
Start	18,475	86,934	369,069	1,182,365	374,357	2,031,200
Stub	4,471	32,392	274,862	2,434,480	868,275	3,614,480
List	4,054	14,611	44,520	139,789	76,578	279,552
Assessed	59,623	213,923	851,448	3,998,784	1,405,905	6,529,683
Unassessed	118	510	1,967	16,501	441,309	460,405
Total	59,741	214,433	853,415	4,015,285	1,847,214	6,990,088

[About this table](#)

How are Wikipedia pages maintained?

- People wage wars on edits, typically not on accuracy but on presentation/representation of info, including (from Lamest Edit Wars):
 - *Was Chopin Polish, French, Polish–French, or French–Polish?*
 - *(On Caesar salad) Is it spelled Caesar, Cesar, César, or Cesare? - 11 years & going*
 - *In the name of the programming language C#, is that # thing after the C a number sign or the musical sharp symbol? (Microsoft said it's an octothorpe)*
 - *A revert war on whether the tiger can properly be described as the "most powerful living cat" (complete with accusations that people were "tiger fanboys") gradually led to arguments about how tigers would match up vs. bears and crocodiles (oh my!), complete with another revert war about the inclusion of a YouTube video showing a tiger fighting a crocodile, eventually leading to the article being semi-protected. The debates about bears and crocodiles continue on the talk page.*

Meme break



Using ads



Respectfully ask it's users to donate money in order to keep the site running without ads

Wikipedia: Please donate, we would be so thankful

Me: Donates

Refreshes page:



How trustworthy is the info on Wikipedia?

- A few Israeli researchers trained an ML classifier to identify writings & edit patterns by professionals and found that 10-30% of Wikipedia contributors have subject-matter expertise
- Articles that are extremely important, commonly viewed or commonly vandalised gets a semi-protected (blocks editing to new users) or protected status (requires edits to be requested and reviewed).



The screenshot shows the top navigation bar of a Wikipedia article. It includes a user status indicator 'Not logged in' with links for 'Talk', 'Contributions', 'Create account', and 'Log in'. Below this is a row of article actions: 'Article', 'Talk', 'Read', 'View source', and 'View history'. A search box labeled 'Search Wikipedia' is also present. The main title of the article is 'COVID-19'. At the bottom right, a black banner with a lock icon and the text 'This article is semi-protected.' is displayed. The footer text reads 'From Wikipedia, the free encyclopedia'.

How trustworthy is the info on Wikipedia?

- Since vandalism typically involves deleting or modifying content that doesn't follow the style guidelines/has citation... intentional vandalism, sockpuppeting, meatpuppeting, canvassing & conflicts of interests are extremely hard to detect
- Systemic bias oh boy
- Wikipedia: *“The online encyclopedia does not consider itself to be a reliable source and discourages readers from using it in academic or research settings. Researchers, teachers, journalists, and public officials do not regard Wikipedia as a reliable source.”*

Concluding remarks

You now have basic tools for applying linear models to practical settings.

Regularizers can be mixed and matched with different losses (e.g., Logistic regression with L_1 regularization).

The overarching framework for linear models is *Generalized Linear Model* (GLM); sometimes covered in COS424. Or see Wikipedia.

Next time: Applying linear models to problems in other disciplines.

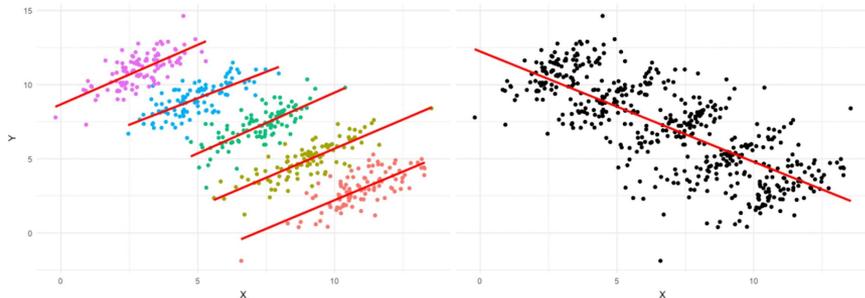
49

COS324

@Tom Griffiths

COS534
@Arvind Narayanan

Simpson's paradox



via Wikipedia

Criticisms of Wikipedia

Systematic bias (especially when it comes to biographies)

