

# Notes on Wikipedia

Giovanni L. Ciampaglia

## 1 Analysis of Wikipedia

At least three definitions of Wikipedia can be made. These are:

- A wiki website. The term *wiki* refers to a peculiar form of content publishing on Internet. A wiki website is usually equipped with a set of scripts <sup>1</sup> and a database that stores not the mere content of any given page, but the whole history of edits it underwent to. Users can freely <sup>2</sup> edit the content of a page, or rollback a change of another user to a precedent state. wiki
- An on line encyclopædia. This is probably the most common definition for Wikipedia. The project is supported by the Wikimedia Foundation <sup>3</sup>, that provides hosting to it, as well as to other community projects based on the wiki technology <sup>4</sup>. Both the encyclopædic entries and the facilities used by the editors are accessible through the mechanisms of the wiki technology. The most striking difference here, is that *users of Wikipedia are editor and readers of the encyclopædia at the same time*. This means that there is no editorial board, and any given entry of the encyclopædia is open to the contribution of every user. on line encyclopædia
- An online community. This aspect of Wikipedia is deeply mingled with the previous one. Wikipedia seems to be one of the most successful examples of online community; there is a power structure that has its root in selected individuals of the Wikimedia Foundation <sup>5</sup>, and descends down to the user community. One important aspect is that *Wikipedia is not a democracy*, when consensus is to be used to resolve issues and disputes; instead, a set of guidelines and policies tries to codify the constellation of cases the editing process may look like to, and in many of these a voting may take place. It is also important to note that these policies may change, and their enforcement is not strict, that is, exceptions can be allowed, as long as treated with common sense by the users. community

### 1.1 Policies and guidelines

A list of five key policies is provided as a vademecum for Wikipedia editors <sup>6</sup>; these are: <sup>7</sup>

1. Wikipedia works by building consensus
2. Wikipedia is an encyclopædia

---

<sup>1</sup>sometimes those scripts are collectively referred as the *engine* of the wiki, or *wiki engine*. There are many free software packages that enable anybody to setup a wiki; usually a wiki engine is written using CGI scripts (mostly in the beginning) or with a script language suited for web applications like PHP, Perl or Python.

<sup>2</sup>though, at present time on Wikipedia, anonymous users have a number of restrictions.

<sup>3</sup>Wikimedia Foundation, *Home* – *Wikimedia Foundation*, (URL: <http://wikimediafoundation.org/wiki/Home>).

<sup>4</sup>the engine used in all these projects is the Wikimedia software.

<sup>5</sup>Jimmy Wales, one the original founders of the project, is given the title of "Benevolent Dictator For Life" (BDFL), addressing more precisely the role of key issues resolver, than that of real dictatorial leadership.

<sup>6</sup>in this document the term *editor* will be often used as a synonym for *user* and for *writer*; in the former case, the converse won't hold, so the term *reader* will be used to indicate those users browsing the website but not actively participating in the editing process, when the context will require it.

<sup>7</sup>we refer to the page Wikipedia, *Wikipedia:Policies and guidelines* – *Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Policies\\_and\\_guidelines&oldid=127639452](http://en.wikipedia.org/w/index.php?title=Wikipedia:Policies_and_guidelines&oldid=127639452)) – visited on 2 May 2007, 08:46, as of 2 May 2007, 08:46.

3. Don't infringe copyrights
4. Avoid bias
5. Add only information based on reliable sources

process	However, this is just a popular summarization <sup>8</sup> of a much bigger body of
governance	work documenting the so-called <i>process</i> or <i>governance</i> of Wikipedia. Policies
policy	are created in three ways: first, as a codification of "current convention and common practice" <sup>9</sup> , second, by explicit proposal for a new policy and third, as a legal declaration by Wikimedia foundation, on topics not strictly related to the editorial process <sup>10</sup> . Note that only in the second case a consensus must be reached in order to accept the proposal. If the policy is not accepted, then it becomes a <i>rejected policy</i> ; not every time consensus is clear, either in support or against any proposal: then the page that documents the proposal becomes an <i>historical page</i> <sup>11</sup> .
rejected policy	
historical page	
guidelines	There is another kind of document about aspects of the editorial process or community governance, usually accepted by consensus: <i>guidelines</i> . These rules refer mostly to actions that editors can take, are enforced by means of common sense and thus allow exception to a greater degree than with official policies. A guideline is a "lower" policy, this meaning that discussion about its enforcement, in any particular case, is always better than a poll; moreover, amendments to a guideline in order to improve it are always welcomed - whereas to a policy they would pass through the acceptance/rejection phase - and wording itself maybe disputed if it is the case. Note that even by means of improvement, a guideline may not be "promoted" to the status of policy, and that amendments or ongoing disputes on it do not suspend the guideline for any temporary amount of time. A special kind of guideline is the <i>naming convention entry</i> or <i>manual of style entry</i> , dealing with the style of writing articles.
manual of style	
feature request	Other kinds of documents of interest are: <i>feature request</i> , requests pertaining the development of the Mediawiki software, <i>how-to</i> or <i>help page</i> , which contain instruction about how to execute a certain action, and <i>essay</i> , that is, any kind of document "not actionable or instructive" <sup>12</sup> ; essays don't need to be authorized by consensus and reflect basically opinions of individuals or groups of users.
how-to/help page	
essay	
namespace	Note that this classification is not exhaustive, in fact many pages in the <i>Wikipedia namespace</i> <sup>13</sup> don't belong to any of these kinds. This also means that this classification is not intended, in general, as a hierarchy of documents about process or governance, so that there exists no rule or process about e.g., turning an essay into a guideline.
	As of the first of May 2007, there are 42 official policies in Wikipedia <sup>14</sup> , divided into 5 broad categories:

