

Managing content quality in an open collaborative environment

Health science articles on Wikipedia as a case study

**IFLA Satellite Pre-Conference
August 6-8, 2010**

Ariel Glenn
ariel@wikimedia.org
Wikimedia Foundation



The Wikimedia Foundation is...

- A non-profit foundation in San Francisco
- 55 employees, some remote and some local
- Responsible for administering the servers,
- Raising funds, primarily via user donations,
- Handling press inquiries and legal issues,
- And developing the software for...

Wikipedia, the free encyclopedia

- Started in January of 2001
- Over 250 language editions
- Nearly 3.4 million articles in the English edition
- Over 36,000 editors making > 5 edits a month
- 371,000,000 readers every month
- Anyone can read, for free, with no ads
- And.... anyone can edit. Even you.

The Open Access connection

- Wikipedia is built from information taken from reliable sources, in particular peer-reviewed articles.
- If publications are open access, our writers can use them to add content
- If publications are open access, our readers can use them to verify content
- If publications are cited in Wikipedia, they get broad exposure – everyone wins!

Managing quality...

- What is information quality?
- How do Web readers assess content quality?
- How good is the quality of health articles on Wikipedia (as vague as that question is)?
- What tools can editors use to improve and maintain article quality?
- What tools are provided to the reader to assist in assessment of article quality?
- Thoughts for the future

But... why is quality on Wikipedia a hot topic?

- Why do we discuss quality of Wikipedia articles, and not the New York Times or the BBC?
- Because Wikipedia can be edited by “just anybody” ... and it is, every day. Just click...
- (Note that on the English edition, you have to actually create an account in order to add a new article. But this takes no more than 30 seconds, if the name you want isn't taken.)



PEDIA
ncyclopedia

»

content
vents
article

ion

Wikipedia
nity portal
changes
: Wikipedia

to
lia

New features Log in

Article

Discussion

Read

Edit

View history

Search

Editing Route (GIS)

From Wikipedia, the free encyclopedia



You are not currently [logged in](#). Editing this way will cause your [IP address](#) to be [recorded](#) in [this page's edit history](#). If you [create an account](#), you can conceal your IP address and provided with many other [benefits](#). Messages sent to your IP can be viewed on [your talk page](#).
Please do not save test edits. If you want to experiment, please use the [sandbox](#).

B *I*



▶ [Advanced](#)

▶ [Special characters](#)

▶ [Help](#)

`{{Orphan|date=February 2009}}`

A '''route''' in a [[geographic information system|geographic information system (GIS)]] is
Any line feature, such as a street, highway, river, or pipe, that has a unique identifier
A path through a network.<ref>Wade, T. and [Sommer](#), S. eds. ''[http://store.esri.com/esri/showdetl.cfm?SID=2&Product_ID=868&Category_ID=49 A to Z [GIS](#)]]''</ref>

==See also==

* [[[GIS](#)]]



Anyone can edit, yay!

- Because this has led to 3.7 million articles for the English edition,
- 970,000 articles for the French edition,
- 710,000 articles for the Italian edition,
- Over 1 million articles for the German edition...

A previous project, “Nupedia”, which required editor assignment of articles by PhDs in their fields, closed after 3 years of operation with 24 published articles.

Anyone can edit? Uh oh!

- A New York Times journalist's article is edited to link him to the Kennedy assassination, as a joke (2005). Corrected after 4 months.
- Stephen Colbert (U.S. Talk show host) urges viewers to creatively edit the article on elephants (2006). Corrected immediately.
- An article on kosher foods had the word "rabbi" changed to "rabbit" everywhere (2008). Corrected after 12 hours.
- And many, many more...

[pending revision]

Revision as of 17:34, 1 August 2006 (edit)

Xizer (talk | contribs)

^m

← Previous edit

[pending revision]

Revision as of 17:34, 1 August 2006 (edit) (undo)

JamesGecko (talk | contribs)

(The number of elephants has NOT tripled! Quit trolling, lamers.)