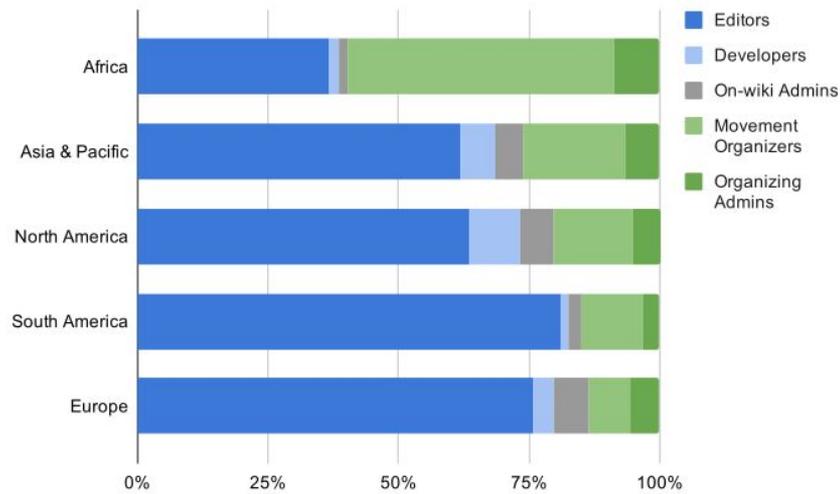
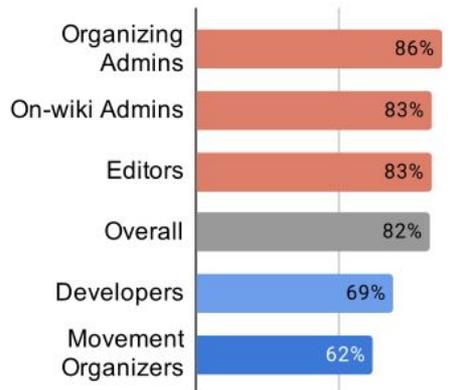
- 'POV pushers' can derail the editing process
- Wikipedia contributors are 82% male
- Almost half live in Europe and a fifth in North America (9.7% and 4.8% of global population respectively)
- Contributors that join in the last two years feel less empowered to succeed if they are women, live in Eastern Asia, or are not fluent in English

Wikipedia articles for a person has to meet the criteria for notability, which, of course, is heavily affected by systemic sexism & racism

My friend Linh wrote this wikipedia article for her class last semester:

https://en.wikipedia.org/wiki/Soraya_Santiago_Solla

% Identifying as Male***



Other crowdsourced encyclopedia attempts

- Uncyclopedia (2005): It's a potato



The screenshot shows the homepage of Uncyclopedia, a crowdsourced encyclopedia. At the top left is a logo featuring a globe made of puzzle pieces with letters, and the text "UNCYCLOPEDIA The Content-Free Encyclopedia". Below the logo is a navigation menu with links like "Main page", "Contents", "Featured content", "Current events", "Random feature", "Random article", "About Uncyclopedia", "Donate", "Shop", "Contribute", "Help", "Community portal", "Recent changes", "Social media", "Votes for Highlight", "Votes for Pictures", "Pee Review", "What's happening", "Discord", "Forums", "Tools", "What links here", "Related changes", "Special pages", "Printable version", and "Permanent link".

At the top right, there is a user status "Not logged in" and links for "Talk", "Contributions", "Create account", and "Log in". Below this is a search bar with the text "Search Uncyclopedia" and a magnifying glass icon. To the left of the search bar are links for "Main Page" and "Talk". To the right are links for "Read", "View source", and "View history".

The main banner features a blue background with a photo of two people in a field. The text reads "Join Uncyclopedia Today!" and "Uncyclopedia is made of people (and intelligent monkeys) like you". A "Create an account" button is on the right.

Below the banner is a "Welcome to Uncyclopedia," section with the text "The content-free encyclopedia anyone can edit" and "36,665 articles in English". To the right of this section is a grid of category links: Arts, Biography, Music, History, Politics, Science, Society, Television, and All portals.