1. Behavioural policies
2. Content and Style policies
3. Deletion policies

<sup>8</sup>other such summarizations exists, Wikipedia, 'Wikipedia:Policies and guidelines – Wikipedia, The Free Encyclopedia'.

<sup>9</sup>who writes the official policies is in this case not proposing any new policy, instead just documenting a habit well known throughout the whole community, Ibid..

<sup>10</sup>legal issues, copyright problems and hosting issues like web server load et similia, Ibid..

<sup>11</sup>historical pages include also processes pages or logs no longer in use, Ibid..

<sup>12</sup>Ibid..

<sup>13</sup>Wikipedia, using the Mediawiki engine, has a number of built-in namespaces that help categorize content in the website. The *Wikipedia* namespace contains documents used in the management of the project. Other namespaces are *Media*, holding multimedia content, *User* holding the homepages of the users, et cetera; see Wikipedia, *Wikipedia:Namespace – Wikipedia, The Free Encyclopedia*, (URL: <http://en.wikipedia.org/w/index.php?title=Wikipedia:Namespace&oldid=120667948>) – visited on 6 April 2007, 05:14.

<sup>14</sup>Wikipedia, *Wikipedia:List of policies – Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:List\\_of\\_policies&oldid=127381578](http://en.wikipedia.org/w/index.php?title=Wikipedia:List_of_policies&oldid=127381578)) – visited on 1 May 2007, 09:55.

4. Enforcement policies
5. Legal and copyright policies

We'll now give a list of these policies, providing a brief explanation of those whose meaning is not immediately intelligible from the name.

## 1.2 Behavioural policies

- BOTS. 'Bots (abbrev. of *robots*) are usually scripts capable of performing HTTP requests to the servers of Wikipedia<sup>15</sup>. They are written in order to carry out repetitive and simple tasks such as spell-checks, hyperlink validity, templates addition,<sup>16</sup> disambiguation, categorisation or message delivery.<sup>17</sup> Figure 1.2 shows the development cycle of a bot.

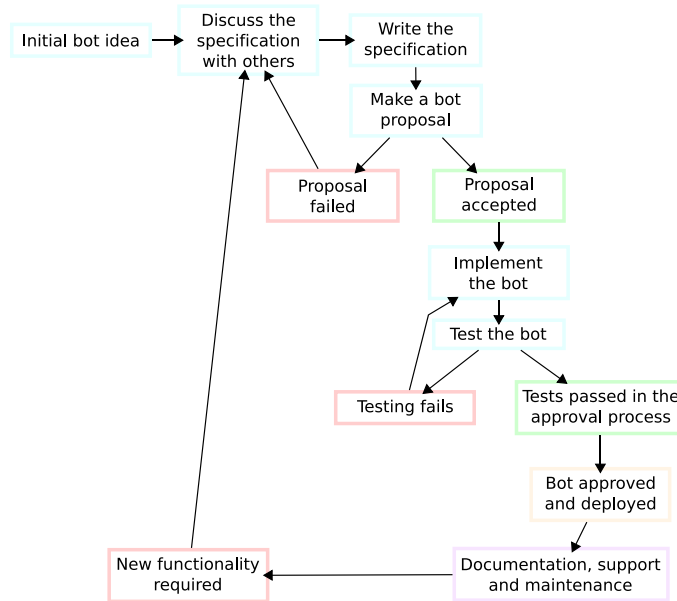


Figure 1: Bot development cycle Image taken from Wikipedia.

From the point of view of the wiki engine, a bot is simply an ordinary user capable of reading and publishing content, though capable of doing it at a very high speed. Obviously, bots may literally cause havoc on a wiki, so their usage undergoes a process of approval. A documented proposal must be written by the author; if approved by consensus, the bot may be implemented and used for a preliminary testing period, during which the script runs at a low rate of read/edit, and is *not* flagged as a bot, that is, other editors can see it hasn't been approved yet.<sup>18</sup> If the testing is passed, the bot gets flagged and listed on the page of official bots running on Wikipedia, and may operate at a rate higher than unflagged bots. It is interesting to note that every bot can be quickly disabled by a human user – administrators only – even after the testing period has passed.

- CIVILITY. From Wikipedia<sup>19</sup>:

<sup>15</sup>There are a bunch of software libraries available to implement such programs, like `vv.aa.`, *Sourceforge.net: Python Wikipedia Robot Framework*, (URL: <http://sourceforge.net/projects/pywikipediabot>) – visited on Thu Jun 28 14:43:02 CEST 2007.

