

Science

AAAS

Magazine

News

STKE

Careers

Multimedia

Collections

All Free Articles

Top 10 Last Month

ScienceShots

Daily News Archive

About ScienceNOW

[Home](#) > [News](#) > [Daily News Archive](#) > [2007](#) > [January](#) > 23 January (Bohannon)

Folk Wisdom for Web Sites

By John Bohannon
ScienceNOW Daily News
 23 January 2007

Web sites such as Del.icio.us (<http://del.icio.us>) and Connotea (www.connotea.org) organize thousands of links by various topics, all without any expert guidance. How so much information gets tidily categorized is a mystery. A study of how visitors to these sites create descriptions of links suggests that they may be following simple rules without realizing it.

One of the hottest areas for social scientists studying online interactions is so-called folksonomies. A taxonomy is a categorization system created by a group of experts, but a

folksonomy is created spontaneously by millions of strangers through a process called tagging. When users of folksonomy-based sites post a link to a Web resource they like, they must "tag" it with a string of descriptive words. For example, to post a link about circus skills, a user might include the tags "circus," "juggling," and "acrobatics."

What amazes scientists is that this simple, uncoordinated system works so well. In spite of the enormous number of alternative words that can be used as tags, people tend to converge on a narrow vocabulary. Plotting the rank of the words versus their frequency reveals that tag vocabulary, like word usage in languages such as English, follows a power law. The most common word occurs twice as often as the next most common, which occurs twice as often as the next, and so on.

To probe the phenomenon, a team led by Vittorio Loreto, a physicist at University of Rome "La Sapienza" in Italy used a simple computer model to create an artificial folksonomy. The program simulated a group of virtual users who add tags to a pool of existing descriptions. They had a fixed chance of making one up or copying one already in the pool. To compare the model's output with real folksonomies, the researchers downloaded tag-related data from Del.icio.us and Connotea.

Although the artificial folksonomy was created by mindless agents that followed simple rules and had limited information, its structure was close to the real ones. For example, the set of tags within the 37,974 posts on Del.icio.us that contained the label "blog" fit the power law curve, and tags from the virtual folksonomy

ADVERTISEMENT

matched it as well, the team reports today in *Proceedings of the National Academy of Sciences*.

Don Steiny, a computer scientist and consultant in Santa Cruz, California, is intrigued that simple rules can mimic the self-organizing behavior of folksonomies. But before believing that humans are following these rules, he says, researchers need to study how people actually create tags.

Related site

- [About folksonomy](#)

[To Advertise](#) [Find Products](#)

[Magazine](#) | [News](#) | [STKE](#) | [Careers](#) | [Multimedia](#) | [Collections](#) | [Help](#) | [Site Map](#)
[Subscribe](#) | [Feedback](#) | [Privacy / Legal](#) | [About Us](#) | [Advertise With Us](#) | [Contact Us](#)

© 2007 American Association for the Advancement of Science. All Rights Reserved.
AAAS is a partner of HINARI, AGORA, PatientInform, CrossRef, and COUNTER.