The "From today's featured article" section highlights "Kitten Vaping" with a small image of a person and text describing the act of vaping kittens' souls. Below this is a paragraph about its history, mentioning Dr. Liao Jing and the coronavirus pandemic of 2020. At the bottom of this section are links for "Recently featured: Kitten Vaping · Mona Lisa · Robbie Rotten · Mosque Archive · Vote for featured articles · Purge".

The "In the news" section lists several headlines, including "Stan Marsh presents: 2022 FIFA World Cup sucks ass", "Biden responds to North Korea's 'Stink' complaints", "BMI songwriters (but not ASCAP) to receive vaccine", "Governor Cuomo from New York is screwed", "Grodin Museum reopens to public", "Old movies are suddenly offensive", "Northeast getting another blizzard", "Just another dead doper and a dirtbag dies", "Marilyn Manson removed from New Radicals hit", and "Biden will recognize Hillary as 45th President". To the right of this section is a small portrait of a woman with the caption "Gongheo!".

Other crowdsourced encyclopedia attempts

- Conservapedia (2006): Conservative, family-friendly, Christian. Alternative to Wikipedia since the founders believe that Wikipedia has a liberal bias.

The screenshot shows the Conservapedia website. On the left is a sidebar with navigation links: Main Page, Recent changes, New Pages, Random page, Statistics, donate, and a prominent yellow 'Donate' button. Below these are 'Edit Console' and 'What links here' sections. The main content area is titled 'Featured on Conservapedia.' and includes a banner for 'Over 800,000,000 page views!' and a section for 'Popular articles at Conservapedia' with two columns of bulleted links. On the right, there is a search bar and a 'Log in' link. Below the search bar is a section titled 'In the News, what the MSM isn't fully covering.' which contains several news snippets with links, such as 'Not a Trump supporter...', 'Vaccine passports...', 'Leftist takeover of professional sports continues...', 'Why 'Global Britain' may be a pipe dream.', and 'Experts believe coronavirus variants will make vaccines ineffective within ONE YEAR.'

Other crowdsourced encyclopedia attempts

- Citizendium (2007): Contributors have to be verified with their real identity & have at least a bachelor's degree. Enforce gentle guidance & strict oversight.
 - It's practically dead



Page [Discussion](#) [Read](#) [View source](#) [View history](#)

Welcome to Citizendium

Agriculture	Earth Sciences	Journalism	Physics
Anthropology	Economics	Law	Politics
Archaeology	Education	Library & Info. Sci.	Psychology
Architecture	Engineering	Linguistics	Religion
Astronomy	Food Science	Literature	Robotics
Biology	Games	Mathematics	Sociology
Business	Geography	Media	Sports
Chemistry	Health Sciences	Military	Theater
Classics	History	Music	Topic Informant
Computers	Hobbies	Philosophy	Visual Arts

Help Write Articles about our World

Welcome to **Citizendium**, a wiki for providing free knowledge where authors use their **real names**. We write the kinds of encyclopedia-style articles that Wikipedia and other sites can't write. We welcome anyone who wants to share their knowledge on virtually any subject. Our online community prides itself on being congenial and supportive.

See **Recent Changes**—an overview of articles we are writing *now*.

Become a member—it's free!

- **Join** via our [application form](#) (Google Account required)
- **Start a new article**
- **Improve existing articles**
 - [Quick Start](#) | [How to collaborate](#) | [How to format pages](#)
- Wikipedians are **very** welcome!

Some of our finest

-  **Citable Articles** (154)
- Developed Articles** (1,174)
(17,291 total articles)

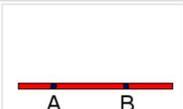
Knowledge is like money: To be of value it must circulate, and in circulating it can increase in quantity and, hopefully, in value.

— Louis L'Amour (1908–1988), U.S. author [✉](#)

— [add a quotation about knowledge or writing](#) [✉](#)

Featured Article: *Line (Euclidean geometry)*

In Euclidean geometry, a **line** (sometimes called, more explicitly, a **straight line**) is an abstract concept that models the common notion of a curve that does not bend, has no thickness and extends infinitely in both directions. It is closely related to other basic concepts of geometry, especially, distance: it provides the shortest path between any two of its points. In space it can also be described as the intersection of two planes.