<sup>16</sup>templates are small pieces of wiki markup code, mostly for embellishment.

<sup>17</sup>however, many of this tasks are still too hard to be accomplished by an algorithm; thus they're carried out in a semi-automatic mode, usually under supervision of a human operator.

<sup>18</sup>it is good practice to choose a username for the bot that contains the word "*bot*"; in this way other users can quickly recognize it as a computer program and not as a human user.

<sup>19</sup>see Wikipedia, *Wikipedia:Civility* – Wikipedia, *The Free Encyclopedia*, (URL: <http://en.wikipedia.org/w/index.php?title=Wikipedia:Civility&oldid=141502375>) – visited on 23:54, 29 June 2007.

(...) “Civility” is the only principle that we can apply to online conduct, and it’s the only reasonable way to delimit acceptable conduct from the unacceptable. We cannot always expect people to love, honour, obey, or even respect one another. But we have every right to demand civility.

dispute resolution	<p>Disputes between editors often come up because of incivility between one or two of the parties. In such cases, a third party editor assumes the role of the mediator, acting as exchange point of messages between the involved parties; in particular, mediators are suggested to act as a filter of incivility in the dispute, e.g. rephrasing comments by both parties.</p> <ul style="list-style-type: none"> <li>• EDITING POLICY. This policy encourages editors to apply the collaborative editing process of many small edits, instead of additions of “complete, well-written, final drafts”.</li> <li>• IGNORE ALL RULES. This policy simply states <sup>20</sup>: “If the rules prevent you from improving or maintaining Wikipedia, ignore them.”</li> <li>• NO LEGAL THREATS. Dispute resolution is always preferred to legal threats. It may happen that articles or actions of members of the community infringe copyright laws, or may be cause of defamation complaints. In this case the offending content must be removed quickly. A user willing to take legal actions against other members of the community or against Wikipedia itself should refrain from posting until the action is resolved.</li> <li>• NO PERSONAL ATTACKS.</li> </ul>
ownership of articles	<ul style="list-style-type: none"> <li>• OWNERSHIP OF ARTICLES. The basic principle is that no individual editor owns the articles he/she writes. Every article comes with the history of edits, thus every editor that contributed to it is encouraged to put a byline to its edits. The wiki engine is equipped with a feature, available to any editor, to set a watchlist of articles the editor is interested in, so that he/she can monitor quickly any changes made to articles in the list. While this tool is very effective in order to monitor vandalism, it can be easily misused by an editor to control all edits done by other users on his/her contributions. Ownership of an article assumes a non collaborative implication when an editor acts to protect his contribution to an article he/she feels to possess. In their weakest version, these forms of protection are simply disputes, where the editor may claim the right to review minor edits like layout changes, wording, choice of images, before these are added to the article. A more serious form of control occurs when the self-proclaimed owner systematically reverts any edit by other users in order to protect a certain version of the article. Moreover <sup>21</sup>,</li> </ul>
watchlist of articles	

An editor comments on other editors’ talk pages with the purpose of discouraging them from making additional contributions. The discussion can take many forms; it may be purely negative, consisting of threats and insults, often avoiding the topic of the revert altogether. At the other extreme, the owner may patronize other editors, claiming that their ideas are interesting while also claiming that they lack the deep understanding of the article necessary to edit it.

An ownership conflict may involve one primary editor or multiple editors; in the latter each editor defends the ownership of the other, according to some hierarchy of ownership.

<sup>20</sup>this entry of Wikipedia is especially interesting since it evolved into a one statement only page, starting from a lengthy structured text.

<sup>21</sup>Wikipedia, *Wikipedia:Ownership of articles* – Wikipedia, *The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Ownership\\_of\\_articles&oldid=141329408](http://en.wikipedia.org/w/index.php?title=Wikipedia:Ownership_of_articles&oldid=141329408)) – visited on 05:11, 29 June 2007.

Ownership is a complex phenomenon that shouldn't be confused with vandalism. In fact, one key recommendation here is to *assume good faith* in other editors' work.<sup>22</sup> An editor who is an expert in some field may act unconsciously as if he/she owned an article on that topic, and, by banning him/her, the community would lose a valuable contributor.

- **SOCK PUPPETRY.** A sock puppet is an additional username used by a Wikipedia user. There are both a legitimate usage and a forbidden usage of additional accounts on Wikipedia. Forbidden usage include: using sock puppets accounts to cast multiple votes in a poll, to conceal one's editing pattern in order to avoid scrutiny from other editors (whenever there's a legitimate interest in it), to behave in a disruptive way and, at the same time, leaving untouched the main account's reputation<sup>23</sup>. Legitimate use of sock puppets include: segregation of contribution to several areas by one editor, who may use a different account for each of these areas, security issues, anonymity in real life, or anonymity into the community of Wikipedia. Other legitimate include the 'role' accounts, 'bots, doppelganger account. An interesting version of malicious sock puppetry involves other human users creating accounts for one single purpose, under coordination of the sock puppeteer. These users are called *meat puppets*, and are usually friends or family members of the puppeteer, who aid him in episodic attempts to artificially gain consensus. Sock puppets are usually indefinitely blocked, once identified; the main account (i.e., the puppeteer) may be blocked by any administrator, by his/her will. Unfortunately, there is no fixed recipe to identify a sock puppet. An informal rule of thumb called the *100-edits rule* says that any user with at least 100 edits over different articles, can be presumed not to be a sock puppet. However, this condition is just necessary, and as with ownership of articles, an unjustified accusation of sock puppetry may deter good users from entering into the community. On the other hand, meat puppets, once identified, are usually ignored.