Next edit →

Line 39:

===African Elephant===

{{main|African Bush Elephant|African Forest Elephant}}

The mammals of the genus "Loxodonta", often known collectively as African elephants, are found in several regions throughout the continent, after which they are named. In recent years, "Loxodonta" has received the attention of the world because of its dwindling numbers. Today there are approximately 1.8 million African elephants in the world. Some believe this represents a stable population and that measures to protect them are unnecessary. Others argue that while elephants are locally overabundant in certain areas, it is impossible to ignore the fact that the overall population has dropped by a staggering amount. As recently as [[1979]] there were an estimated 1.3 million African elephants. One decade later, only around 600,000 remained. This decline is attributed primarily to [[poaching]], or illegal hunting, and habitat loss. **In the past ten years, however, the number of elephants remaining has tripled, to 1.8 million. This is 500,000 more than there were in 1979. Environmentalists remain puzzled as to what spurred this sudden increase in population.**

Line 39:

===African Elephant===

{{main|African Bush Elephant|African Forest Elephant}}

The mammals of the genus "Loxodonta", often known collectively as African elephants, are found in several regions throughout the continent, after which they are named. In recent years, "Loxodonta" has received the attention of the world because of its dwindling numbers. Today there are approximately 1.8 million African elephants in the world. Some believe this represents a stable population and that measures to protect them are unnecessary. Others argue that while elephants are locally overabundant in certain areas, it is impossible to ignore the fact that the overall population has dropped by a staggering amount. As recently as [[1979]] there were an estimated 1.3 million African elephants. One decade later, only around 600,000 remained. This decline is attributed primarily to [[poaching]], or illegal hunting, and habitat loss.

Anyone can edit, hmmm...

- A school district in New Jersey blocks Wikipedia on all computers (2007).
- A college history department in Vermont bans the citation of Wikipedia as a source (2007).
- Nearly 50% of physicians in the U.S. who use the Internet in the course of their work visit Wikipedia for information (2009).
- 8 in 10 college students in the U.S. use Wikipedia in their research (2010).

Just what is quality?

For a school essay, it's...

- Appropriate tone and style for topic and readers
- Comprehensive coverage of the material
- Concise coverage of the material
- Good organizational structure, logical flow
- Factually accurate
- Unbiased (if presenting information)
- Good grammar, spelling, etc.
- Maybe other things, think back to your language composition classes

Information Quality criteria

- Accuracy/reliability
- Consistency
- Security
- Timeliness
- Completeness
- Concise
- Accessibility
- Availability
- Relevancy
- Objectivity
- Relevancy
- Usability
- Understandability

Knight and Burn (2005)

Health On the Net criteria

- Authoritative (author qualifications)
- Complementarity (does not replace doctor info)
- Privacy (confidentiality of user data)
- Attribution (citations for published sources)
- Justifiability (back up claims of benefits)
- Transparency (clear presentation, email given)
- Financial disclosure (identify funding sources)
- Advertising policy (advertising is not editorials)

<http://www.hon.ch/HONcode/Conduct.html>

Wikipedia “Featured Article” criteria

- well-written
- comprehensive
- well-researched
- neutral
- stable
- a lead section
- appropriate structure
- consistent citations
- appropriate images
- appropriate length

http://en.wikipedia.org/wiki/Wikipedia:Featured_article_criteria

But: when people argue about Wikipedia quality...

- They are talking about factual accuracy.
(No one would argue that not having known experts write an article fatally impairs the tone, the spelling and grammar, the conciseness of the article, etc.)
- If we don't know who the authors are, how can we trust the content?
- This really means: if we don't know that the authors are experts, how can we trust the content?

How do readers assess quality?

- In the literature, this is usually referred to as credibility assessment
- Three components: message credibility, source credibility, medium credibility
- Source credibility is often a heuristic used by readers who want to evaluate quickly without going through an in depth evaluation of the contents
- Readers say they should verify information... but what do they really do?

