



Side Lines

If You Have Kinky Tastes

William Baldwin, 01.09.06

If You Have Kinky Tastes You might find it creepy how much marketers know about you and what you buy. See Helen Coster's "[Consumer Spy](#)". Get ready for far more intrusiveness as your shopping migrates from printed catalogs to electronic ones. Now L.L. Bean is going to know not just what you bought but how long your eyes lingered on each image.

Jeff P. Bezos has made this kind of data mining into an art form. Go online to check out a particular massage showerhead and Amazon will suggest that you also consider a Nat King Cole CD, reruns of *Moonlighting*, and TurboTax. The connections are a bit obscure, perhaps. But this computer is very smart. If you are looking for John Wayne movies, it knows not to serve up a recommendation for *Midnight Cowboy*.

What connects product X to product Y? Don't ask Amazon to tell you how the software works. Whether from fear of offending people squeamish about privacy, or from some desire to lend a mystique to its merchandising, the firm is extremely cagey about its methods. It refuses even to reveal how many mathematicians it has working on the problem.

In the outside world, though, information is leaking out about what these geeks are up to. Bruce D'Ambrosio, a former professor at Oregon State, is now the chief technology officer for Cleverset, a Seattle firm helping lesser Web vendors emulate Amazon's skill at anticipating customers' needs. The science is part graph theory, he says--the six-degrees-of-separation riddle--and part Bayesian statistics, which describes the connectedness of two products (or two customers) in probability terms. What is the a priori probability that you like Cole? How does the probability change if we also know you are into showerheads? Also relevant, according to Michael W. Berry at the University of Tennessee: text mining, the automated detection of semantic connections between two documents.

Andreas Weigend, a former chief scientist at Amazon who now teaches data mining at Stanford, lets out one little secret: The difference between a good online marketer and a great one is not in the algorithms, which are getting to be well known, but in the richness of the data fed into them. In a recent interview with a German magazine he said that Amazon collects a terabyte of data on a busy day. Emulating Amazon, all manner of vendors will be watching your every move. Just to serve you better.