On the face of it, the idea of Claude Shannon and L. Ron Hubbard being friends in any sense seems nearly preposterous. One the father of Information Science as a discipline that had deep impact on communication and much of the fundamental infrastructure of modern life; the other a college drop out of some literary talent who founded a religion with a bizarre hagiographic narrative. Yet, there is a letter from Shannon to Warren McCulloch with the phrase “I am writing on behalf of a friend L. Ron Hubbard…” What are we to make of such a line?

We could easily become mired in this question; however, there is another take on this matter. This is just seriously cool! We have spent several years linking information retrieval practices to Shannon’s model and we have developed concepts for maximizing the utility of browsing. How cool is it then to stumble across something neither of us expected about Shannon?
We have taken this as an occasion to revisit the role of browsing in seeking information especially the information of new connections no matter the realm. When one is attempting to go beyond the norm, the recognized public knowledge, one has to search without ordinary finding aids. Finding aids are generally based on what is known. Documents are assigned subject headings from within a list of terms for what is known. Going beyond the norm requires a level of comfort with the unknown and with ambiguity. Setting aside any judgments about Hubbard and Shannon, we see two men capable of striking out into the unknown and making new connections.

The following letter was the catalyst that initiated revisiting conversations in creative browsing and interdisciplinary connectivity.
Salutation: Correspondence, Friendship, and Influence

The letter was written in 1949 by Claude E. Shannon, a mathematician, logician, and father of information theory; written to Warren S. McCulloch, a psychologist whose work was some of the first connecting neurophysiology to cybernetics; and written about his friend L. Ron Hubbard founder of Scientology, science fiction writer, and author of *Dianetics*, which teaches conscious relationships between mind and body. In this letter, Shannon calls Hubbard his
“friend” and asked McCullough to read a copy of Hubbard’s original thesis (not named in the letter) because the topics are out of the scope of his expertise.

At roughly the same time in the 20th Century that Shannon wrote this letter, he published his *A Mathematical Theory of Communication* (Shannon, 1948) and Hubbard wrote his original thesis (Hubbard, 2007). Originally, we had been interested in structural and content analysis of these two documents, Dianetics and *A Mathematical Theory of Communication* in order to speak to degrees of influence of each over the other and to establish just how strong was their friendship.

Dianetics was the new-to-us piece of this puzzle, so we combed through it and ran it through some basic content analyses to compare commonly used phrases and sentence structures. Essentially, we learned nothing and felt like we were making too many assumptions to qualify as contextual influence. For example, Hubbard describes *fields of thought* as a binary relationship between the knowable and unknowable (Hubbard, 2007); while some of Shannon’s work has been abstract to explain the communicated message as a binary relationship between structure and content (Anderson, O’Connor, & Kearns, 2007).

From another perspective, Shannon expresses that every communicated message is made up of *successive symbols* and that the receiver of the message decodes the message one symbol at a time, as long as the receiver understands the code for decoding the message. With the successful decoding of each successive digit, the final message becomes more predictable and eventually successfully decoded. For example, we look to the child’s game of hangman. If we present you the puzzle _ _ | _ _ _ _ , then you would probably assume under widely practiced though not officially sanctioned rules that our message is made up of two words, the first with two digits, the second with four. You might guess some vowels first, so that we can fill in _ _ | U
You examine the puzzle and plug in some possible options for the second space in the second word: is it A? Unlikely. Is it B? Unlikely. And so on until you hit the letter N. Reasonable knowledge of the frequency with which certain letters in the Roman alphabet appear in the English language would make decoding the second word fairly predictable until you have UNI T. If you last a bit longer in the game before your representative stick figure hangs in a noose, you might finally come to realize the puzzle is unsolvable using the Roman alphabet as your selected and presumed code. If you had known you were required to use an alpha-numeric system, however, you would have quite effortlessly solved this puzzle: K 9 | U N I T .

Information theory is about decoding unknown symbols in sequence in a communicated message. If the communicated message is a thought in your brain, Hubbard and Shannon might have agreed that the Knowable and the Unknowable could fit in similarly to the hangman clue. Do you know it? Yes or no. One problem in this analysis, however, is that Hubbard expresses concern only with the Knowable.

