

The Science of Information

Just as I was ready to say in no way are library or information sciences “true” sciences, I decided to check the definitions. My *Webster’s* defines “science” as “knowledge regarding any one department of mind or matter coordinated, arranged, and systematized. . .”¹ “Information” is defined more simply as “news or intelligence communicated.”² That is the *denotation* of “information.” Meadow brings up an important point in stating its *connotation* is “evaluated, validated, or useful data.”³ He explains that “information is what is used to affect a decision.”⁴ This usage matches our 580 class notes (8/28/95) and what Dr. Sinkankas taught in SIS 504. Meadow’s view, especially when considering computer databases, that information is “a basis for subsequent query” is bolstered by the North American Air Defense anecdote regarding the system operators being able to disbelieve the computer reading saying a missile attack against North America was in progress. The operators based their assessment on the absence of other world preconditions the computers could not “know” about.⁵ However, in nuclear work the opposite experience is more common. At both Three Mile Island and Chernobyl, the instruments were reading that meltdown was imminent, but the operators ignored or refused to “believe” the computers, with disastrous results. In both cases, though, information was (or should have been) a basis for subsequent query.

My hasty position that our field is not a science is somewhat weakened by the presumption (from our SIS 580 8/28 notes) that definitions for and in all sciences are hazy. The only real difference between the “hard” sciences and the emerging ones is the

¹ Virginia S. Thatcher and Alexander McQueen, editors, *The New Webster Encyclopedic Dictionary of the English Language*, (Chicago: Consolidated Book Publishers, 1971), 751.

² *Ibid.*, 441.

³ Charles T. Meadow, *Text Information Retrieval Systems*, “Chapter 2: Data, Information, and Knowledge,” (Academic Press, 1992), 21.

⁴ *Ibid.*, 22.

⁵ *Ibid.*, 26, 29.

former are mature enough to have already moved beyond these self-defining issues. I am willing to concede that the Shannon-Weaver Communication Model, Bradford's Law of Scattering, and the formulae of bibliometrics (Lotka's Law, information doubling, discovery/use relationships) are each valid and useful, but I cannot get by the notion that they are all at least pretentious as well, if not contrived. It is almost as if Library Science, as a "new" science, is working extra hard to look like one.

This is not to say Library Science or Information Science are not true sciences. There is a difference between the two, but obviously the "science" part is not the discriminator. Library Science became Information Science when storage and retrieval gained sophistication. An outward sign of this subtle transition (according to the SIS 580 notes) occurred in 1968 when the American Documentation Institute renamed itself the American Society for Information Science. Harold Borko, as quoted by Miranda Pao, advocated use of the term "library information science" in 1984 to "denote the close relationship between these two areas."⁶ So the simple answer regarding the next question—in which of the two fields does information retrieval (IR) belong?—is it depends on if we mean card catalogs and loaded shelving or the OPAC and online databases. This is an oversimplification, of course, but the even simpler answer is IR belongs to both. This begins to become clear when considering modern responses to ever-growing information needs: namely, new computer power and storage technology, and improving electronic communication and information organization/dissemination methods (SIS 580 notes, 8/28/95). Borko defines IR as a system-oriented "interdisciplinary science. . .directly applicable to information processes."⁷ Meadow ties the definition of IR to that of "information" saying IR is concerned with "how to represent information, and how to interpret its structure."⁸ He explains that IR starts when the user recognizes he or she has a knowledge gap, and therefore an information need.⁹

⁶ Miranda Lee Pao, *Concepts of Information Retrieval*, "Chapter 1: Communication and Information," (Englewood: Libraries Unlimited, Inc., 1989), 5.

⁷ *Ibid.*, 5.

⁸ Meadow, *Text Information Retrieval Systems*, "Chapter 1: Introduction," 1.

⁹ *Ibid.*, 7.

But how does information become knowledge? My first look at the knowledge continuum from data to wisdom was last autumn, in SIS 490. In a small group discussion, we decided first that it *is* a continuum and second that the order of acquisition is data, information, knowledge, then wisdom. Data is the content of the information and information is the content of what is communicated. In short, the data becomes information when it is communicated, knowledge when the user processes it, and perhaps wisdom if the user retains or builds on it. Of course, not even among the small group participants did this view gain unanimous support. Even Meadow admits that while “*information, news, meaning, and wisdom. . .*are sometimes use[d] synonymously. . .[.] other times we may recognize differences among them.”¹⁰ He does, however, clarify that “wisdom” is the ability to “provide insight into matters of importance.”¹¹

This view of wisdom ties into my last point of discussion, T.S. Eliot’s quote from “The Rock”—“Where is the wisdom we have lost in knowledge? Where is the knowledge we have lost in information?”—and Gerald Lundeen’s follow-up: “Where is the information we have lost in data? Where do we lose data?” Both information and data, as well as the knowledge and wisdom that could have resulted, are lost at every stage of the communication process despite the previously mentioned improved sophistication and capacities of our communication systems. The basic inhibitor is that data itself is like money: once a user has it, he or she can spend it anyway they want to. So data and information are not only lost through lack of comprehending (or listening in the first place), but also through lack of belief on the receiver’s part. This is true to some degree in the man-machine interface, as my nuclear accident examples illustrate. To be fair though, although machines do not lie, what Vannevar Bush wrote fifty years ago is still true and still a cause of human mistrust: machines breakdown, but their reliability is improving.¹² However, to invert the computerphobic saw: computers make mistakes but to really screw up things takes a human lie. Lies are information (more accurately, misinformation), like fiction. Both are usually based on the truth. However, fiction differs from outright lying in that the receiver knows from the outset the situation is

¹⁰ Meadow, *Text Information Retrieval Systems*, “Chapter 2,” 20.

¹¹ *Ibid.*, 25.

¹² Vannevar Bush, “As We May Think,” *Atlantic Monthly*, 176, no. 1, 102.

untrue, but can still gain wisdom from the work's underlying (excuse the pun) truth. All a recipient may gain from a lie is the knowledge that the teller is a liar. The wisdom of a person once viewed as a liar is always suspect.

This unhappy circumstance calls to mind Weaver's warning regarding information usage in his follow-up to the Shannon-Weaver Communication Model (as explained by Pao). Weaver says lies not only do not count as wisdom, but cannot even be information since the first consideration is the degree of change the information engenders in the user. That is, the message has to be not only received and understood, but validated and acted on.¹³ He continues, explaining that "information is a measure of freedom of choice" not only in delivery and interpretation, but also in use—that is, the action the recipient takes as a result. Weaver's warning works well with Romanczuk's definitions that *knowledge* is information one uses and *wisdom* is information one keeps.

¹³ Pao, *Information Retrieval*, "Chapter 1," 7.