

Musings on Information Seeking Support Systems

Michael D. Levi
U.S. Bureau of Labor Statistics
2 Massachusetts Ave., NE Suite
4110
Washington, DC 20212
1-202-691-5900
levi.michael@bls.gov

1. A USE CASE

Even simple questions tend to get out of hand.

Let's say I have been following current events and I want to know "Is the US in a recession?" I start with the common sense notion that a recession is "hard times." A little searching on the Web gives me a more sophisticated definition: a recession is two consecutive quarters of negative growth in the gross domestic product. Further effort gives me the correct definition: in the United States, the National Bureau of Economic Research determines whether the economy has fallen into a recession, and they can only establish the beginning (and end) of business cycle phases after the fact. NBER hasn't made that call yet, so technically nobody knows whether the US is in a recession

That does not satisfy me. I keep poking around. I find plenty of economic pundits offering their opinions. Some say "yes," others say "no," some say "maybe," still others say "it doesn't matter what you call it, people are in trouble."

Being a data geek, I am intrigued by some of the economic indicators the pundits cite. I find the Bureau of Economic Analysis Web site and look at GDP for a while. Then I find the Bureau of Labor Statistics Web site and hunt for unemployment data. 5.5% in May, 2008. Is that good or bad? I locate a data retrieval tool and get a graph of unemployment rates for the last twenty years. I print it off, go back to the NBER site, and mark the dates of previous recessions onto the BLS graph.

During this exercise I remember having read in a magazine somewhere that BLS unemployment numbers are bogus. A little searching gives me lots of references to the article I kind of remember but the article itself is only available to subscribers. The charges strike me as overblown and possibly ill-informed. I go back to the BLS site to see if they have an explanation of how they calculate unemployment. I learn that BLS publishes six different unemployment rates, U1 through U6. I also learn that BLS publishes unemployment rates for the county in which I live.

Cool! I wonder if I can locate international data, too, and all of a sudden I want to know the principal export of Bolivia. The first page I find suggests it is probably coca. I don't know whether I can believe that. And I'm off in an entirely new direction...

2. CHARACTERISTICS OF EXPLORATION

The above is a plausible narrative of information seeking behavior. Starting with the same question, of course, I could have made a substantially different journey. If "Is the US in a recession?" really meant "Am I going to lose my job? Am I

going to lose my house? Will I be able to take care of my children?" then I might have spent my time looking at job postings, financial advice columns, or the Web site of a political party that promises to save me from destitution.

Some characteristics of open-ended, discovery-oriented exploration emerge:

1) I may not know, at the beginning, whether a seemingly straightforward line of inquiry will expand beyond recognition. Sometimes it will, sometimes it won't. A lot depends on my mood at any given moment.

2) I can't predict when the exploration will end. It may be when I'm satisfied that I have learned enough (which also would vary from day to day and query to query.) It may be when I get tired or bored. It may be when I've run out of time. Or it may be when I get distracted by dinner or the allure of the swimming pool.

3) I can't determine, objectively, whether the exploration has been a success. There is usually no "right answer" against which I can measure my progress.

4) My exploration is not a linear process. I could get interested in a tangent at any time from which I may not return. I am also likely to backtrack, possibly with some regularity, either because a tangent proved unfulfilling and I want to resume my original quest, or because I thought of a new question (or a new way of formulating a previous question) to direct at a resource I visited previously.

5) I am likely to want to combine, compare, or contrast information from multiple sources. One of those sources is my memory – which may or may not be reliable in any given circumstance.

In addition, some characteristics of the on-line medium itself also emerge:

1) Though an enormous quantity of information exists on the Web, an huge volume of material is not on the Web or is not accessible to all users.

2) The material that does exist is often of unknown quality and reliability.

3. MANUAL TOOLS

In the pre-Web days, if I was in a rigorous frame of mind when I began an information seeking exploration, I typically assembled and used the following set of tools to assist me in organizing materials and my thoughts:

- A pen
- A pad of paper
- Index cards
- A highlighter
- Sticky notes
- A large flat surface (desk or table.)

The pen and paper allowed me to take notes. The pen, highlighter, and sticky notes allowed me to bookmark and annotate source material. The index cards became the repository of themes, ideas, and facts. The flat surface allowed me to view many index cards at the same time and arrange them into a sequence, into groups, into a tree or network, and into combinations of these structures.

4. HUMAN ASSISTANCE

In the pre-Web days I was also more prone to draw on human expertise: teachers, librarians, and subject matter experts. Though I was not hesitant to ask for specific answers or pointers to specific resources, the highest utility of these contacts lay in the questions they would ask, which helped me clarify my quest and raised interesting new directions to explore.

A valuable human resource typically exhibited some combination of the follow characteristics:

- Discernment (“What is he really trying to get at?”)
- Judgment (“Can he do statistical analysis by himself?”)
- Subject matter knowledge
- Research expertise
- Patience.

5. INFORMATION SEEKING SUPPORT SYSTEMS

A useful information seeking support system, then, would require the following minimum functionality:

- 1) It should not interfere with my behavior as listed under Characteristics of Exploration above.
- 2) It should give me capabilities at least as good as those listed under Manual Tools above.
- 3) It should positively assist my explorations by making them easier or faster or more comprehensive or less error-prone or...

In addition, an ISSS might give me capabilities that I never employed before because they were not possible or because I didn't think of them.

But, to be truly a leap forwards, an ISSS would need to exhibit at least elements of discernment, judgment, subject matter expertise, and research savvy.

Honestly, I am skeptical – though not utterly despairing – that such an ISSS is feasible. The readily automatable functionality is unlikely to be much better than the manual tools and processes with which I am intimately familiar. The “human” characteristics that would make a real difference to me are unlikely to be readily automatable.

But I have been pleasantly surprised before. And maybe I will be, again.