

New Tricks for Keeping Philosophically Fit

by
John C. Fakan

The Philosophical Club of Cleveland
9 March 1999

A **not** so infamous date in history is December 7, 1938 -- the date that an unusual *Special Edition* of the St. Louis Post-Dispatch made its debut. It carried no advertising, made no money at the newsstands -- and, in fact, only 15 copies of each edition were printed.

But, it was a big success. It was featured in a Paramount newsreel and shown in 6,000 theaters across the nation, and was featured in 20 foreign countries. Franklin Delano Roosevelt's son, Elliot, pronounced it "a very good advertising medium".

The Special Edition was not printed at the plant nor delivered to newsstands, but was sent via radio from station KSD's studios to 15 special receivers with built-in printers.

This "radio newspaper" as it was commonly known (the name "fax" hadn't yet been coined) was predicted to become as much a part of American life as had the radio and as television soon would become. The predictions promised something for everyone in the family including today's news and current events, general interest stories, weather reports, and even printed comics for the kids (of any age).

The Post-Dispatch system was just the first of a number of attempts at meeting the desire to make current information universally and easily available right in our homes. Everyone was getting into the act. David Sarnoff, president of RCA, previewed the future of radio broadcasting at a 1940 shareholder's meeting, and said, "Three new services beckon those who seek to expand radio's usefulness. They are Facsimile, frequency modulation (FM radio), and television".

A Scottish electrician and clockmaker, Alexander Bain, patented facsimile transmission in 1843. His system was largely mechanical and re-created images of specially prepared originals formed with tiny raised metal blocks in a wooden frame. In 1902 the first practical system for transmitting images directly from the original document or photograph was demonstrated in Germany by Dr. Arthur Korn. This system used a selenium-based light-sensitive scanner -- as still used in most copy machines today. By 1926 the Newark, New Jersey radio station, WOR, owned by the Bamberger Department Store, started sending photographs via fax. The READO, a facsimile system designed by William G. H. Finch and built by Powell Crosley, owner of the Cincinnati Reds, was featured at the 1939 Worlds Fair.

During World War II facsimile development was put on hold. Most of the developers got into the development of other systems for the war effort. In the Fall of 1944 WOR's president, Theodore Streibert launched an effort to get facsimile back in front of the American public as soon as possible, " . . . so as to determine the type of service and editorial content most useful to the public in home broadcasting".

The problem then was that Facsimile was hard to justify. If people wanted late breaking news bulletins why not just send them via conventional radio? Proponents of Facsimile argued that people wanted their news in tangible form -- hard copy. They were wrong. Over the period from 1950 through 1997 the number of evening newspapers dropped from 1,450 to just 816, and circulation dropped by more than two-thirds. Locally the *Press* and the *Cleveland News* were victims of that period. The morning papers proved more durable. The evening news shifted to television.

Curiously, the trend to television was unpredicted. From an article in *Current History*: "...in evaluating the radio (or television) as an entity, critics state that it violates the basic laws of knowledge -- i.e. (it does not provide) a permanent record and continuity. For in the history of humanity it is never the spoken word, the mouth to mouth method of communication still practiced by many backward tribes, but rather the inscribed tablet embodying the written word that has built column by column the knowledge to which we now have access. But, if this is the criticism suffered by radio in the past, competent people say that facsimile broadcast will make it valid in the future."

Another major advantage of Facsimile was the freedom it provided its users. As one promoter wrote, "Facsimile releases you from your radio set; you do not have to be within hearing or sight of your receiver while the news is coming in." Broadcasters also saw this as an advantage. WOR's vice president wrote in 1946, "Radio broadcasters have always felt handicapped because their programs could be effectively received only when listeners were giving their attention to their radio sets". The newspaper's interest was more motivated by the fear of being left behind in the event this new paradigm caught on. Colonel Robert R. McCormick, editor and publisher of the *Chicago Tribune* said, when the paper launched its fax edition in 1946, "I do not know what facsimile is anymore than I knew what radio was 20 years ago, but we are going to find out all about it. There is no doubt that radio is constantly developing. FM, television and facsimile are all new. We can't resist these advances. We've got to go with them."

