Information Clustering and Problem Solving

ALAN R. SAMUELS

ABSTRACT

People absorb information holistically rather than separately from different levels ranging from the smallest to the highly visible, from the subconscious mind to the written record. A major problem in the transfer of information lies in recognizing that matching information dissemination techniques at all organizational levels to the characteristic ways people absorb information is critical in ensuring the success of providing the right information at the right time.

INTRODUCTION

This article looks at some perceptions of information that are not yet part of the mainstream of information management but may have substantial impact in the future. The purpose of this discussion is to suggest that problem solving in library administration requires alternatives to traditional uses of information not only in solving but also in understanding problems. Such a reorientation involves an interdisciplinary approach to the nature of information. In this process it becomes apparent that various levels or “clusters” of information sources exist. A similar view was suggested by Taylor (1986), who identified information “chunks” as representing “in cognitive psychology...a means for talking about a mental grouping of data in which complexity, recall, understanding, and familiarity are significant factors” (p. 8). Library management involves information transfer using a variety of media. Understanding how different staff both individually and in groups receive and use information and the medium by which it is transferred for problem solving is essential to ensure efficient library service.
Lack of Adequate Definition

Instead of decreasing, the confusion over the nature of “information” has increased. This has caused a major problem in many areas of library management such as job descriptions that not only do not adequately convey exact responsibilities to employees but, more importantly, confuse those writing these descriptions. In normal use the term is viewed as a noun, verb, adjective, or synonym for something else (e.g., data). Conceptually information has been described as a process, a record, an abstract relational term such as meaning or a surrogate for something as yet undefinable. Catalogs are “information banks” while librarians are “information specialists.” We communicate information, store information, access information, transfer information, describe “information rich” environments, prepare “chief information officers,” develop “management information systems,” and discuss “value-added information.” However, as Maricic (1987) and others have pointed out, “no consensus as to the scope of the [information] concept, let alone its definition, has been reached” (p. 34). In some cases we avoid defining the term and emphasize process. For example, Taylor’s (1986) work on “value added information” is process oriented and hence ignores the problem completely by referring to information as a “portmanteau word to cover data, information, and knowledge” (p. 9).

Nor have information retrieval systems been exempt from this confusion. In library management, information retrieval systems have been variously defined as “management information systems,” or “decision support systems” directed by a “chief information officer.” This situation is at least partially caused by the lack of any clear distinction between the commodities with which information retrieval systems deal, resulting in what Teskey (1989) calls “no clear distinction between data, information, and knowledge” (p. 8). In short, we do everything to information but define it. Thus we may be unable to focus on what it is that libraries actually do.

Some propositions can be made. Information is discipline and activity dependent, contingent upon an arbitrary use of the term according to ordinary practice within each field of study or locus of activity (Fox, 1983, pp. 4-5). In an attempt to bring some order into this discordant situation, Debons, Horne, and Cronenworth (1988) have described information as defined by the way it is used (pp. 2-3). They characterize information as a “commodity, energy, communication, fact, data,” and “knowledge.” Although this does not particularly contribute to understanding of the term, Debons, et al. represent the information concept as a type of “knowledge spectrum,” one that encompasses a variety of different levels ranging from the cognitive to the external record. The authors consider that events lead to symbols representative of such events. These symbols are arranged according to a set of rules in order to establish a context. In management terms, a context represents a department or service. When a context is established, awareness follows.
Templates or models are constructed within which "chunks" of information can be combined (Teskey, 1989). With awareness comes a broad clustering of information to which meaning in context is ascribed. Information is formalized, perhaps in policy statements, goals, objectives and the like; is given an intellectual value; and becomes part of the "body of knowledge" or procedure manual that guides the way a library is governed (Debons et al., 1988, fig. 1.1, p. 5).

The importance of Debons's spectrum of knowledge lies in its implicit recognition of information as a continuum, not harmonious with any single definition but rather dependent upon the cognitive level at which it originates. If one recognizes the cognitive nature of information, it becomes necessary to look at information in a cross-disciplinary way. One such attempt was made by Machlup and Mansfield (1983) who brought together specialists in both the hard and soft disciplines and asked them to define "information." The chaotic results forcefully confirmed the contingency character of information. There is no adequate and universal definition of information. What does emerge is a variable dependent upon the context in which it is used. Since information means different things to different people, understanding how people view information within the contexts of their own lives can contribute markedly to the development of information delivery systems by library management.

