

# Avoiding the idiot savant of search

*The purpose of searching has changed. We need to adapt to that change.*

Philip C. Murray  
Chief Knowledge Architect  
The Semantic Advantage  
phil.murray@semanticadvantage.com

An Internet search engine is a smelly, leering idiot savant who remembers everything but can't tell you what's important. Just sits there farting and belching until you ask a question in a way that he understands it. Except he doesn't understand your question at all. He just recognizes that your question has some of the same superficial properties of language as stuff in his memory. When he's done remembering, he vomits results. On you.

And you thank him. And you reduce yourself to his level of behavior again and again. Because, after all, he "knows" everything, right?

So, why did we all reduce ourselves to his level of behavior? Well, the teacher told us to. The amount of information skyrocketed, and Ms. McGuffey reduced importance and "results" to an issue of memory. Speed and scope of recall got all the rewards of being the teacher's pet.

But work doesn't work that way. We are, indeed, using far more information in our work — and we have to accommodate that new reality. That doesn't mean we should fundamentally change or oversimplify our understanding of work itself. That's a lot like saying we should focus our entire educational system on driver education because there are more roads than ever.

Think about it: One person searching all the information in the world and expecting instant, meaningful results. That's silly. Having multiple employees in an enterprise do that for much of their work days — without a formal understanding of how that activity can be performed well and no understanding all of how the results of that activity become value — that's ludicrous! Now who's the idiot?

## Putting the idiot back in the village

The explosion of information, an accelerating pace of change, new technology, and the reshuffling/redefinition of job roles led to a pervasive disruption of traditional methods of turning information and experience into value. But instead of examining how these new forces affected our work, we ignored, deprecated, or discarded the value of those traditional methods. We chose the idiot savant of global search because he was there, because it was easy.

We have to look at work itself more closely. How, exactly, does information become value? Increasingly, that's the domain of "semantic activities." Of course we need "search." How else are you going to find *Mxyzptlk*? Collecting information obviously does have value: It produces a more manageable subset of all possible information. But how well did you conduct your search? How easily can you find the information you discovered in your own new personal library — or on the Internet — the next time around? And — from the perspective of the enterprise — does that personal collection produce some value for the enterprise?



The answers to those questions usually are: “poorly,” “with great difficulty,” and “only accidentally.” Those aspects of “searching” don’t get the attention they deserve because we don’t recognize them as essential core activities of business itself. We view “searching” alternately as a new kind of freedom or as an act of desperation.

## Deconstructing semantic activities

The impact of the Information Age isn’t limited to superabundance of information and the need for powerful, global search. Information still has to undergo a transformation before it becomes value in the form of products, services, and new information — as it **always** has. There are important new considerations, too. We have expanded the processes of communicating and managing information, changed the way we arrive at decisions, and created many new layers of information between original idea and execution.

You can’t address these problems just by making each individual “search” activity faster or bigger. A quick look at the impact of incremental improvements in search engines demonstrates that failing. The more appropriate response is to make the chain of research, reasoning, and decision making explicit and carefully deconstruct those activities. It’s not simply a matter of overpowering the increased volume of information. We need to stop entertaining and honoring the idiot savant!

The first step is to look at how we used to do this stuff — typically when researching topics in libraries. Then we can create an understanding of how that differs from what we do now, often as “knowledge workers” in business environments. If we artificially limit the scope of our discussion to *searching for information on a topic (as an individual in isolation)*, we can see that these processes may include:

- Formulating a query
- Referring to subject, author, and title entries in a library catalog
- Examining cross references in those entries
- Selecting and recording the identities of documents that seem relevant
- Opening those resources to scan them for relevant information
  - Scanning the table of contents and index, if they are available
  - Scanning the content itself
- Exploring documents that are physically close — for example, the books next to one of the books we located in the catalog
- Asking a librarian for guidance
- Recording information about the document (author, title, etc.) for those documents you have judged to be relevant
- Copying sections of the documents into a personal resource — a set of notes
- Appending comments to those snippets
- Organizing those snippets

Even without a more detailed analysis of what happens at each step, the process is moderately complex. It’s the classic model of what we do when conducting research in an academic setting, and it’s based essentially on a paper-document paradigm, even though the library catalog went online two decades ago.