We can force some similarities between Hubbard’s expressions of aggregation of colonies and Shannon’s sequencing of symbols and perhaps between Dianetics’ first axiom of survival and the understanding that the mind’s purpose is to survive by solving problems in order to ensure survival, though many obstacles may impede survival and A Mathematical Theory of Communication’s discussion of noise interferences in a clear communication channel. However, we cannot continue these contrived similarities and speculate about scholarly influence between friends. Especially when a letter from Hubbard to Shannon (Hubbard, 1949) so clearly states “Dianetics is not concerned with structure.” We leave it alone.

Resoundingly, So What? was our response to ourselves. Knowing if the two were friends or if we were successful in establishing influence of one over the other does not add, well,
anything significant to the body of literature that addresses topics of information that concern us, such as information retrieval, document representation, or idiosyncratic information seeking strategies.

    Or does it?

**Body: Browsing**

For some time we have been refining a matrix of question. We purposely use the singular “question” to indicate that we are working out a construct that includes the many sorts of requirements and activities that are included in human questioning. In brief, we posit a matrix based on the work of Maron and Levien (1967) when they were working out their model of relational databases, together with a taxonomy of seeking activities suggested by O’Connor, then refined with Copeland and Kearns (2003).

<table>
<thead>
<tr>
<th>Reference Datum</th>
<th>Deductive</th>
<th>Inductive</th>
<th>Conversational</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-articulated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vague Awareness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Browse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stumble</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2: MLO Matrix of question circumstances**

We assert that most library systems, most bibliographic agency systems for information retrieval are founded on the upper left quadrant. That is, they are designed to supply answers or documents likely to contain sanctioned answers, to well articulated questions for which there would be agreement by authorities on the validity of the answer. These hang on the “public
knowledge," those discoveries and assertions made by the best available observations and processes for analysis.

Such systems are not without utility; however, they do not privilege or enhance those searches taking place in the conceptual space lower and more to the right in our matrix. Here there would be less ability of a seeker to specify the subject area in which a solution might be found and there would be less certainty over just what would constitute or could constitute a “correct” response.

One approach to retrieval in this area of the matrix is to shift the locus of representation from the agency or system to the seeker. Here we operate under an assumption that a motivated seeker wants to be seeing the forest AND the trees. The seeker makes the determination of what clusters of things make sense.

In the 1950s and 60s when Morse (1973) was adapting his submarine warfare algorithms to information seeking and when the MIT INTREX project was examining the possible role of computers in scholarly information seeking, it became evident that there were situations in which the seeker was the only one who could make some of the decisions about whether a particular document “shook up the knowledge store” in a meaningful way.

In 1982, O’Connor wrote during his doctoral qualifying exam on the question of whether or not scholars valued browsing and if they did, how could we incorporate computers into the process. At that time, computers were holding only small representations of documents, basically they acted as catalogs that could do things quickly. The primary things that a computer could do to assist in browsing were to help present a random sample and to present data from that random sample very quickly. If one is presented with a large collection of documents and there is little likelihood of having time to go through all of them seeking eureka connections, then one can
simply assume all documents are of equal probability in stimulating connections and use some form of sampling.

Whether one is keeping an eye out for a clever link between a possible sermon and local conditions, attempting to model lensless photography, make a connection between the influence of Christianity in Rome and the shape of letters, redefine verisimilitude, link creation stories and particle physics, or trying to come up with some new connection between as yet unarticulated notions, one is operating within the realm of creativity.

We are coming to understand the neurological bases for creativity and the image is one of integration of several sorts of processes, some of which have bases in the artistic, religious, and inventor practices of the past. The image is in large part an intentional, and conscious attempt to reconcile or reformulate seemingly antithetical elements into a new concept embodying the ‘truth’ of each of the seminal elements. Such reconciliation or reformulation of such seeking requires the ability to ‘know thyself’,” that is, to be in a ‘state of readiness for catching similarities’ and to be ‘gullible’, making free-associations. All manner of connections are made willingly, especially those that seem ‘illogical’ or unsanctioned by current knowledge or models: yet the connections remain subject to critical evaluation. Several authors have provided thoughts on creative activity e.g.