Others have equated "information" with "knowledge"—an equation considerably distant from Shannon and Weaver's (1949) mathematical formulation of information transfer as the communication of messages regardless of their semantic content or perceptual reception. According to some commentators, information should not be separated from knowledge. For example, Kemp (1976, p. 12) has argued that "information" and "knowledge" are interchangeable terms, metaphorically similar to Taylor's "portmanteau." He suggests that the distinction between "sources of information" and "sources of knowledge" is a tautology that merely confuses rather than clarifies the role libraries play in society and, on a more micro level, the responsibilities of library staff.

However, the expansion of information to include the process of informing or transfer as well as that which is transferred, represents a view that, though problematical, appears to have many adherents resulting in the now respectable phrase "information transfer." Finally, William Paisley (1968) summarizes the current view of process and product by noting that "communication is a process of social exchange...while information is the object of the exchange" (p. 124).

In brief, there is consistent agreement that whatever information is, it is a multifaceted entity that has the dynamics of transfer and the substance of a commodity. The reductionist view of information as solely the document itself (a memorandum, letter, or other written item) has gone out of favor. The current view seems best expressed by Shera
who long ago commented that "in the generic sense, [information] is that which is transmitted by the act or process of communication, it may be a message, a signal, a stimulus. It assumes a response in the receiving organism and, therefore, possesses response-potential" (p. 164). In an organization, directives to employees are communicated in many ways without necessarily receiving a predictable response.

Buckland (1983) gives a stricter definition of information as a process and provides a map of how informing works. This process expands upon Shera's formulation by placing considerable emphasis on such variables as personal knowledge, personal values, cognitive skills, perceptual skills, and "distressing ignorance" (p. 96). Culnan (1985) provides an added dimension to the information concept by noting how perceived accessibility is an important component in the conceptualization process. Citing Mooers, who said that, "an information source or system will tend not to be used whenever it is more painful and troublesome for the customer to have the information than it is not to have it" (p. 302). Culnan stresses that information accessibility, like information itself, is a multidimensional concept consisting of at least three elements:

1. Physical access—the actual mechanism by which access is achieved.
2. Interfacial access—the way in which the access mechanism is made usable by the accessor.
3. Informational access—the "aggravation quotient," the physical effort to obtain information compared to its value in the accessor's context.

Culnan concludes by suggesting that perception of information access mechanisms by the accessor is the major factor determining its use. Thus it is perceptual barriers that must be examined at least as much as the information itself.

Others have related the administration-staff information and acquisition process to a series of psychological variables that, while apparently not of much interest to library managers, are basic to industrial psychology. In a series of articles, Samuels and McClure investigated the role that organizational climate plays in the acquisition and dissemination of information. They discovered that the use of specific sources of information is related to the perceived organizational climate of the library. In libraries where there are mutually supporting relationships among staff and administration, information flow is expedited and used to solve problems in a much more expedient way than in highly bureaucratic and rigidly structured organizations (McClure & Samuels, 1985, pp. 483-98; Samuels, 1979, pp. 237-54; Samuels & McClure, 1983, pp. 1-20).

In organizations not specifically concerned with information transfer, the term information is used in a wider sense. For example, the business sector has been particularly involved in describing the components of information without worrying about its definition. Marchand
and Horton (1986) represent this utilitarian view of information by distinguishing between "information resources" and "information assets." According to them, information resources are:

1. individuals having information-related skills;
2. information technology hardware and software;
3. information facilities such as libraries, computer centers, communication and information centers; and
4. information handling and processing suppliers.

While information assets include:

1. the formal data, document, and literature holdings of the company;
2. the know how it possesses both in the form of intellectual properties like patents and copyrights, and in the form of individual expertise; and
3. the business intelligence and information it possesses about its competitors, its business environment, and its political, economic, and social environment (p. 71).

This categorization of information resources and assets serves to delineate the components of information. Because easily understandable terms are used, the value of information to employees increases substantially and, by extension, to the organization as a whole.