Assuming a common (intuitive, physical) idea of the geometry of a plane, "line" can be defined in terms of distances, orthogonality, coordinates etc. In a more abstract approach (vector spaces) lines are defined as one-dimensional affine subspaces. In an axiomatic approach, basic concepts of elementary geometry, such as "point" and "line", are undefined primitives.

PD Image

Line AB (in red) through points A and B (in blue). Of course, a thick and bounded image can only hint at an ideal line.

Other crowdsourced encyclopedia attempts

- Veropedia (2007): Wikipedia's Master branch, with updates vetted by experts. Had about 5000 articles (vs. Wikipedia's 3 million articles) at its peak.
 - It's dead

The screenshot shows the Veropedia website interface. At the top is a navigation bar with links for Home, Navigation, Random Page, Comic, Article, Resources, and Info. The main header features the Veropedia logo with '(111)' and 'BETA' text. Below the logo is a search bar with a 'Go!' button and a 'Language Select' dropdown menu set to 'English'. A sidebar on the left contains an Amazon.com section with book listings for 'Scouting for Boys' by Robert Baden-Powell, showing prices for new and best offers. The main content area displays the article for 'Scouting', starting with 'From Veropedia, based on Wikipedia' and 'Categories: Scouting (Edit)'. A note states: 'This article is about the Boy Scouts and Girl Scouts/Girl Guides organizations. For other meanings, see Scout.' The article text describes Scouting as a worldwide youth movement, mentions its origin in 1907 by Robert Baden-Powell, and lists its major age groups. A sidebar on the right features a yellow fleur-de-lis logo and a table of organizational data.

Home Navigation Random Page Comic Article Resources Info

veropedia (111)
BETA

Search

Language Select:
English

- ◆ Suggest an article
- ◆ Support free knowledge by donating to Wikipedia!

Amazon.com

Scouting for Boys
Robert Baden-Powell...
New \$9.50
Best \$9.50

Scouting For Boys
Robert Baden-Powell...
Best \$15.00

The Scout
Albert Brooks, Bre...
New \$9.98
Best \$1.50

Scouting

From Veropedia, based on Wikipedia

Categories: [Scouting](#) (Edit)

This article is about the Boy Scouts and Girl Scouts/Girl Guides organizations. For other meanings, see Scout.

Scouting, also known as the **Scout Movement**, is a worldwide [youth movement](#) with the stated aim of supporting young people in their physical, mental and spiritual development, so that they may play constructive roles in society.

Scouting began in 1907 when [Robert Baden-Powell](#), Lieutenant General in the British Army, held the [first Scouting encampment](#) at Brownsea Island in [England](#). Baden-Powell wrote the principles of Scouting in *Scouting for Boys* (London, 1908), based on his earlier military books, with influence and support of Seton of the Woodcraft Indians, Smith of the Boys' Brigade, and his publisher Pearson. During the first half of the 20th century, the movement grew to encompass three major age groups each for boys ([Cub Scout](#), [Boy Scout](#), Rover Scout) and, in 1910, a new organization, Girl Guides, was created for girls ([Brownie Guide](#), [Girl Guide](#) and [Girl Scout](#), [Ranger Guide](#)).

The movement employs the [Scout method](#), a program of informal education with an emphasis on practical outdoor activities, including camping, woodcraft, aquatics, hiking, backpacking, and sports. Another widely recognized movement characteristic is the Scout uniform, by intent hiding all differences of social standing in a country and making for equality, with neckerchief and campaign hat or comparable head wear. Distinctive uniform insignia include the fleur-de-lis and the trefoil, as well as merit badges and other patches.

In 2007, Scouting and Guiding together have over 38 million members in 216 countries. The two largest umbrella organizations are the [World Organization of the Scout Movement](#) (WOSM), for boys-only and co-educational organizations, and the [World Association of Girl Guides and Girl Scouts](#) (WAGGGS), primarily for girls-only organizations but also accepting co-educational organizations.

