



The Atomization of Information

As an infopreneur, I often have to think about how I price my services; rather than simply estimating the number of hours that a particular service requires, I also look at the value of that service to my client. Of course, I have to make sure I'm compensated for all the hours I put into a project; X hours at \$Y/hour gives me a baseline price. Then I factor in the outcome of my work—the better-informed decision made, the workshop participants' improved search skills, the new market identified—and adjust my final price accordingly.

The professional econtent providers are also trying to figure out how to price their services in a way that reflects the value provided and fairly compensates them for their costs. Longtime online searchers remember the CPU-based pricing of LexisNexis and Dialog, the per-line charge of Dow Jones News/Retrieval, and the common practice of connect-time fees—all methods the online services used to assign a price to the value provided. While most enterprise subscriptions are flat-fee now, negotiations are still often based on a calculation of the retail value of articles downloaded.

However, this pricing structure assumes that the primary value of the service is delivery of the full text of licensed content. Increasingly, data-intensive organizations are seeing value beyond “mere” full text. Knowledge managers and data scientists are building or buying tools for semantic enrichment of content to enhance interoperability with internal decision tools. Information center managers are embracing text and data mining initiatives that require access to large datasets of licensed content—and they are finding that their econtent providers have not yet figured out how to price this type of access in a way that is cost-effective. When online vendors place all the value on full-text content, they fail to see the value of many tiny portions of their content.

I started thinking about the atomization of information, and, just as when you're thinking about green cars, you notice that green cars are everywhere, I've noticed instances—and consequences—of atomized information wherever I turn.

Take Google's “zero-click” search results, the goal of which is to provide the answer to a query without the user needing to click any of the hyperlinked results. Google's Knowledge Panels, Featured Snippets, and other “rich results” are drawn from structured data that tell the search engine *about* the content on that page. Not only do these no-click results decontextualize information by presenting only small extracts from a larger body of information, they enable smart speakers to provide a single answer to a spoken query. While this

feature is convenient when asking a factual question, nuances and alternate answers are lost. As an example of the weakness of Google's zero-click results, try Googling **who said If you want to test a man's character, give him power**. While the quote is often attributed to Abraham Lincoln, according to Snopes, it was actually said about Lincoln by writer and orator Robert Ingersoll almost 2 decades after Lincoln's death. In fact, Google's first organic search result is often the Snopes article. But the featured snippet appearing just above the Snopes debunk? Yes, Google's “answer” to the query is to attribute the quote to Lincoln. (See my September/October 2019 column, “Would You Trust a Free Taxi?,” for more about the impact of zero-click results.)

I've noticed another example of the consequence of atomized information in Amazon's search results. Recently, I have seen a proliferation of items being identified as “Amazon's Choice.” While the algorithms behind these designations are proprietary, Amazon suggests that it uses factors such as product popularity and customer reviews. *The Wall Street Journal* published an analysis of products designated with this apparent endorsement and found items that violated Amazon's policies and the listings for which appeared to have been designed to game the algorithm. (See “Amazon's Choice' Isn't the Endorsement It Appears,” *The Wall Street Journal*, Dec. 22, 2019.)

As disturbing as this is, what alarmed me more was that the Amazon's Choice designations are used when a customer uses an Amazon Echo smart speaker to order a product. If you ask Alexa to purchase an item, Alexa first checks your Amazon order history; if you haven't ordered that item in the past, Alexa recommends an Amazon's Choice product for you. Without a screen to show you alternative products, you do not have an opportunity to evaluate options, and, as the *Journal* article found, you might wind up purchasing an unsafe or counterfeit product.

Increasingly, the job of info pros is to help our users identify and understand the impact of atomized information in casual search engine queries and to see the value of atomized information in econtent. As information becomes more fragmented, we must lead the conversation about both the resulting benefits and pitfalls.

Mary Ellen Bates (mbates@BatesInfo.com, *Reluctant-Entrepreneur.com*) never trusts a quick answer.

Comments? Email the editor-in-chief (marydee@xmission.com).