

THE HIDDEN BIOSPHERE: Is There Life Beneath the Ocean Floor?

# SCIENTIFIC AMERICAN

JUNE 2000

www.sciam.com

## THE BIRTH OF MOLECULAR ELECTRONICS

DWARF GALAXIES  
AND STARBURSTS

THE NETWORK  
INSIDE A CELL

SPECIAL REPORT:  
THE NEW FACE OF  
WAR



\$4.95



# Accounting for Taste

**O**AKLAND—Personal taste is notoriously tricky to quantify. Opinions are subtle. Non-linear. And just barely associative. "Taste is idiosyncratic," says Ken Y. Goldberg of the University of California at Berkeley, who studies the subject as a computer software problem. "The best example is that you don't always like all of your friends' friends."

But most of the time, you do. And thanks more to reason than RAM, one long-hyped method of automated recommendation is finally proving itself. Called collaborative filtering, it predicts individual preferences based on the preferences of others. Amazon.com rolled out one of the first commercial applications of collaborative filtering in 1997, recommending books that your nearest neighbors in taste, as determined by their click history, have bought. Its filtering engine was designed by Net Perceptions, which also built CD-Now.com's system. Net Perceptions was co-founded by the "father of collaborative filtering," John Riedl, a computer scientist at the University of Minnesota, who in 1994 co-authored a paper on the collaborative filtering of newsgroup postings.

Although the first publicly accessible academic experiments were novel and showed promise, it has taken some time to shake out the kinks. All too often, Amazon's book suggestions proved to be so general or off-base that you'd have had better luck throwing darts at the *New York Times Book Review*. Firefly, an early music recommender, was fun to fiddle with, but you were just as well off (if not better) chatting up the music fiend behind the counter at the nearest Tower Records.

Naturally, the engines grew smarter as ever more Internet users fed them with data. But the software engineers grew smarter as well, developing novel algorithms, customization features and more user-friendly interfaces. An important innovation was tuning the engines. An early customer, Riedl recalls, was an on-line grocer that expected collaborative filtering to expand the scope of what customers put in their shopping carts. Not quite. "They called and said, 'We don't need your fancy software package to tell us that our customers like bananas,'" Riedl says. The solution was to enable clients to adjust the

software themselves to recognize items that are already big sellers and "recommend others that are more of a surprise."

Now researchers are pushing personalization further. In Riedl's university lab, Jon Herlocker invented a feature to appear on the MovieLens site, which will translate the reasoning behind a recommendation into a language the user can understand and respond to. For instance, MovieLens might recommend *Titanic* if your neighbor in the profile database enjoyed it. Then, if you watch it and give it a thumbs down, MovieLens will provide you with the option of shutting out the opinions of that anonymous neighbor.

Unfortunately, most people haven't



used sites enough for their profiles to be sufficiently developed, says Dan Greening of Macromedia eBusiness Solutions, makers of the LikeMinds collaborative filter engines used at Levis.com, WeddingNetwork.com and other sites. The key, Greening believes, is that his software is elitist when determining who is dropped in the "mentor pool" of user profiles that are actually mined for recommendations. Good mentors have rated many things over a wide spectrum, making them general "opinion leaders." But if they also prove to be good mentors for other mentors in the pool, the lesser candidates will be flushed out.

While Greening has been coding the makings of a good mentor, Goldberg and his colleagues have taken a different approach, using pending patents accrued

from their joke-recommending site, Jester. They founded PreferenceMetrics; its demonstration site, Sleeper, is eerily accurate at recommending books based on ratings of books users may not have even read. The site polls you on your level of interest in a particular book, given a brief description. Accuracy is also increased because your user profile is determined only by the ratings you actively provide; other sites don't distinguish between items you buy for yourself and those you choose for others.

Sleeper's recommendations are based on an algorithm that employs a mathematical technique called principal component analyses to lower the number of variables, or dimensionality, of the problem. That speeds up the software's recommendation process without compromising accuracy, according to Goldberg.

But the unique and most noticeable element in Sleeper is its continuous rating bar. Traditionally users pick from a five-level rating system, like a newspaper's movie reviews. Goldberg's rating bar spans from "very interested" to "not interested," enabling the user to click anywhere in between. The computer translates the clicked position into a number between 1 and 500. Taste is more visceral than rational, Goldberg says, and "moving the mouse along the bar feels a lot more kinesthetic than the rational process of clicking on buttons."

Most of today's collaborative-filtering Web sites are based on "personalizing" a retailer's relationship with a customer because, as Riedl bluntly puts it, "that's where the money is." But Riedl, along with Greening and Goldberg, are optimistic that as the technology continues to improve, myriad applications will follow. They predict that their brainchild will imminently return full circle to its roots as an information filter and become, Riedl maintains, "one of the most important changes in the way information is disseminated." Goldberg agrees, pointing out that customization of what you see on your monitor is increasingly mandatory as the screens on emerging Internet portals, cellular phones and wearable computers continue to shrink.

Yet whatever the access point is, one of the ultimate hopes of collaborative filtering is that on-line individuals will each have their own intelligent agents, crawling the network and seeking out news you can use before you even ask for it. After all, in some sense, your agent may know you even better than you do. —David Pescovitz