The main obstacle to progress is not ignorance, but the illusion of knowledge (Boorstein).
The trick in coming up with good ideas is to think up a great many ideas and simply get rid of the bad ones (paraphrase of Linus Pauling).

Type T personality: lives on edge of uncertainty, thrives on novelty, and is tolerant of ambiguity (psychologist Frank Farley).

Immersion, incubation, illumination (Helmholtz). (Adapted from Weisburd).
Such descriptions of creative behavior provide a framework for the discussion of bibliographical activity as it relates to creativity. We want to enter a document collection maximizing the probability of a ‘eureka’ situation – the discovery of new knowledge, of new syntheses. Increasing the probability relies on making new connections between attributes of documents, between concepts in the user’s knowledge store (with document attributes as catalyst), and between the user’s concepts and attributes of documents. This is essentially what we have modeled as browsing of the sort intended to ‘shake up the personal knowledge store [by] searching in a literature not obviously relevant’” and to ‘monitor the information environment’.

The various activities generally recognized as browsing attempt to accomplish within real world constraints the possibilities of an ‘ideal’ retrieval system which would screen all user attributes against all document attributes of all works in the collection. It is a matter of sampling texts at various levels of penetration to make one of two judgments: accept/reject document; insufficient data/ sample more. We assert that serendipitous discovery is not a matter of blind luck, rather it is the recognition of a valuable document attribute/connection discovered by means outside normal retrieval rules and relying on the user’s self-knowledge.

Morse (1973) had adapted his wartime algorithms for hunting submarines to scholarly work and had asserted that one aspect of maximizing the utility of browsing was to maximize the number of what he termed glimpses – spending more time inputting useful data and less time finding that data. So, for example, having a computer present a number of citations with just title, subtitle, and author on a screen in a long list saved time over flipping through card catalog cards or even looking at full citations on a screen.
Still, there was an issue. Many browsers used external clues such as title, publisher, book cover color, to find documents that might warrant deeper examination. So, in the early days of computers in browsing there was still a disconnect between the tool and the documents, between the forest and the trees. As computer speeds and storage increase and connectivity increases, more individual trees become available. Some forms of browsing are pushing just a little bit out from the known, while others are forms of exploration with no known target. The latter type often hinges on simply stumbling onto something and having the ability to recognize the value of something upon which one has stumbled—even if that value is not immediately evident or describable.

Accomplishing such stumbling may be said to be aided by presenting more data, by making available leaves and cones and branches and roots of trees within a copse within a woods within a forest. A library or archive is, for all practical purposes, a large document, divided into chunks of a few hundred pages. Many seekers will need many chunks, while others will need but a page. Typically we have not made the single page available; indeed, we have elevated the notion of presenting the document as a whole to the level of theory. Yet, the creative seeker may have little use for this.
Dissecting big works into component parts is becoming easier and less time consuming. Dissecting roots and leaves and branches and scattering them about for stumbling upon may be one of the most meaningful ways to change retrieval systems.

Closing: Fostering Creativity Redux

It may well be true, as George Lois tells us, “creativity can solve every problem” (NPR Staff, 2012), but stumbling upon is also about solving problems we didn’t know we had. Much of the work we have done on browsing has gone hand in hand with notions of fostering creativity. Recent research on creativity (see Lehrer, 2012) gives us a more brain-based and elaborate model of creativity than has been available in the past, yet the fundamentals remain the same. The seeker making connections is among those. The retrieval system may best foster
creativity by standing out of the way; that is, by not imposing representations but by presenting more data, more glimpses, more effectively.

We had a “eureka” moment and something of chuckle. Stumbling upon a letter in which Shannon calls L. Ron Hubbard a friend has caused us to do quite a bit of thinking about the two so different men. None of that thinking has resolved or generated anything to a publishable point, yet. We have not seen a fundamental connection between Information Science and what would become Hubbard’s religion. We have found some utility in simply thinking about what there might be to think about in this instance. The phrase “scholarly discord” arose from the discussions as a cute play on “scholarly discourse.”

Bibliography


Hubbard, L. H. (1949, December 6) [Letter to Claude E. Shannon]. Manuscript Division, Library of Congress (Box 1, Folder 9), Washington, D.C.


