

www.futurebrief.com

I Don't Think So!

Jeffrey R. Harrow
Principal Technologist, The Harrow Group

[\(download PDF version\)](#)
[\(read his bio\)](#)

The Alta Vistas, Googles, Yahoos, MetaCrawlers, and the other Web search engines are (perhaps) second only to Email as the 'killer apps' of the Internet. There are petabytes (one petabyte is one-million gigabytes) of information scattered throughout both the surface and the "depths" of the Web ("deep" info is "hidden" in online-accessible databases, but which is not normally visible to typical search engines). Yet without today's "surface" searching, the Web is so large that it would be virtually impossible for us to know where to begin to scratch our info- itches!

Indeed, the volume and types of Web-based information are growing by leaps and bounds, expanding into some surprising areas. For example, Google has worked out [deals](#) with some of the most prestigious libraries in the world (including Oxford's, Harvard's, Stanford's, the New York Public Library, and others) to digitize a vast number of rare and not so rare books, making them Googleable.

("To Google" something is already a recognized verb (although Google's lawyers are [less than thrilled](#)), so why can't I create another variant as well?).

This "Google Print" [program](#) represents an inflection point analogous to the invention of the printing press, since for the first time this project will easily bring content, which had historically been accessible only to those who studied and taught in those hallowed halls, to the masses.

Given the tremendous amount of information that is already available at our fingertips, plus projects such as Google Print, it might seem that Search technology can already do most of what we need. Indeed, according to Raul Valdez-Perez, founder and chief executive of Vivisimo Inc. and a computer scientist at Carnegie Mellon University, he states in the Aug. 18, 2004 [New York Times](#):

"I would say that search technology on the Web is pretty mature..."

BUT...

But -- I don't think so!

Today's search engines are not perfect or, in my opinion, "mature" by any means. Useless or non-authoritative "hits" are a constant occurrence, and it's often frustrating to come up with just "the right" search terms to yield what we want. Nevertheless, today's search engines are already enormously empowering, and text-based searches will continue to improve.

Yet I believe that search technologies will also have to move in important new directions.

The Big Picture.

For example, here's one particular area that leaves an enormous amount of information that's already online, untapped.

www.futurebrief.com

I have the antithesis of a "green thumb." I know next to nothing about gardens and plants, yet I do have a small garden behind my house.

The problem is that it's becoming overgrown with weeds, yet I can't safely pull them out. Because by looking, I don't know which green things are weeds, and which are the flowers and herbs that are supposed to be growing. Since I don't know their names to start with, and I don't have any other identification words that I can type into a search engine, I'm relegated to finding a knowledgeable human neighbor who can simply look at the garden and tell me what's what. The Web (currently) isn't much help for the vast array of similar visual identification tasks.

What a waste of the vast resources of the Internet, since I'm sure that the information is out there in the form of pictures, probably captioned, that could tell me what I want to know. But what I really need is the ability to take a photo of my garden and show it to a search engine. It should then recognize the growing things in the picture, tell me what they are, and then take me to the specific resources that can turn my Internet-connected thumb a bit greener.

Or, how about being able to search the incredibly large amount of audio that is generated every day? Or find just the video footage you need because the search engine would recognize and index the video?

These are just a few of those archetypal "non-trivial tasks," but solutions would be SO valuable for such a wide array of personal and business and industry and governmental uses, that it deserves (and I believe will get) significant attention.

Even MORE Useful.

I don't have a solution for Google, Yahoo! or the others, but such search engine extensions would dramatically (again) extend the usefulness and value of the Web. And handled correctly, their just might be (a lot) of money to be made from such a capability, even if it's freely given away for personal use while charging businesses -- essentially today's model.

Given the flurry of attention going to "machine recognition" of various media for security and industrial and scientific purposes, and our always-increasing amounts of processing power and memory and storage space, ideas for accomplishing these new recognition and perception and indexing tasks should be regularly coming to light. Wouldn't it be something if a simple digital picture of, say, a specific building, could return a Google's worth of information?

Of course that picture, when applied to a picture of "you," illustrates the potential for a significant "dark side" -- from stalking to Big Brother to far more. We already, though, have to deal with this issue even with today's technology -- just try typing your name into your favorite search engine!

Although it might seem desirable to keep such technologies from coming to market, we continuously learn that technology and innovation genies can *never* be stuffed back into their bottles -- if one developer or country does repress such advances, others will surely take the lead.

Non-textual searching is still an area where humans rule. But it doesn't have to (and it won't!) remain that way forever. The day will come when I can send that picture of my garden, and I'll receive meaningful results. (Of course in the case of my garden, it might well tell me to till the garden under and plant grass.)

Now -- imagine if you or your business or agency is the one who meets these new search challenges...!

This essay is original and was specifically prepared for publication at Future Brief. A brief biography of Jeff Harrow can be found at our main [Commentary](#) page. Other essays written by Jeff Harrow can be found at his [web site](#). Jeff receives e-mail at jeff@theharrowgroup.com. Other websites are welcome to link to this essay, with proper



www.futurebrief.com

credit given to Future Brief and Mr. Harrow. This page will remain posted on the Internet indefinitely at this web address to provide a stable page for those linking to it.

To download a PDF version of this essay, [click here](#). Please feel free to share the PDF with others who may be interested. To hear about future **Commentary** essays, take a few seconds to read about [Daily Brief](#), one of the "briefest" Internet updates offered anywhere.

© 2005, Jeffrey Harrow, all rights reserved.