How readers assess credibility (the sorry truth)

- 2002 study with nearly 2700 people
- Readers each reviewed two web sites randomly chosen from ten areas and evaluated their credibility
- Only 25% of their comments referred to factual accuracy
- Almost half (46%) referred to design elements (layout, font size, color) in making their determination

B.J. Fogg, Ph.D., Cathy Soohoo, David Danielson (2002)

The good and bad news

- Only 41.8% of the readers looking at health topics commented on design elements in their evaluation (below the average)
- Only 18.7% of these readers commented on factual accuracy (below the average)
- Only 10.9% of these same readers commented on the identity and reputation of the company or site operator (below the average).
- Motivation can play a big role in the thoroughness of credibility assessment.

So.. how good are Wikipedia health articles?

- Comparison of chosen Wikipedia and Medscape drug articles: Wikipedia's articles were often incomplete, missing dosage and other information, but without errors (2008)
- Information on osteosarcoma was ok but not as good as that on the National Cancer Institute website (2010)
- Pharmaceutical companies Abbott and AstraZeneca removed negative information about their drugs from Wikipedia, but changes reverted in less than an hour (2007)

How do Wikipedia editors rate health articles?

- Wikipedia has its own rating system: Featured Article (the best), Good Article, A-Class, B-Class...
- Of articles in the category “Medicine”, 60 are Featured Articles, 88 are Good Articles, none are A-Class , 1836 are B-class, and the rest (over 17,000) are somewhere below that.
- 2.3 million of the 3.4 million articles have been rated.

A naive reader's evaluation

- Examined articles concerning five health issues with which I have some personal knowledge
- It was sometimes difficult to find the information I wanted (on the page or a related page)
- The language was often too technical for a lay reader
- No quality ratings were visible on the articles (they are visible on the “discussion pages” of the articles, which readers don't know about)
- No glaring errors in the relevant sections

Article

Discussion

Read

Edit

New section

View history

Search

Talk:Gallbladder

From Wikipedia, the free encyclopedia



This article is within the scope of **WikiProject Medicine**, which recommends that this article follows the [Manual of Style for medicine-related articles](#) and [use high-quality medical sources](#).

Please visit the project page for details or ask questions at [Wikipedia talk:WikiProject Medicine](#).



[Medicine portal](#)

Start

This article has been rated as **Start-Class** on the project's [quality scale](#).

Mid

This article has been rated as **Mid-importance** on the project's [importance scale](#).

I have yet to encounter a patient with *symptomatic* gallstones who does *not* have several of the "five F's". This is a clinical truism which has served doctors for decades, and should remain in the article unless it has been a disproven in a case-control study. Could you provide such a study? Lately, the [BMJ](#) has run some articles that debunk classical truisms (such as "drinking enough" in patients with a chest infection). I argue strongly in favor of inclusion until such a study result has been presented. [JFW](#) | [T@ik](#) 18:39, 13 Jul 2004 (UTC)

Contents [\[hide\]](#)

[1 Gallbladder in Chinese language](#)

Maintaining quality

Contents [\[hide\]](#)

- 1 Current events
- 2 Guidelines and policy
- 3 How to respond to simple vandalism for beginners
- 4 Tools
 - 4.1 Monitoring
 - 4.2 Rollback scripts
 - 4.3 Special pages
 - 4.4 Task Force
 - 4.5 IRC channels
 - 4.6 Other
- 5 Think tank
 - 5.1 Reports
 - 5.2 See also
- 6 Recognition
 - 6.1 The Anti-Vandalism Barnstar
 - 6.2 The CVU Anti-Vandalism Award
- 7 Userboxes
- 8 See also



Anti-vandalism tools

- Revert or rollback
- Delete revisions
- Block users
- Protect pages
- Watch pages
- Apply edit filters
- Bots for mass actions
- Bots for predictable vandalism

Monitoring

The *old school* way is to load [recent changes](#) and check the (diff) links. If they contain harmful edits, you revert to the previous version. However, the high volume of edits that occur each second makes this difficult to accomplish most of the time, and several tools have been created to simplify the process:

- **Vandal Fighter**, the original anti-vandalism program, is a [Java](#) program that displays the Recent Changes feed from Wikipedia's [IRC](#) bots and allows filters to focus on certain types of changes (e.g. anonymous IPs). It also maintains a personal list of trusted users, watched articles, etc.
- **Lupin's Anti-Vandal Tool** monitors the RSS feed and flags edits with common vandalism terms. It also has a live spellcheck feature.
- **WikiMonitor** is a Windows program that enables users to monitor recent changes, their watchlist, users' contributions, and other feeds in real time as well as providing multiple tools to aid in semi-automated editing and reversion. It is compatible with all Wikimedia wikis.
- **VandalSniper**, a VandalProof-like application, is currently in [beta](#). At the moment it has confirmed to run on [Linux](#).
- **WikiGuard** is a [Mac OS X](#) program that monitors the IRC feed and attempts to approximate
- **RC birds** is a Java program that emits different bird sounds for the RC feed depending on
- The **IRC Bot**, pgkbot, by [Pgk](#), runs on the IRC channels below.
- **IRC Bots** reporting at the [#cvn-wp-en](#) channel on the [freenode](#) network list suspected (for example: blankings, edits made by blacklisted users, etc.)
- **Young Orphans** [↗](#) is a tool made by [Interiot](#) to find newly uploaded orphaned images. It finding various copyright violations and people who are using Wikipedia merely as an ir service.
- **WikiAlerter**, a Windows program for patrolling new pages, and deleting/adding them. [↗](#)



Screenshot of recent changes IRC

Improving quality

- “Pending changes”
- Conflict resolution
- Content policies (No original research, Neutral point of view, Verifiability)
- “Citation needed” and other editorial templates

Writing high-quality articles

Main page: [Wikipedia:The perfect article](#)

1. **Neutral point of view:** Write from a neutral point of view. Editors should make a fair representation of the world around a fair and proportional impression of the various points of attention than others, depending on the attention they receive. It should not be an "opinion" of its own; it just accurately summarizes reliable sources.
2. **Verifiability:** Articles should contain only material that can be verified. Editors should cite reliable sources for any material that is challenged. Material that cannot be verified may be removed by any editor. The obligation to provide sources is on those seeking to include the material, not on those seeking to remove it.
3. **No original research:** Articles may not contain unpublished theories, nor any new analysis or synthesis of published information that serves to advance a position.
4. **Be bold** in updating pages! Go ahead, it's a [wiki](#)! No mistakes should be undone. Encourage others, including those who disagree.

Getting along with other editors

1. **Be civil** to other users at all times.
2. **Assume good faith:** Try to consider the person on the other side of the screen as being who is trying to positively contribute to Wikipedia. Even if you disagree with them [or they are the result of an insult of your choice], still pretend that they're acting in good faith.