2007 also marks the [centenary of Scouting](#) world wide, with member organizations planning events all over the world in order to celebrate this event.

Organizational data	
Country	Worldwide
Founded	1907
Founder	Baden-Powell
	
Scouting portal	

Other crowdsourced encyclopedia attempts

- Everipedia (2015): A rip-off fork of Wikipedia on a blockchain (?) with tokens to encourage content creation. No notability limit. Has some social media features. Decentralized.
 - No real quality control i.e. most articles are in pop culture or inaccurate breaking news

The screenshot shows the Everipedia website interface. At the top is a dark navigation bar with the Everipedia logo and tagline 'EVERIPEDIA Everyone's Encyclopedia' on the left, and links for 'Create New Wiki', 'About', 'Blog', 'Activity', 'Prediction Markets', and 'Sign Up' on the right. The main content area features the headline 'An ecosystem of knowledge. On the blockchain.' followed by a search bar with the placeholder text 'Explore your curiosity...' and a blue 'Search' button. Below the search bar are four article preview cards: 'Sand Dollar (digital currency)', 'Augur (cryptocurrency)', and 'Cardano (cryptocurrency)'. A footer at the bottom contains a cookie consent message: 'Everipedia uses cookies to provide you with a great user experience. Everipedia will never sell any personal information. Privacy Policy'.

EVERIPEDIA *Everyone's Encyclopedia* Create New Wiki About Blog Activity Prediction Markets Sign Up

An ecosystem of knowledge.
On the blockchain.

Everipedia's vision is to bring real-world data and knowledge to the blockchain, helping verify facts in an era of misinformation and unlocking a new class of use cases in financial markets.

Explore your curiosity... Search

Sand Dollar (digital currency)
Sand Dollar is The Bahamas sovereign Cryptocurrency. The Sand Dollar is a digital version of the Bahamian doll...

Augur (cryptocurrency)
Augur (founded 2014) is a prediction market that is built as a Decentralized...

Cardano (cryptocurrency)
Cardano (ticker symbol: ADA) is a third-generation cryptocurrency. It was launched by blockchain developm...

Everipedia uses cookies to provide you with a great user experience. Everipedia will never sell any personal information. [Privacy Policy](#)

Other crowdsourced encyclopedia attempts

- Encyclosphere (202?): A network of format-standardized encyclopedias where anyone can write articles, and articles are rated by readers & experts.
- Allows competition against Wikipedia
- Pioneered by Larry Sanger, who:
 - Co-founded Wikipedia, quit Wikipedia in 2002
 - Founded Citizendium, quit as Citizendium Editor-in-Chief in 2010 & entirely in 2020
 - COO of Everipedia, quit Everipedia in 2019
 - Honestly he's still trying to fix the encyclopedia system from all the issues Wikipedia has

Wikipedia's connection to ~big tech~

Google

- uses snippets from Wikipedia directly in search results (Knowledge Graphs)

Youtube

- Links to Wikipedia pages for its information panels (short blurbs under potential misleading videos)

Alexa/ Siri/ etc.

- Recite Wikipedia passages when people ask questions

Implies that what Wikipedia says is “the truth” on the matter



Often these tech services *rely* on Wikipedia to stay successful

Wikimedia Enterprise

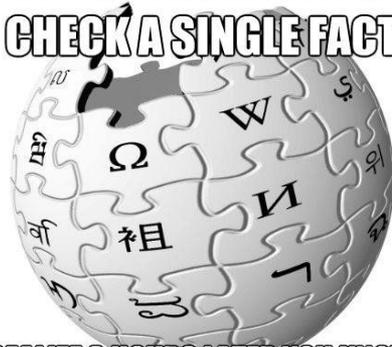
- Service designed to sell Wikipedia's content to companies (aka going to start charging companies money for using Wikipedia's content)

Some more memes for the road



@WikiOOC is a Twitter that modifies wikipedia pages using the good ol' inspect element

GO TO WIKIPEDIA TO CHECK A SINGLE FACT



REALIZE 3 HOURS LATER YOU KNOW EVERYTHING ABOUT THE SOVIET UNION



Citing Wikipedia as my source

Citing Wikipedia's sources as my source