Lacking a utilitarian description of information may invalidate many performance measures developed by managers to measure library effectiveness by not clearly specifying what library service or activity the information is supposed to measure. Smith (1980) has pointed out that measuring information is really measuring only coded signs without regard to their meaning (p. 22). It is only in the daily work life of members of an organization that information assumes meaning and utility. From the systems point of view, members of an organization must compare their perceptions of similar information with each other. In other words, what means something to one recipient may have an entirely different meaning to another.

There are many examples of ambiguity of meaning with which library managers have to deal: what is meaningful to a public service librarian may not be so to one in technical services. For example, although considered a "written record," the catalog card is a part of information transfer as well. The components of a catalog card, ranging from classification number to subject tracings, are the data that are transferred into meaning by the value added processes of cataloging and classification. The end product is then used by others to determine whether or not a particular item in the library is relevant to their specific information needs. The systemic nature of information represented by the written record (the catalog card) depends on whether or not the producer of the card is able to understand the context within which the user will view it, a user whose frame of reference may be entirely different from that of the producer.
Even automation may not alleviate the problem. A MARC record, for example, consists of a series of data elements preceded by identifiers in a sequential order. In other words, there is a clear syntax to the record similar to that of a verbal or written sentence in which different data are linked together by tags, terminators, and other syndetic mechanisms placed at fixed points in the record. For exchange purposes, an unlinked MARC record serves as text. For interpretation, however, the linkages must be reestablished, or placed in context, to be of any use (Gerrie, 1986, p. 5; Crawford, 1986). Eventually the collection of data, links, tags, and text make the bibliographic record.

However, every record must be translated into terms which the recipient can understand. A raw MARC record does little to help a librarian understand how best to serve patrons, nor does a poorly written memorandum aid in the problem solving process. Managers who ignore potential ambiguities in what they convey to their employees risk misunderstanding or, even worse, incomprehension by employees which results in total inactivity—i.e., the job never gets done. Librarians who develop mission statements, goals, objectives, and performance measures, neglecting linkage and communication with those outside of the library, risk disillusionment and funding cuts. In summary, everything that is communicated has an information component and a purpose. That component and purpose must not only be understandable to the sender but also to the receiver, and it should be delivered at a level understandable to the user. Just as advertisers of commercial products carefully choose the media and programs to convey their messages, so also must libraries develop the marketing acumen to match transfer medium, construction of information statements, and psychological life-style of the recipient (Devore-Chew et al., 1988).

**Toward a Taxonomy of Information Clusters**

If it is reasonable to suppose that information is as multidimensional and interdisciplinary as claimed, we are required to seek its components outside traditional library and information science areas. In doing so, categorizing information clusters might be considered as one way of mapping this multidisciplinary approach. The map serves to convey the type of information that needs to be communicated to problem solvers as well as the format in which the information is transmitted.

The taxonomy consists of thirteen groupings of data elements that coalesce to form what might be called an information template. Specifically, these information clusters are: information paradigms, conventional wisdom, written records, semantic intersubjectivity, language and syntax, syndetic structures, context, symbolic representations, linguistic utterances and speech acts, data bonding, data reduction, data assimilation, and template formulation.

Less understood, less measurable, and virtually ignored by infor-
Information studies are those information clusters that are below the level of consciousness. In the list given earlier, these would be data bonding, data reduction, data assimilation, and template formulation. It is these clusters that make conscious activities possible. The subconscious coalescence of information clusters forms a "latent understanding" just below the level of verbal expression which impels the individual to seek satisfaction for that need. It is the manager's responsibility to provide guidelines for the employee's understanding for job clarification and its consequent need gratification. Workers do far better jobs when they know why and how their efforts contribute to the well being of the company.

A good discussion of this level of information clustering is given by Emanuel Peterfreund. In stressing a psychoanalytic and biological approach to information clustering, Peterfreund (1971) notes that, "in general, the biological organism possesses information from (a) phylogenetic sources—transmitted by the genetic code; (b)—ontogenic sources—memory or residues of irreversible experience; (c) current, ever-present stimuli of various kinds, from inner or outer sources; and (d) feedback from the organism's monitoring of its own operation (p. 119).