All of the dreams and aspirations for facsimile have come true, of course, but many years later and in a form no one appears to have imagined. Fax machines themselves remained in limited use until they took off in the 1980's -- thanks to technological advances that made them practical and affordable as a standard business tool. But the realization of the promise of information that is universally available to all has been brought to us by the Internet -- and more specifically by the World Wide Web.

The convenience of not having to be present when information is placed into the system in order to use it (a key feature of facsimile) is inherent in the design of the Internet. Once information is recorded -- or "published" in the common vernacular, it then becomes available thereafter. It does not disappear -- as does a radio or TV broadcast -- unless the publisher actually removes it.

I would argue that the appearance of the information on the cathode ray screen fulfills the requirement for information to be tangible if it is to be a part of *knowledge*. Granted the image disappears when a new image is displayed or the machine powered off, but it can be restored exactly as before whenever needed. It is not subject to the vagaries of memory to recall the spoken word or the recollection of an image seen at an earlier time on the television screen.

The first commercial facsimile newspaper transmissions were only available to the 15 families that had one of the special receivers and printers. Later systems were available to much larger audiences, but the carrying capacity, or bandwidth, of radio is actually quite limited. That is why Cleveland had only three television stations for quite a long while, and even with the use of the UHF channels the offerings were still quite limited. Cable television has managed to get around some of the limitation by creating its own "radio universe" within its shielded cables so that it can exploit the entire radio spectrum for its individual use without concern for other broadcasters who are using either the airwaves or their own closed systems.

The Internet, however, utilizes enclosed hardwired networks to carry information to and from users, so has the advantages of the cable systems, but it is far more ubiquitous because of the almost universal existence of telephone circuits. Such telephone circuits can function as a valid network within the Internet system. Actually, the particular vehicle used to transport Internet traffic is irrelevant to the system. The only requirement is that all such networks have at least one interconnection to the overall system. It is this characteristic that is the "web" in "World Wide Web".

If published information is to remain available for long periods of time -- or even forever -- a means for storage must be available in huge quantities and at very low cost. Modern technology has been taking care of this requirement very well right in step with the growth of the user population on the Internet. In a paper I presented a few years ago I told you that information doesn't weigh anything and doesn't take up any space. It is the medium that represents the information that determines the physical requirements for storing any given amount of information. For example, at that time we learned that the book, Moby Dick, could fit on a 3.5-inch floppy diskette with room to spare. The cost of that diskette was about \$1 and it easily fits in a shirt pocket. Today's 3.5-inch super-floppy costs just under \$10 but will store 120 copies of Melville's book and will still fit into the same shirt pocket. Thus we see that, at least for the highly portable floppy diskette we have gained about a factor of 10 in lowering the cost and a factor of 100 in size since that time. Hard drive technology has done even better. A 16-gigabyte drive has sufficient capacity to store 16,000 copies of Moby Dick and costs about \$320, or about 2-cents per copy.

What is the paper and ink cost of storing this work on printed pages? And, storage technology is still improving at a very fast pace.

Thus we can see that the cost of storing information for availability to users of the Internet is no longer an issue. Electronic digital storage is now the lowest cost means for information storage regardless of how that information is to be used.

Advances in the delivery of information in hardcopy form -- books, publications, newspaper, etc. -- have made next-day delivery available at reasonable cost and second or third-day delivery fairly standard. Air transportation and efficient, well-organized service organizations such as FedEx, UPS and others have made this happen. Even the United States Postal Service has been able to utilize these same means to move the public mail in good order. But, your postal address among Internet users is referred to as your "snail-mail" address -- and this says it all. When

Internet users use the unmodified noun *address* they are referring to your e-mail address or POP (*point of presence* on the Internet). Times have indeed changed.