## What’s changed ... and what hasn’t

We haven’t changed that model too much, even as we moved those documents online ... and added billions more documents on the Internet. The biggest differences are (1) the size of the resource being searched, (2) the accessibility of content itself to search engines, (3) the number of searches executed against those resources, (4) the absence of a single, authoritative top-down categorization for those resources (The library catalog became a search engine and directories like Yahoo! and DMOZ.), (5) the lack of vetting for the vast majority of those



resources, and (6) the proliferation of personal digital resources — copies of online information often amounting to gigabytes of data for each individual.

There are also potentially important, but as yet unrealized, differences in the online networked environment, including the implementation of highly expressive personal profiles that enable search engines to know us — much as our friends become familiar with our interests and expertise and avoid sending us information they know we don't want. (You can't always get you what you want, as the 'Stones say. Hell, you can't always **remember** what you want.) Such profiles even allow some resources to “push” relevant information to us.

The most ignored difference, however, may be accommodating the changed **purpose** of searching. Because in enterprise environments we really aren't doing “research” in the traditional sense. We're trying to turn information into money or results, turn the raw material of information into value. And we're trying to do it as a group, even though we can only act as individuals.

So we need to deconstruct our “search” activities into more explicit, discrete steps so that we can eliminate duplication, make individual aspects of searching easier, and apply new tools to the more structured information produced by these new understandings of how search affects results. That may seem like making a simple activity vastly more complex. It's not, and we seem to have ignored the fact that the “simple” activity of *researching a topic* on the Internet is actually so complex that the average person fails to produce effective results predictably from that activity. The Google search box only makes it seem simple. Why, then, should we expect **any** enterprise to produce value from the search activities of tens or thousands of individuals, who each perform such activities dozens of times a day?

Talk about idiots!

## When do knowledge-based enterprises succeed?

Knowledge-based enterprises succeed **not** because they perform these “research” activities well and with great frequency but because **they short-circuit these inappropriately defined activities as much as possible**. Successful enterprises rely on people who make better evaluations of the ideas in the information gathered by multiple workers. Those “deciders” are responsible for decreeing that some ideas (and the information associated with them) are not relevant to the purposes of the business ... or are especially important to those purposes. This historic part of the process of turning information into value — once typically seen in peer reviews and management fiat — is largely unaccounted for today. It's still there — often abetted in online forums and proliferating as information itself proliferates — but the process of evaluation is not acknowledged as a formal activity of subject-matter experts and managers ... and it doesn't produce an explicit chain of results and refinements that you can see, measure, re-use, and build upon.

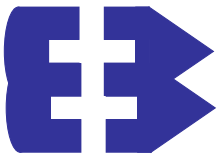
Such hidden or informal evaluation processes do not create a predictable, scalable competitive advantage, especially when the amount of information over which these judgments must be exercised is growing rapidly. So starting from square one — deconstructing the new processes of turning information into value — isn't simply advisable; it's an utter necessity. “Square one” includes transferring the focus of our knowledge-based activities from *information* to *representations of enterprise realities*. Because interposing vast amounts of unstructured (essentially meaningless without interpretation) information between ideas and results is no way to do business.

For this reason, the recent obsession with “metadata” is a bit distracting. Metadata, in the broad sense of *information about information*, is still, er, information. It is information with a good heart and a college education. Some refinement. It wants to help us locate its unkempt charges, those unstructured documents. But it is an idiot savant with a pedigree, and it still vomits information on you. Just less of it. Better chunks, maybe. It still focuses on your interaction as an individual with the whole world's knowledge.



The good news is that we're beginning to see semantic tools and management strategies that respond to the need for extracting, measuring, retaining, and managing the meaning within information. The bad news is that we may be subjected to a new, extended round of isolated, specialized applications for those purposes before we see the broad business benefits that the Information Age really needs.

---



The Semantic Advantage  
3 Sullivan Terrace  
Barrington, RI 02806  
401-247-7899

Blog: <http://semanticadvantage.wordpress.com/>  
Web site: <http://www.SemanticAdvantage.com>

*Turning Information into Assets*