What seems to separate subconscious and latent information clusters and conscious clusters is language. When internal symbols become visible signs, information becomes transferable. When information becomes transferable, intersubjective understanding can be reached. In other words, organizational coordination of tasks is accomplished only when understanding (though not necessarily agreement) on the meaning of what needs to be done is achieved by organizational members: workers might not agree that a particular service is necessary but at least understand what is being done. Data are also converted into transferable information by the computer programmer. For example, Lachman et al. point out that computer manipulation of data is, in essence, the placement of data in a context that makes it both accessible and understandable to the user. "Computers take symbolic input, recode it, decide about needed input, make new expressions from it, store some or all of the input, and give back symbolic output. By analogy that is most of what cognitive psychology is all about (Lachman et al., 1979, p. 99).

It is through cognitive psychology that the clustering of information can best be understood. The recent emphasis on the study of linguistic behavior by cognitive psychologists (psycholinguistics) has recognized that gradual clustering and codification of information takes place prior to its use in problem solving (Gardner, 1987, pp. 214-15). This clustering rejects the lawlike explanation of information processing in favor of a more environmental set of linkages arrived at through what Lachman et al. call an "information processing paradigm." The paradigm consists of a number of cognitive processes that are performed on data in order to convert it into conscious information.
Among these are symbol manipulation, symbol representation, and analogous thinking (Lachman et al., 1979, p. 90).

Another strategy for clustering information has been proposed by Johnson (1984) who suggests that the isolation of data is reduced by a process of “dialectical synthesis” (p. 205). Johnson proposes that the common underlying structure of information (the context) is established through a reductionist interaction among different data elements. This interaction serves to: (1) reduce redundancy by forcing a selection of admissible versus inadmissible data, (2) combine like data elements to form clusters, and (3) gather each cluster of data elements into a unified system or totality. After this process occurs at the preconscious level, the information thus acquired is subjected to interrogation at the conscious level by the individual. Interrogation allows the individual to compare newly emerged information needs with prior experiences and existing programmed knowledge, separate the heuristic from the already known, and reformulate the residual into a message that is, in itself, subject to translation by others into the language of an information system. Put in library terms, an analogy may be drawn by examination of the managerial decision-making process. An individual “feels” a need for information that can resolve a problem created by either external or internal stimuli. That individual engages in a search process beginning with the most familiar and expanding outward until a possible source for obtaining the information is arrived at. If the needed source happens to be a librarian, the individual contacts a librarian who in turn translates the information need voiced by the individual into the language of existing information sources either inside or outside of the library. An acceptable response will then be given to the decision-maker and incorporated into that seeker’s long or short term memory bank where it forms another part of that decision-maker’s knowledge. Thus the process continues in a constantly changing hermeneutical circle. Decisions are resolved by information which in turn is programmed into policies. Information gathered through performance measures is analyzed and used to develop services.

The transference of latent to manifest occurs through many processes that can be subsumed under the term hermeneutic. There is already a vast literature on hermeneutics (Belicher, 1982; Howard, 1982; Bernstein, 1983; Rosen, 1987). Here a brief review of the concept is given in order to place it in the context of this article.

“Hermeneutics” is basically the study of meaning arrived at through experience. That is, a person understands a phenomenon by participating in it. Thus, in dialogue, people understand the phenomenon of speech by the act of speaking. Derived primarily from literary criticism, hermeneutics has intruded into many different disciplines ranging from sociology to education. It has even touched librarianship through the works of Michael Harris (1986) and H. Curtis Wright (1986). The classic interpretation of hermeneutics describes it as a means
by which we continually refine a phenomenon until it becomes clearer and clearer (Gadamer, 1988). As we repeatedly reflect on a phenomenon, that phenomenon assumes a recognizable shape. In management terms, the more a problem is studied, the more it becomes familiar and the more we begin to understand its particular nuances. Another analogy is online searching. One becomes a good online searcher by constant and ever changing repetition and not by classwork.