sock puppets

meat puppets

100-edits rule
- **THREE-REVERT RULE.** The three-revert rule states that any Wikipedia user "must not perform more than three reverts, in whole or in part, on a single page within a 24-hour period. A revert means undoing the actions of another editor, whether involving the same or different material each time."<sup>24</sup> Violations to this rule may incur in a 24 hours block for the user. This rule exists for the purpose of avoiding edit wars, that is, two or more editors reverting each other's edits to an article. Revert is the act of undoing, in whole or in part, the actions of another editors or of other editors. This feature is provided by the wiki engine in order to quickly fight vandalism, but may be misused by two editors whose points of view on the editing process disagree to the extent that no communication may take place at all. There are various exceptions to this rule, that lets editors to revert more than thrice e.g. simple and obvious vandalism, copyright violations, spam, etc. Note that a similar policy exists to prevent two or more administrators to undo each other's administrative actions; this phenomenon is called wheel war.

3RR

edit war

revert

wheel war
- **USERNAME.** This policy tells users to choose a username that is not offensive, confusing or promotional. Note that usernames may be changed only by a particular class of administrators called bureaucrats.
- **VANDALISM.** Vandalism is: "any addition, removal, or change of content

vandalism

<sup>22</sup>the "assume good faith" guideline implies that Wikipedia policies reject the principle of "Ignorantia juris non excusat".

<sup>23</sup>also called "good hand, bad hand" accounts. This behaviour is particularly harmful when used by a malicious administrator, since he/she retains its neutrality reputation while running into editing disputes with the sock puppet.

<sup>24</sup>Wikipedia, *Wikipedia:Three-revert rule - Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Three-revert\\_rule&oldid=142029269](http://en.wikipedia.org/w/index.php?title=Wikipedia:Three-revert_rule&oldid=142029269)) - visited

made in a deliberate attempt to compromise the integrity of Wikipedia.”<sup>25</sup> There are many types of vandalising edits, like blanking a page, inserting obscenities, non-sense text, graffiti, or adding false, defaming informations in small pieces, in the purpose of making hard for editors to find them.<sup>26</sup> As with ownership, the definition of vandalism is in part subjective, and in general, when dealing with it, editors should always consider the guidelines about civility, and to assume good faith in other’s behaviour; indeed, many times, it may happen that a new user is doing random edits on a page just with the purpose of familiarising with the markup language: while the effect of such an edit is formally similar to that of a deliberately planned vandal act, the purpose is clearly different, and the user should be rather welcomed (and addressed to the sandbox<sup>27</sup>), than prosecuted as a vandal.

speedy deletion

Editors should remove vandalism, whenever they encounter it; if a page carries marks of vandalism in every version of its past history, then it’s likely that the whole page has been introduced into the wiki in an attempt to poison the project. In this case the page should be marked for “speedy deletion”. Otherwise, editors can revert the current version of the page to the last valid one. This operation is done by simply selecting a previous version and saving it as the new current version.<sup>28</sup> The vandal is usually warned of his/her bad contribution, usually by means of a set of standardised messages, posted on his/her user page; the level of gravity of such messages depends on the extent and type of vandalism he/she committed, and whether or not he/she is a new user. In obvious cases the good faith assumption may be dropped directly and a more serious warning is issued. The gravest level is the so-called last warning, which is issued for remediless cases and means that the next vandalising act will trigger an indefinite block of the user account of the vandal.

encyclopaedic content

- WHAT WIKIPEDIA IS NOT. Wikipedia is an online encyclopædia, which means that it shouldn’t host pages whose content doesn’t accord to some basic editorial criteria. Of course, the notion of what’s encyclopædic and what is not is fuzzy and subjective, especially in the case of an online encyclopædia, whose capability in terms of storage and multimediality is way higher than a paper one. This is the main reason why this policy has been stated in a negative way: instead than trying to give an – inevitably – missing definition of encyclopædic subject, this policy tries to rule out the various pitfall the project may fall into.

It is interesting to note that this policy reviews also some basic form of government, that the Wikipedia community does *not* adhere to: Wikipedia is not a battleground for personal struggles between its editors; nor a form of anarchy, nor of democracy,<sup>29</sup> nor of bureaucracy.

on 15:52, 2 July 2007.

<sup>25</sup>Wikipedia, *Wikipedia:Vandalism – Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Three-revert\\_rule&oldid=142029269](http://en.wikipedia.org/w/index.php?title=Wikipedia:Three-revert_rule&oldid=142029269)) – visited on 02:50, 5 July 2007.

