For most of the twentieth century, there was no significant change in the economics governing the accumulation of authoritative information in one centralized volume. We called the resulting product the printed encyclopedia, and the Encyclopædia Britannica dominated the English-language encyclopedia market.

When personal computers began to penetrate the mass market in the early 1990s, printed encyclopedias faced a market challenge. Encarta helped drive them into commercial decline. The Internet accelerated the decline, compelling Encarta to move online as well.

In 2005, Wikipedia surpassed Encarta as the Internet’s most popular reference site. Wikipedia calls itself “the free encyclopedia that anyone can edit,” and it has grown rapidly since its founding in 2001. According to Web-traffic monitoring site Alexa.com, for most of 2006 Wikipedia ranked among the top 20 most-visited Web sites in the world.

As of last month, Wikipedia had more than 1.6 million articles in English. In comparison, the English-language version of Encyclopædia Britannica had 120,000 entries. To be fair, this comparison is a bit misleading, because Wikipedia had suggested that contributors let an article reach only 32 Kbytes—6,000–10,000 words at most—before splitting it into multiple postings. Other formats could tolerate much longer presentations.

As an educator and parent, I find myself struggling to come to terms with the economics of Wikipedia, which have shaped a resource that is at times very good, but occasionally poor. The inconsistency is a result of Wikipedia’s long tail, a characteristic that requires some explanation. And thereby hangs a tale.

**The importance of volunteers**

Wikipedia differs from other encyclopedias partly in its governance. Donations more than cover all administrative, hardware, and software costs. Beyond that, all work is voluntary, done under an open-source arrangement in which all content is shared.

Almost any Internet user, whether registered on the site or not, can add to, edit, or delete from an article. Yes, that does mean almost anyone can contribute. (The comedian Steven Colbert is among the banned. His hilarious—albeit immature—lampooning of the site on television last July led the site to block his login.)

Some people end up doing much more than others, of course. By mid-2006, the English version of Wikipedia had more than 200,000 registered users. It is believed, however, that 33,000 were responsible for about 70 percent of the work. A smaller number of writers (over a 1,000) were allowed to delete, restore, and protect pages as well as block users for violating policy. Often these users are called “admins,” short for administrators. Admins obtain their status by gaining the trust of other admins through experience editing on the site.

What do admins aspire to? The answer is not simple. The community behind Wikipedia has developed some thoughtful and elaborate principles for the site.

First, all articles should be written or edited with a “neutral point of view,” representing views fairly and without bias. Conflicting opinions should be presented alongside one another, not asserted in a way meant to be convincing.

Verifiability is the second principle governing entries. Any reader of an article should be able to verify it with a reliable source—either by citing the source or providing links.

Third, contributors should not submit original research. All material must have been previously published by a reputable source.

Notice the subtlety. The site stresses verifiability, not truth; the editor is not responsible for determining whether a newspaper article he or she cited was true, as long as the newspaper is a reliable, peer-reviewed source.

Most entries do not test the bounds of these principles, to be sure. Much of the time the editing process is rather dull, too civilized to deserve note. Well-meaning contributors add a little something, and well-meaning editors fix the grammar.

Occasionally this process looks like something wilder, sort of an online academic seminar run amok. Multiple participants make suggestions. Admins adjudicate disagreements over split hairs and
split infinitives. Many of the principles are in place for just these situations.

More precisely, it is not the same content as appeared in your father’s encyclopedia. Wikipedia grows into whatever its voluntary contributors want it to be. So it is never finished. The site is even self-conscious about it. It has multititudes of “to-do lists,” or “stubs,” where volunteers have posted pages that need to be categorized, linked, completed, and referenced, ready for any eager contributor to take up.

The long tail

Three features of Wikipedia get attention from popular commentators. First, admins spend time eliminating vanity entries—from teenagers, musicians, athletes, and self-aware politicians. Second, the entries for pop stars Britney Spears and Michael Jackson, as well as those for Hitler, George Bush, Jesus, and Mohammed, ranked among the most frequently edited in 2006. Third, the site has a lot of entries about sci-fi.

This is all true, and in my humble opinion, mostly irrelevant. It is like assessing the quality of a grocery store’s produce, dairy, and meat sections by whether the store carries the National Enquirer. The presence of the eye-catching stuff does not affect what really matters.

The site actually does contain plenty of intellectual meat and potatoes. Here is a shorthand rule: If it might appear on a standardized college entrance exam, it’s in Wikipedia. These basic entries do not differ from those in a printed encyclopedia: They offer basic history, science, geography, and politics.

Some commentators have worried about and even tested the accuracy of these basic entries. While this concern might have been valid several years ago, at this point I think it is misplaced. After several years of editing, it is hard to mess up the presentation of the Pythagorean theorem or the habitat of Emperor Penguins.

One feature of the site is more novel and deserves closer scrutiny: Wikipedia faces no binding hard-disk space constraints. In brief, no editor need cut the details about British royalty to save space for great personal qualities and an abundance of professional accomplishments, Huchra cannot enunciate Shakespearean dialogue like a television star. So Huchra has a Wikipedia entry of fewer than one hundred words. While his entry also has a link to his Web site as well as to several other relevant entries on the Great Wall, there is no fan base improving the entry every week.

Huchra is an unassuming man, so I am sure he could care less. Yet, something seems amiss: His discovery is too important to receive an entry less extensive than one devoted to the Borg, which is a fictional space foe, after all.

What I am trying to say is this: Wikipedia’s long tail is better at being wide than deep. For example, as a professional economist I could find fault with some of the entries that go beyond Econ 101. As an Internet business specialist, I see many biographies of Internet movers who suffer the same celebrity biases evident with space explorers. Other specialists make similar remarks about their fields of knowledge, whether it is the scientific basis for biotechnology or the origins of the Hapsburgs. It just depends. Sometimes a group got a passage together, and sometimes not.

What is it good for?

Teachers and parents who worry about Wikipedia becoming an authoritative source have a valid concern. Kids easily can use it improperly. And it is correct but not helpful to say that most admins would not advise any sane adult to use Wikipedia as an exclusive source.

I want to convey that observation with the right attitude. I find myself tiring of those who moan about Wikipedia’s faults. It is actually a pretty good resource if supplemented with outside material.

More to the point, everyone is an expert on something. Experts do not have to depend on Wikipedia, but others do. If you see an error, don’t whine. Just fix it.

For more information on this or any other computing topic, please visit our Digital Library at http://computer.org/publications/dlib.