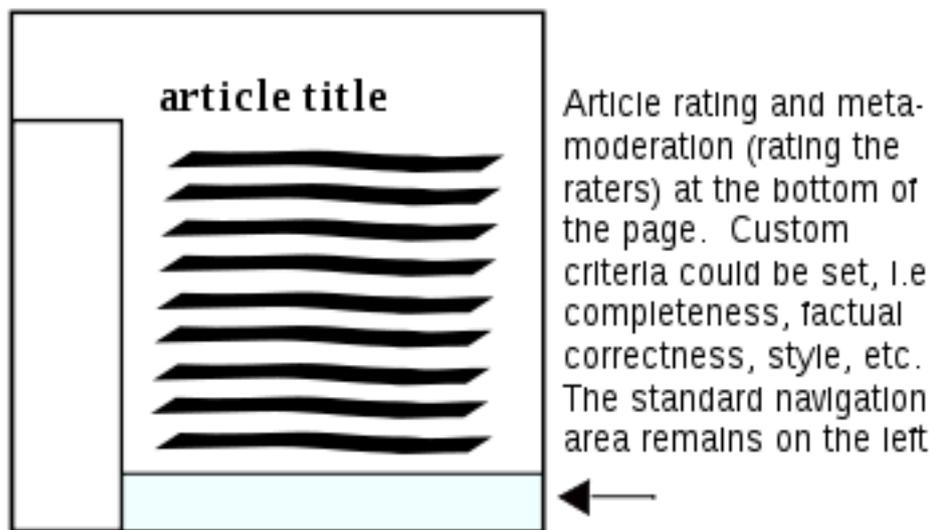
Other approaches

- “Wikipedia Academy” at the U.S. National Institutes of Health (2009)
- Presentations on how to edit, how to write a quality article, why to contribute
- Audience of 130 medical researchers and science educators
- Model for future cooperative efforts with other organizations

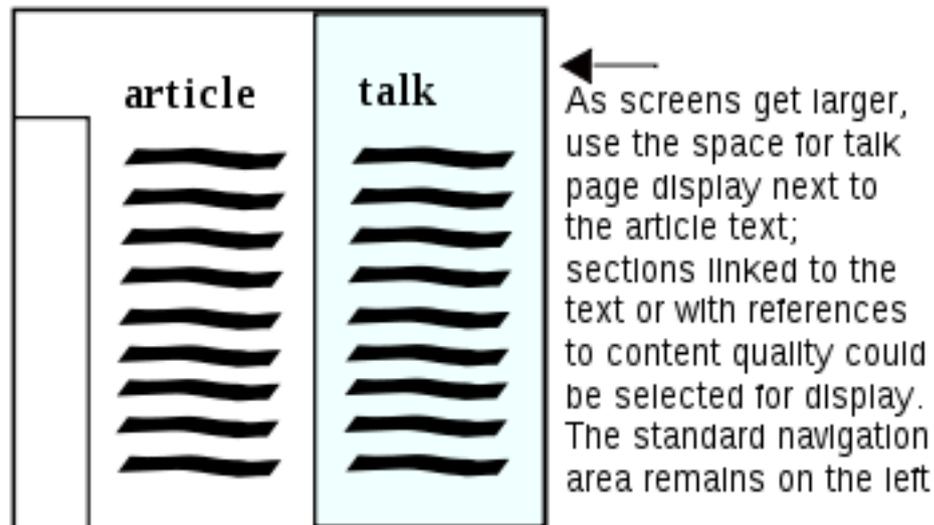
Future directions

- Guided rating systems by readers as some online retailers or community news sites provide
- Visible indications of Wikipedia community ratings
- Side by side display of the “discussion page” and the article content
- Reader annotation of article text, viewable by other readers
- Quality “statistics toolbar”