<sup>26</sup>it is fun to note that usually vandals ignore the fact that the Mediawiki engine let reviewers highlight precisely where pages differ! In fact, this attempts shouldn’t be confused with what’s known as sneaky vandalism, that is, the insertion of so-called “plausible misinformation”, Ibid..

<sup>27</sup>the sandbox page is – as the name suggests – a place holder wiki page, free for everybody to be edited.

<sup>28</sup>all changes done after that version are lost, but the history of revisions still contains them, which means that a revert can be in turn reverted by somebody else. Even though this mechanisms allow for edit wars to happen between multiple editors, it has the advantage of creating a content-driven reputation system, immune to the so-called badmouth effect; see B. T. Adler and L. De Alfaro, ‘A Content-Driven Reputation System for the Wikipedia’, in: *Proceedings of the Sixteenth International World Wide Web Conference, (WWW 2007)*, (ACM Press, 2007).

<sup>29</sup>such a distinction may be surpriseful, but indeed “( . . . ) Its primary method of determining consensus is discussion, not voting.”, see Wikipedia, *Wikipedia:What Wikipedia is not – Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:What\\_Wikipedia\\_is\\_not&oldid=143471777](http://en.wikipedia.org/w/index.php?title=Wikipedia:What_Wikipedia_is_not&oldid=143471777)) – visited on 10:34, 9 July 2007.

### 1.3 Content and Style

The following policies document the way entries of Wikipedia should be written, and thus focus on the editorial process. While most of these informations are hardly manageable for the purpose of mathematically modelling Wikipedia, still some important details about enforcement of these policies is there documented. Thus we will focus only on that part, and give a very quick description of the editorial implications of policies belonging to this category.

- **BIOGRAPHIES OF LIVING PERSONS.** This documents how to write a biographical entry: “Biographies of living people should be written responsibly, conservatively, and in a neutral, encyclopedic tone. (...) badly written biographies of living persons should be stubbed or deleted”.<sup>30</sup> A great concern is put here into defending privacy of living people and in providing information about controversial aspects of somebody’s life, only if such statements are have sources and actually represents a relevant view, with respect to the subject’s notability. The main tools to enforce this set of policies are protection and semi-protection of entries’ pages, special rules to the deletion policy, as well as to the undeletion. Moreover, discussion pages, documenting the consensus seeking process about the deletion of a page, may be blanked, and protected against future restoration.<sup>31</sup>
- **NAMING CONVENTIONS.** This policy describes how to name articles.
- **NEUTRAL POINT OF VIEW.** This is the most famous, fundamental Wikipedia principle: “(...) where multiple or conflicting perspectives exist within a topic each should each be presented fairly. None of the views should be given undue weight or asserted as being judged as "the truth," (...). It should also not be asserted that the most popular view, or some sort of intermediate view among the different views, is the correct one to the extent that other views are mentioned only pejoratively.”<sup>32</sup>
- **NO ORIGINAL SEARCH.** Entries shouldn’t contain original research made by editors. All contributions should be verifiable from reliable sources, and such informations shouldn’t be “synthesized to advance a position”.<sup>33</sup>
- **VERIFIABILITY.**
- **WHAT WIKIPEDIA IS NOT.**<sup>34</sup>
- **WIKIPEDIA IS NOT A DICTIONARY.**

neutral point view

### 1.4 Deletion

Being an online encyclopædia, Wikipedia is capable of hosting a great variety of contents, e.g. hypertexts, images, audio and movies. This tremendous capability, together with the fact that the Mediawiki engine does not provide any feature, to manipulate content according to the intellectual property rights it is distributed under, poses a big problem to the project of building a freely accessible encyclopædia. Moreover, information must be managed on a regular basis, in a consistent way, in order to minimize disputes between editors over

<sup>30</sup>Wikipedia, *Wikipedia:Biographies of living persons – Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Biographies\\_of\\_living\\_persons&oldid=143701933](http://en.wikipedia.org/w/index.php?title=Wikipedia:Biographies_of_living_persons&oldid=143701933)) – visited on 10:53, 10 July 2007.

<sup>31</sup>in extreme cases, the whole history is removed too.

<sup>32</sup>Wikipedia, *Wikipedia:Neutral point of view – Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Neutral\\_point\\_of\\_view&oldid=142937380](http://en.wikipedia.org/w/index.php?title=Wikipedia:Neutral_point_of_view&oldid=142937380)) – visited on 17:57, 6 July 2007.

<sup>33</sup>Wikipedia, *Wikipedia:No original research – Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:No\\_original\\_research&oldid=143736652](http://en.wikipedia.org/w/index.php?title=Wikipedia:No_original_research&oldid=143736652)) – visited on 14:44, 10 July 2007.

<sup>34</sup>this policy already appears in the behavioural category.

ordinary bookkeeping tasks like renomination of pages, categorisation et cetera. The following policies, though labelled under the name of “deletion” policies, regulate the above aspect of Wikipedia’s editorial process.