Perhaps the best description of "the hermeneutical circle" has been given by Palmer (1969) who writes that:

Understanding is a basically referential operation; we understand something by comparing it to something we already know. What we understand forms itself into systematic unities, or circles made up of parts. The circle as a whole defines the individual part, and the parts together form the circle. A whole sentence, for instance, is a unity. We understand the meaning of an individual word by seeing it in reference to the whole of the sentence; and reciprocally, the sentence's meaning as a whole is dependent on the meaning of individual words. By extension, an individual concept derives its meaning from a context or horizon within which it stands; yet the horizon is made up of the very elements to which it gives meaning. By dialectical interaction between the whole and the part, each gives the other meaning; understanding is circular, then. Because within this "circle" the meaning comes to stand, we call this the "hermeneutical circle." (p. 87)

**Information for Problem Solving**

According to Robertshaw, Mecca, and Rerick (1978), problems "are characterized by three factors: a multiplicity of interactions (the problem is 'complicated'); a necessity to identify what is good and what is bad; and gaps in our knowledge of the situation" (p. 3). The authors continue by describing four processes by which problems are solved: the problem is defined; alternative solutions are generated; the solutions are evaluated; and an iterative procedure is followed. These processes do not exist in isolation but rather form a system in which solutions are arrived at through reconsideration of variables (or "information") that emerge through each phase of the problem solving process.

Problem definition is the most crucial step in the entire process and can only come with recognition that problems, like information, exist at many levels. Although the problem solving process can begin with recognition of immediate decisions to be made, as each decision is made and programmed it is likely that other problems will emerge, many of which have been unrecognized or are new. As dialogue between supervisor and supervised occurs, new understandings are reached about what is required to get the job done.

According to current linguistic thinking, "understanding" is a matter of translation. In Debons's scheme cited earlier, the translation process consists of formulating rules which can organize the symbolic representation of events in the information seeker's mind. These rules are not necessarily the property of any one participant in the problem solving process but rather are arrived at by mutual agreement as to the contextual meaning of the visible symbol. For example, if a supervisor
writes down his/her understanding of the objective of an employee in the management by objectives (MBO) process and then shows it to that employee, that supervisor has no way of knowing beforehand whether or not the employee will understand what is written. However, if the supervisor has taken the time to evaluate the environment within which the employee normally works, the supervisor's ability to translate accurately his/her understanding of the employee's objective is strengthened.

Problem solving is basically a cognitive activity. Data gathered through community surveys, needs assessments, information audits in organizations, and similar techniques may be objectively summarized but must be subjectively interpreted. The Public Library Association's role setting manual clearly recognizes the importance of cognitive interpretation through its urging that those who would plan services that are of use to their communities take whatever time necessary to "look around," to gain a subjective impression of the context in which services must be provided. Impressions are obtained through a combination of cognitive processes, among which are archetypical representation, data assimilation, data reduction, and data bonding.

Archetypical Representation

At the most basic level of the human mind lies the archetype. In psychoanalytic terms, the archetype cannot be exactly defined. It is a pattern that can only be recognized by the effect it produces (Jacobi, 1959, p. 31). An archetype is a mode of apprehension, a prototype of perception, that establishes a psychic context for data assimilation and template formulation. Although the existence of the archetype can never fully be verified, adequate empirical support for its existence has been suggested (Mattoon, 1981).

The function of the archetype is to utilize whatever innate characteristics the individual has to begin the process of ordering the thought process through the creation of templates within which data can be placed. Precisely how this mechanism works is unknown, but we can gain a dim glimpse of the process whenever an idea arises in our minds from no discernible source. When the template is formed, data assimilation begins.

Data Assimilation

The senses continually absorb stimuli to an overwhelming degree. Most data are not recognized as such since they are assimilated at the subconscious level. In many ways this is similar to Abraham Maslow's (1970) basic "instinctoid" needs—needs which are "vague, unquantifiable, and hardly scientific" (p. 96). It is also what distinguishes the elusive "managerial style" from the more concrete "management science." As more and more data are assimilated, it becomes necessary to reject as well as to gather data to avoid an overloaded state produced by
excessive redundancy. When a balance is reached between a person's ability to combine data into recognizable information, data assimilation is stopped. When achieving such a balance is ignored, the result is information overload and indecision.

**Data Reduction**

Data reduction is the process of eliminating redundancy. Like data elements are combined or clustered. These clusters are the origins of preconscious needs whenever there is a part of the cluster that is missing. An analogy is a puzzle in which all but one piece is present. It is the search for that piece that bonds different data clusters together.