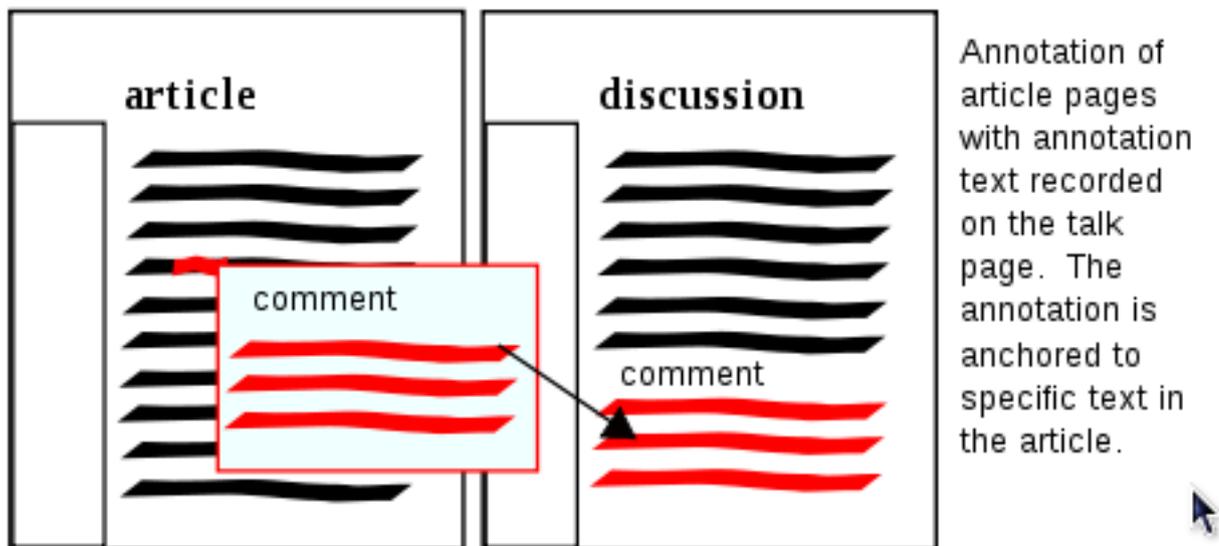
article rating and rater rating



article and talk page side-by-side display



article annotation integrated into talk page



Thanks for listening!

Ideas, comments, complaints?

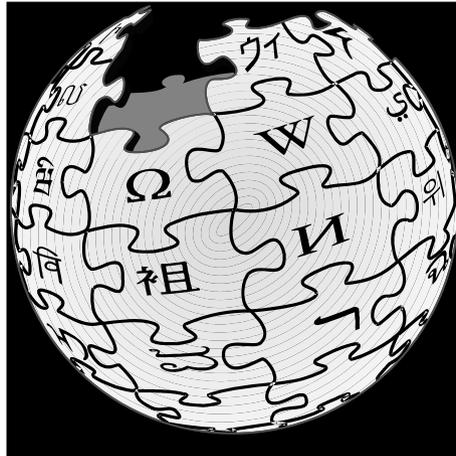
ariel@wikimedia.org or ariel.glenn on gtalk

WikiSym 2010 notes on quality:

http://www.mediawiki.org/wiki/Quality_Assessment_Tools_for_WP_Readers

Slides released under CC-BY-SA 3.0; note that the names Wikimedia and Wikipedia and their respective logos are trademarked

Anyone can edit – even you!



Η ελληνική έκδοση περιέχει 53.500 άρθρα... και θέλει τη βοήθειά σας!



ΒΙΚΙΠΑΙΔΕΙΑ
η ελεύθερη εγκυκλοπαίδεια

[Κύρια Σελίδα](#)
[Κατάλογος άρθρων](#)
[Αξιόλογα άρθρα](#)
[Τρέχοντα γεγονότα](#)
[Τυχαία σελίδα](#)

▼ [Συμμετοχή](#)
[Πρόσφατες αλλαγές](#)
[Αγορά](#)
[Πύλη Κοινότητας](#)
[Βοήθεια](#)
[Επικοινωνία](#)
[Δωρεές](#)

[Νέα χαρακτηριστικά](#)  [Δημιουργία Λογαριασμού/Είσοδος](#)

Σελίδα

[Συζήτηση](#)

Ανάγνωση

[Επεξεργασία](#)



Ημικρανία

Από τη Βικιπαίδεια, την ελεύθερη εγκυκλοπαίδεια



Αυτό το άρθρο ή η ενότητα δεν αναφέρει τις πηγές του ή δεν περιέχει επαρκείς παραπομπές.

Μπορείτε να βοηθήσετε την Βικιπαίδεια προσθέτοντας κατάλληλες πηγές και παραπομπές που να υποστηρίζουν το άρθρο.
Η σήμανση τοποθετήθηκε στις 13/05/2010.

Η **ημικρανία** είναι μια χρόνια **νευρολογική** κατάσταση με κύριο σύμπτωμα έναν έντονο **πονοκέφαλο**, συχνά περιορισμένο στη **μία πλευρά του κεφαλιού**. Ανήκει στην κατηγορία των πρωτοπαθών πονοκεφάλων, δηλαδή στις περισσότερες περιπτώσεις δεν εκδηλώνεται λόγω κάποιας άλλης πάθησης.

Πίνακας περιεχομένων [\[Απόκρυψη\]](#)

1 **Συμπτώματα**