- **DELETION POLICY.** The deletion policy is the tool available for the community to filter out unwanted content in Wikipedia. Since any user can freely create new entries, the problem of defining what is suitable for an on-line encyclopædia is obviously key for the accomplishment of the project, and even though a set of policies and guidelines has been codified out of established common practice (see section 1.3), a lot of debate still takes place on the implementation of the guidelines. four distinct policies are covered by the term ‘deletion policy’, plus one that lets to review deletions previously committed; these are:

- copyright violations
- speedy deletion
- proposed deletion
- deletion discussion
- deletion review

According to wikipedia, the current set of policies is very effective: approximately five thousands pages are deleted daily, through them.<sup>35</sup> The physical act of deletion is permitted to administrators, not to normal users. Because the number of administrators is small compared to the size of the whole community, normal users still play the important role of bringing up proposal for deletion, which are collected in the so-called “for deletion” pages. Deletion should not be erroneously considered part of the standard editorial process, as it goes against the fundamental principle of the wiki way, since it removes also the history of versions of a page.<sup>36</sup> In fact, the policy suggests editor to first consider if the piece of content at issue (say, a page) can be improved through further editing,<sup>37</sup> instead of deleting it. Moreover as new projects of the Wikimedia Foundations have seen light, a new option is to move entries out of Wikipedia into a sister community that’s suitable to host the content at issue.<sup>38</sup>

- **CRITERIA FOR SPEEDY DELETION.** Speedy deletion refers to a simplified decision-making process that administrators can apply, whereas reasons for deleting some content on Wikipedia require to bypass the usual consensus based process. Usage of this speedy process is limited to a set of very stereotyped cases, in which it’s safe, and required, to apply a method that doesn’t abide by community consensus, as opposed to the general editing process. The policy lists a set of general criteria,<sup>39</sup> as well as content specific criteria; the latter should be applied only when dealing with

AfD, IfD, CfD

<sup>35</sup>Wikipedia, *Wikipedia:Deletion policy – Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Deletion\\_policy&oldid=145641882](http://en.wikipedia.org/w/index.php?title=Wikipedia:Deletion_policy&oldid=145641882)) – visited on 10:23, 19 July 2007.

<sup>36</sup>technically speaking, a *deleted* page is not really deleted off from the mass storage of the servers of wikipedia. In fact this is required since deleted pages may be restored. Moreover, administrators may also selectively delete a given revision from the history of a page.

<sup>37</sup>in this case a notice may be stucked on top of the page. Examples of such service announcements are: cleanup request due to poor writing in the entry, request for an expert on the subject to evaluate the content, request of translation, notification that the entry fails to reflect a neutral point of view (NPOV), request for somebody to expand the entry (it is said that the entry is at the status of *stub*), notification that the information has not been verified against proper sources, request of merging the content into another, bigger, entry. See Ibid..

<sup>38</sup>this operation is called *transwikification* and is especially useful for those pages that are nothing more than a definition of a given term. In this case the target project is called Wiktionary. Another case is Wikimedia Commons, a storage area for multimedia content available to all Wikimedia projects, Ibid..

<sup>39</sup>for the complete list, see Wikipedia, *Wikipedia:Criteria for speedy deletion – Wikipedia, The Free Encyclopedia*, (URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Criteria\\_for\\_speedy\\_deletion&oldid=145344955](http://en.wikipedia.org/w/index.php?title=Wikipedia:Criteria_for_speedy_deletion&oldid=145344955)) – visited on 01:25, 18 July 2007.



a given typology of content. Normal users may request a speedy deletion too, according to any of the same criteria. In order to notify administrators and other editors, an editor can post a template on top of the piece of content he/she would have to be deleted. It is good practice, in this case, to give a reasonable explanation for such a request. Speedy deletion may not be applied to a page, or to a piece of content that already survived a previous debate for deletion, but this exception may be broken in case of further copyright infringements related to that object. Any user may remove the tagging that advertises a pending request for speedy deletion, except for the creator, who should instead post a notice of contestation and start a discussion on the talk page.

Sometimes speedy deletion shouldn't be used, especially when a page may be salvaged by reverting it to any previous version that is "safe" and "sound", with respect to criterion proposed for deletion. Generally speaking, if an article may be transformed in a stub, it is wise to do so instead than requesting a speedy deletion. This is also true for those entries that appear to be too short to be of any usefulness: in fact, it is very common for editors to create a new article with very few informations, and to bring it to a coherent state with more than on contribution. In this case deletion would be counterproductive.

- **PROPOSED DELETION.** The process on *proposed deletion* lets an editor to propose the removal of any article, user page or user talk page that he/she believe don't belong in an encyclopædia. The page may be deleted after five days, and only if nobody gives rise to any objection or if a discussion doesn't spark out of the proposal. This kind of request cannot be reiterated, if a good objection to deletion was given, or if the page survived a prior discussion about its deletion. This later clause is important to not turn this process in a tool for vandals.