With storage cost and capacity no longer an issue, and access times essentially instantaneous there remain about three significant challenges to fulfilling the dream of the early facsimile pioneers. These include

- providing everyone with a point of connection (port) to the Internet,
- an effective and efficient means for cataloging and indexing the vast quantities of information,
- and finally, the user training and experience required to navigate this immense and rapidly growing warehouse of knowledge.

So, where do **you** look for information nowadays?

Just a few short years ago the top of the list of answers to that question would have included the library, technical publications, the media and books -- and, perhaps, a few others. Ask that question of today's college age generation and those who have matriculated over the last half-dozen years and the source at the head of the list would be the *Internet*. Nothing else would get many votes.

Have you noticed the list of required supplies for the student heading off to his freshman year at the college or university? Almost certain to be on the list will be a computer of some kind -- probably a laptop --, and some form of network interface device or modem. Our schools have already made the change to the new methods for collecting and distributing information. And, that is just what the college generation expects. It is what they are already used to, and they have an advantage none of us ever had.

It should be noted that today's college student has never even seen a black & white TV set -- except perhaps in a museum.

So what's an old dog to do?

How many of you do have access to the Internet at some level?

One in Two US Households Own a PC

Feb 19 1999: At the end of 1998, 50 percent of US Households owned a PC, according to Dataquest. This compares to 43 percent of US households in 1997, 36 percent in 1996 and 27 percent in 1995. Falling prices have had a major impact on PC penetration rates in the US consumer market, with an increasing perception of the PC as a standard household appliance, according to Van Baker, director of consumer market research at Dataquest. While the report found that PC ownership has increased across the board, the number of first time buyers in the lower-income brackets had increased significantly.

How many feel comfortable with it?

Who regularly uses e-mail as a means of communicating with others?

To whom do you still send snail-mail?

How much has that changed in, say, just the past year?

Who among you considers the Internet your PRIMARY source of information?

Finding Information on the World Wide Web:

Information on the World Wide Web (WWW) is published in the form of websites. A website is a specific place in the WWW universe, and it has a specific purpose. Think of it as a specific information resource. Every website is identified with a unique address known as a Uniform Resource Locator or URL (pronounced like the name "Earl"). If you know the URL of a particular website you can access it directly and make use of the information published therein.

But what if you don't know where such information is to be found?

Is there anyone here who has not heard of Yahoo.com? Yahoo is known as a 'search engine'. It is actually just another website, and its specific purpose is to list other websites and provide a brief summary of the information published on each. The big challenge, though, is that the WWW contains a far greater number of websites than the number of books found in even the biggest library in the world. (eg: In Jan, 99 it was 43,230,000.) So how do all these references get entered and catalogued? A major function of a search engine, such as Yahoo, is to literally "visit" individual websites and "read" their information, produce a summary, and then publish that information on the search engine website. Quite an undertaking! Yahoo's computer center includes computers and programs that 'snoop' through the many sites on the WWW and produce the summaries. This process is completely automated, so, the information on the search engine website is being constantly expanded and modified.

The primary principle behind the cataloging on today's search engines is the presence of specific words. Automatic programs that gather the information are not yet able to evaluate context, so the person doing the looking has to provide the logical evaluation and guide the search.

Another challenge, or perhaps opportunity, comes because we live in America. There are more than one schools of thought on how a search engine should set its priorities and on how search criteria are evaluated. In fact each of the search engines will have its own 'personality' and will provide different results for the same query.

Lets try the same search on different search engines to see what happens.

Let's see what we can find about the Philosophical Club of Cleveland.

We know that the philosophy club doesn't have a website so we don't expect to find one. However, others may be mentioning us and so we might show up in the listings.

If the club's name, "Philosophical Club of Cleveland" is entered as the search field (with and without quotes" this is what I found a few days ago:

Yahoo:

without	670
with	0

Web Crawler:

without	82,749
with	0

Excite:

without	1,251,878
with	0

Infoseek:

without	651
with	2

The first of the two hits on Excite was an Events Calendar that showed we are meeting this evening in this very room!