**Data Bonding**

As a decision-maker approaches a problem, he/she may have some idea of what information is needed to solve it. This may be quite conscious or still little understood. This phenomenon is often seen occurring in libraries. “Browsing” is one example of information-seeking without necessarily having a conscious purpose, and “searching the card catalog” is another. Most cataloging systems are predicated on the presumption that people are familiar with the area of information they are seeking. However, it is common enough to observe people exercising a random search pattern for “something interesting”—what might be called the serendipity factor.

Data bonding creates information clusters that act as stimuli to conscious, rather than latent, acts. The speech act itself is a verbal representation of internal need even though it may not be understood by others. However, it is through the process of socialization and interaction that mutual understanding is achieved by the establishment of contexts in which speech becomes meaningful—i.e., for informative.

**The Passage from Information to the Written Record**

At the conscious level, environmental and contextual factors take hold of those formerly internalized information clusters and continue to transform them. Speech acts are combined in a “semantics of interaction” and become symbolic representations of visual phenomena (Rasmussen, 1985, p. 57). Syndetic structures are identified and created in order to facilitate the retrieval of these symbolic representations. “Languages” of varying kinds are produced through the combination of symbols and syndetic structures. These languages have both semantic and semiotic aspects. A programming language, for example, may be entirely incomprehensible to the nonprogrammer but still has the same structure as normal discourse. All written records have languages of their own. Without observing a strict sequence of placement of data elements, neither a MARC record nor a printed catalog card would have meaning. Without understanding the overall mission of an organization, a manager would find it very difficult to develop goals and objectives.
Information clusters are continually refined through the addition of values, environments, connectives, and mutual agreement about interpretation. Eventually they become "programmed"—i.e., become part of that conventional wisdom or body of knowledge that managers use to guide their selection of techniques and styles. Selection of the management mix of techniques and styles is the major problem with which managers must contend.

Problem solving requires presentation of information that is comprehensible and transferable. It does little good to present a written document to others as "information for problem solving" unless the recipients are aware of the context within which to view that information. Each individual or group has preconceived perspectives about a problem developed through a combination of preconscious information clustering and external experiences. It becomes essential for the information communicator to understand and not necessarily agree with these preconceived perspectives. Understanding what these perspectives are can lead to training programs which enable individuals to translate their own perspectives into the language of the environment in which the problem must be solved—a method well known to educators who, in dealing with multicultural environments, describe the process as linking understanding of the world with the word. "From the beginning, in critical and democratic practice, the reading of the world and the word are dynamically linked. The command of reading and writing is achieved beginning with words and themes meaningful to the common experience of those becoming literate, and not with words and themes linked only to the experience of the educator" (Freire & Macedo, 1987, p. 42). When the written record becomes recognizable to the user in his/her own terms, it is suitable for problem solving.

**Utility**

The value of any theory lies in its utility. Utility itself is a charged word and requires the addition of purpose: "utility for what?" There is considerable reason to believe that an understanding of the psychology of information clustering has significant utility in promoting library service. Herbert Goldhor's own extensive work in studying the best way to display library materials (prime display areas) is a very practical example of translating the theoretical to everyday use. It also recognizes the tendency of people to seek that which appears to be intersubjectively recognized as "good." Library users may not like a best-selling book but sometimes read it because everyone else supposedly does. Labeling is yet another mechanism that facilitates the search process by describing the content of many different items through one or two symbols.

From the overall perspective of what a library is supposed to accomplish, perhaps the most important use of the process of information clustering is in directing the library patron. It is now part of the "conventional wisdom" of library practice that patrons use libraries for
many reasons other than reading. Libraries accommodate both readers and nonreaders. With this recognition comes the need to communicate in ways that deal with all classes of library users ranging from the illiterate to the scholar. A good example of recognizing the visible communication of symbols occurs every time someone takes a vision test for a driver’s license: he/she is asked to interpret wordless symbols rather than labels, interpretations that come about through experience and common agreement about their meaning. The implications of this for developing unified information transfer systems such as signs or online public access catalogs in libraries and other information agencies are substantial, especially in the multicultural environments within which libraries exist. Can we communicate the possible uses of the library without words or other limited mechanisms? It is the library manager’s responsibility to use every means possible to match the needs of employees with those of the organization. Promoting unity of action and mutual understanding by library employees at all levels through careful use of information transfer devices, recognizing which such devices match the capacities of the receiver, and sensitivity to the many different points of view held by organizational members will contribute to providing good library service.

REFERENCES