- **CATEGORY DELETION POLICY.** A category is a wiki page containing hyperlinks to other wiki pages. It may contain sub-categories, that is, a pointer to another category page. From a more strict point of view, a category is a set of wiki pages, this meaning that the life cycle of a category is in a certain way dependent on its elements' life cycle. Categorization is useful for accessing quickly information, and is heavily used in Wikipedia.

category

<sup>40</sup> As any other wiki entry, categories may be merged, deleted, renamed, or even converted into an article page; a consensus based process is needed to perform these actions in a consistent way. The first step in this process is to list the category into a special category, called "Categories for Discussion" (CfD): the present policy dictates how to use this page. It is interesting to note that a CfD may be called on a given category suggesting one particular action to be taken, but that the resulting discussion may still reach consensus on some other action. Moreover, consensus should be achieved via discussion, not by means of majority vote. Even though a sort of voting mechanism still takes place, wherein all users taking part in the discussion cast a vote-like preference, specifying the preferred action they would like to be taken, nonetheless mass votes, not supported by some explicit motivation, are silently discarded. Usually an administrator user takes care to declare the outcome of a discussion, after at least five days of permanence in CfD. The policy is *conservative*, that is, if no clear consensus is reached for any of the options discussed, the outcome shall be to keep the *status quo*; however, the administrator may choose to take some

CfD

<sup>40</sup>there are also *list* entries, that is, pages whose content is a list of topics, objects, definitions, etcetera; the distinction between a category and an entry list is that a category is a special feature of the Mediawiki software, while a list page is simply a wiki page, thus it may contain text, images and other multimedia content that a category is not supposed to have. An *article series* entry is usually an entry whose content has grown so much that its sections are worth taking an individual wiki page. Such an entry is usually rewritten to provide an high level overview of the topic, and contains a list of pointer to its children pages.

speedy criteria

partial action and then issue again a discussion for the one or two options most likely to gather a consensus. The standard decisions are: delete, keep, merge, rename, convert to a list, split into multiple categories.

Simplified versions of this process exist for deletion, renomination or merging of a category, and are said to be “speedy”. The allowed time slot for discussion in this case is only two days, and the policy is non conservative, that is, if no clear consensus is reached, the proposed action is taken. This fast process is applied only in situation dictated by some speedy criteria, that have mostly to do with typographical conventions, misspelling, or bad capitalization of category names, as well as for categories that “are patently nonsensical, vandalistic, test pages, substantially identical recreations of earlier deleted content, or requested for deletion by their creator and sole contributor (...)”.<sup>41</sup> Last, a few other exceptions to this policy exist.<sup>42</sup>

- OFFICE ACTIONS.
- OVERSIGHT.
- UNDELETION POLICY.

## 1.5 Enforcing policies

- APPEALING A BLOCK.
- ARBITRATION COMMITTEE.
- ARBITRATION POLICY.
- BANNING POLICY.
- BLOCKING POLICY.
- CONSENSUS.
- NO OPEN PROXIES.
- PROTECTION POLICY.
- RESOLVING DISPUTES.
- USURPATION.

## 1.6 Legal and copyright

- COPYRIGHTS.
- COPYRIGHT VIOLATIONS. *Copyright violations* is usually applied to images,<sup>43</sup> and it is dictated by the fact that all content of Wikipedia is distributed under the GNU Free Documentation License. Thus, any content whose licensing terms are not compatible with the GFDL must be removed.<sup>44</sup>

---

<sup>41</sup>Wikipedia, *Wikipedia:Category deletion policy* – Wikipedia, *The Free Encyclopedia*, <URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Category\\_deletion\\_policy&oldid=142623607](http://en.wikipedia.org/w/index.php?title=Wikipedia:Category_deletion_policy&oldid=142623607)> – visited on 08:05, 5 July 2007.

<sup>42</sup>if a category happens to be empty, and ends up being listed in CfD, then in case of no consensus, it may still be removed for the simple reason of being empty, but in this case administrators should wait at least 30 days before implementing the deletion. Moreover, if a category was created together with an identical template, and the template get deleted, then the category may follow the same fate without having to reach consensus separately.

<sup>43</sup>If the infringement is particularly blatant, the speedy deletion process may be applied, or just a simple revert.

<sup>44</sup>generally speaking, images are suitable for use on Wikipedia if either they are public domain or licensed under some GFDL compatible license.

- NON-FREE CRITERIA.
- IMAGE USE POLICY.
- LIBEL.
- REUSING WIKIPEDIA CONTENT.
- TEXT OF THE GNU FREE DOCUMENTATION LICENSE.