The second was the CV for our good friend Larry Landsroner.

Sometime within the past month or so Yahoo made a major change in the way they parse and process information. Last Fall I happened to enter a search for the USS COD Submarine. At that time Yahoo found over 32,000 hits. A few days ago the exact same search found **one** hit -- and that was the website for COD.

Web Crawler finds 7,818 hits and the COD website is not among the first 25 -- I didn't look further. If quotes are placed around the search words -- meaning that the three words have to be found together the result is zero hits.

The same search on the Excite search engine provides 131,847 hits without the quotes and 4 hits with quotes. The 4 hits were all pertinent.

Infoseek finds 567 without quotes and 15 with. The 15 are all good hits.

So the lesson here is that the search engines use different "rules" for how they interpret and process both the websites on the WWW and the search criteria. They also differ in the way their "Advanced" search option works.

In the "Advanced" searches the user may enter combinations of phrases and/or individual words, and also specify where these words and phrases are to be found -- for example: in the title, in the URL, or in the document. You can also specify that the search is to only consider websites that fit into a category such as: Automotive, Business, Career Opportunities, Travel, etc. Finally you

can limit the search to websites within a country and/or type such as: .com, .org, .gov, .edu, or .mil.

Some search engines, such as Infoseek, allow you to establish your own subset within which to conduct searches. This process can be nested so that you continue to slice and dice your way down to the specific websites of interest. For example a good friend was looking for a specific website that reported the results of an experimental study to determine the effects of a certain procedure in cardiology. She had attempted to find the site through various combinations of words and phrases but was constantly coming up with thousands of hits -- far too many to actually review. She finally gave up after about two frustrating hours. The next day I happened to be at her home and she told me about her wasted evening. We went back to the computer and she tried the procedure I was just describing above using Infoseek. In less than a minute she had narrowed the hits down to exactly one -- the website she was looking for!

To conduct this search she first had Infoseek grab everything that contained the word "cardiology" -- many thousands of sites. These results were then used as a category within which to search for the next most general word -- which I believe was "density". This resulted in a new category of websites that dealt with cardiology and mentioned the word "density". Just a few hundred. Her next search within that category produced a couple dozen, and the final search just the one she was after.

I find this last technique the surest means to chase down information. It allows me to apply my own ideas about the relative importance of the matching words -- that is, by ordering the words by their relative importance I can zoom in on the desired sites in a hurry with less chance of missing something that might otherwise be overlooked.

Another feature I find useful is the ability to have specific websites notify me whenever they are updated or whenever they contain information dealing with a specific subject. For example, because my son has been stationed in Monrovia, Liberia for the past three years I am very interested in any news that mentions Liberia. Accordingly, I have set up an automatic procedure that sends me any *Associated Press* or *Reuters* news report that contains the word "Liberia". These show up as e-mail. Using this technique I could actually design my own "newspaper". It would be a daily compilation of news dealing with only those specific areas of interest that I specify.

As newer algorithms are conceived and implemented in the various search engines even more exacting means for finding specific information will become available. As I mentioned earlier, none of these engines can relate to context matches though some are getting close. Just a few short years ago they were very crude and incomplete. They still have a long way to go, but the rate of improvement actually seems to be accelerating.

According to a quote from "Newsweek" magazine: "For something that has spread with all the forethought of Kudzu, the Internet isn't half bad." - Newsweek, 2/27/95

The difference between what today's Internet users can find on the WWW compared with what those 15 families had available back on December 7, 1938 is difficult to comprehend. Yet, it is

just the natural evolution resulting from our insatiable need for information or knowledge. And, to paraphrase what Colonel McCormick said back in 1946: "I have no idea where this Internet idea is going or what it will be good for -- but we are going to find out."

Thank you.