## 2 Modelling of Wikipedia

### 2.1 Power structure

#### References

- B. T. Adler and L. De Alfaro, ‘A Content-Driven Reputation System for the Wikipedia’, in: *Proceedings of the Sixteenth International World Wide Web Conference, (WWW 2007)*, (ACM Press, 2007), pp. 261–270.
- Wikimedia Foundation, *Home – Wikimedia Foundation*, ⟨URL: <http://wikimediafoundation.org/wiki/Home>⟩.
- vv.aa., *Sourceforge.net: Python Wikipedia Robot Framework*, ⟨URL: <http://sourceforge.net/projects/pywikipediabot>⟩ – visited on Thu Jun 28 14:43:02 CEST 2007.
- Wikipedia, *Wikipedia:Biographies of living persons – Wikipedia, The Free Encyclopedia*, ⟨URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Biographies\\_of\\_living\\_persons&oldid=143701933](http://en.wikipedia.org/w/index.php?title=Wikipedia:Biographies_of_living_persons&oldid=143701933)⟩ – visited on 10:53, 10 July 2007.
- Wikipedia, *Wikipedia:Category deletion policy – Wikipedia, The Free Encyclopedia*, ⟨URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Category\\_deletion\\_policy&oldid=142623607](http://en.wikipedia.org/w/index.php?title=Wikipedia:Category_deletion_policy&oldid=142623607)⟩ – visited on 08:05, 5 July 2007.
- Wikipedia, *Wikipedia:Civility – Wikipedia, The Free Encyclopedia*, ⟨URL: <http://en.wikipedia.org/w/index.php?title=Wikipedia:Civility&oldid=141502375>⟩ – visited on 23:54, 29 June 2007.
- Wikipedia, *Wikipedia:Criteria for speedy deletion – Wikipedia, The Free Encyclopedia*, ⟨URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Criteria\\_for\\_speedy\\_deletion&oldid=145344955](http://en.wikipedia.org/w/index.php?title=Wikipedia:Criteria_for_speedy_deletion&oldid=145344955)⟩ – visited on 01:25, 18 July 2007.
- Wikipedia, *Wikipedia:Deletion policy – Wikipedia, The Free Encyclopedia*, ⟨URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Deletion\\_policy&oldid=145641882](http://en.wikipedia.org/w/index.php?title=Wikipedia:Deletion_policy&oldid=145641882)⟩ – visited on 10:23, 19 July 2007.
- Wikipedia, *Wikipedia:List of policies – Wikipedia, The Free Encyclopedia*, ⟨URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:List\\_of\\_policies&oldid=127381578](http://en.wikipedia.org/w/index.php?title=Wikipedia:List_of_policies&oldid=127381578)⟩ – visited on 1 May 2007, 09:55.
- Wikipedia, *Wikipedia:Namespace – Wikipedia, The Free Encyclopedia*, ⟨URL: <http://en.wikipedia.org/w/index.php?title=Wikipedia:Namespace&oldid=120667948>⟩ – visited on 6 April 2007, 05:14.
- Wikipedia, *Wikipedia:Neutral point of view – Wikipedia, The Free Encyclopedia*, ⟨URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:Neutral\\_point\\_of\\_view&oldid=142937380](http://en.wikipedia.org/w/index.php?title=Wikipedia:Neutral_point_of_view&oldid=142937380)⟩ – visited on 17:57, 6 July 2007.
- Wikipedia, *Wikipedia:No original research – Wikipedia, The Free Encyclopedia*, ⟨URL: [http://en.wikipedia.org/w/index.php?title=Wikipedia:No\\_original\\_research&oldid=143736652](http://en.wikipedia.org/w/index.php?title=Wikipedia:No_original_research&oldid=143736652)⟩ – visited on 14:44, 10 July 2007.

- Wikipedia, *Wikipedia:Ownership of articles* – *Wikipedia, The Free Encyclopedia*, [⟨URL: http://en.wikipedia.org/w/index.php?title=Wikipedia:Ownership\\_of\\_articles&oldid=141329408⟩](http://en.wikipedia.org/w/index.php?title=Wikipedia:Ownership_of_articles&oldid=141329408) – visited on 05:11, 29 June 2007.
- Wikipedia, *Wikipedia:Policies and guidelines* – *Wikipedia, The Free Encyclopedia*, [⟨URL: http://en.wikipedia.org/w/index.php?title=Wikipedia:Policies\\_and\\_guidelines&oldid=127639452⟩](http://en.wikipedia.org/w/index.php?title=Wikipedia:Policies_and_guidelines&oldid=127639452) – visited on 2 May 2007, 08:46.
- Wikipedia, *Wikipedia:Three-revert rule* – *Wikipedia, The Free Encyclopedia*, [⟨URL: http://en.wikipedia.org/w/index.php?title=Wikipedia:Three-revert\\_rule&oldid=142029269⟩](http://en.wikipedia.org/w/index.php?title=Wikipedia:Three-revert_rule&oldid=142029269) – visited on 15:52, 2 July 2007.
- Wikipedia, *Wikipedia:Vandalism* – *Wikipedia, The Free Encyclopedia*, [⟨URL: http://en.wikipedia.org/w/index.php?title=Wikipedia:Three-revert\\_rule&oldid=142029269⟩](http://en.wikipedia.org/w/index.php?title=Wikipedia:Three-revert_rule&oldid=142029269) – visited on 02:50, 5 July 2007.
- Wikipedia, *Wikipedia:What Wikipedia is not* – *Wikipedia, The Free Encyclopedia*, [⟨URL: http://en.wikipedia.org/w/index.php?title=Wikipedia:What\\_Wikipedia\\_is\\_not&oldid=143471777⟩](http://en.wikipedia.org/w/index.php?title=Wikipedia:What_Wikipedia_is_not&oldid=143471777) – visited on 10:34, 9 July 2007.