Chapter 7
Information Redux

Information exists at the core of human activities such as observing, reasoning, and communicating. Information serves a foundational role in these areas, similar to the role provided by mass or velocity in the field of physics, or emotions or ideas in psychology. Information Science is the discipline responsible for discovering and studying the informational laws, methods, and applications that range across a universe of potential problem domains.

Processes produce information, just as energy produces the movement of objects studied in physics. Thus, understanding the nature of processes is necessary for a full understanding of information. Processes and their outputs may be understood on two levels: (1) the basic nature of any process and its output, consistent with certain assumptions, which is the focus of *theoretical information science*, and (2) the nature of individual processes and their outputs with specific capabilities in specific domains. These processes include processes that assign economic value to information, or processes that encrypt information, or processes that make information into knowledge in the human brain. Studying types of processes and their outputs, or individual processes and their outputs, is the focus of *applied information science*.

Much can be said about processes. For example, computational complexity has suggested various models of the complexity of processes, including the amount of time needed for processes to operate, often based on the nature of the details of the processes or the amount of storage needed. The different types of operations may be studied (e.g., digital devices or quantum devices), as well as the communication and time considerations that occur with operations. Processes may also be decomposed into subprocesses, or smaller processes may be brought together into larger processes. The nature of the equivalence of processes, and other relationships between processes, may be formally studied. Similarly, the inputs and outputs to a process may be described and modeled in different ways (162).

A number of models have been developed about processes, largely produced by computer scientists and mathematicians (162; 94; 90). One prominent idea that was discussed in Chapter 2 was based on the Turing and Church work that lead to the common treatment of processes as computer algorithms. Given the Church-Turing hypothesis, the Church-Turing information principle suggests that the information produced by any process consistent with Turing's constraints can be produced by a universal Turing machine. Related to this is Turing's halting hypothesis, which suggests the Church-Turing halting principle, that not all information values can be produced by a process consistent with the Turing constraints.

Information has been defined in this work as the characteristics or the state of a process's output. The characteristics of the output provide information *about* the input to the process, as well as *about* the process itself. Because these definitions may be applied to all types of processes and applied to all types of information, such definitions are highly useful for those focusing on information, how it comes to exist, and what information and which forms of information can exist. Understandably, someone who focuses on human behavior might wish to limit the information they study to human communication or the output of cognitive processes. An economist might wish to limit their focus to information that has value to those that use the information. Mathematicians may choose to focus on information defined by the assumptions and rules of a particular area of mathematics, such as probability theory or topology. The Church-Turing principles apply in all of these domains.

Information professionals have two unique strengths: (1) their ability to understand what and how information is produced in general, and (2) the ability to apply this knowledge to understanding the information in existing systems. To fully understand the nature of information and its production requires some broad knowledge of processes and their outputs, a knowledge that has some depth, along with a detailed and rigorous knowledge about specific types of processes in their range of interests, whether it is the study of behavior or of gravity or of accounting or of assigning an economic value to information. Some individuals may take a more theoretical route, studying formal methods for describing information processes and their outputs. Computer science and mathematics provide many of the basics for the theoretical study of information.

The distinction between information science and other disciplines may be understood through an examination of their different goals. When the focus of study or practice is on the emphases of the disciplines, then the study or practice is part of that field. When the focus of the study is on information and its production, then the work is part of information science. A degree of respect for the interests of others is necessary when observing and describing information, given that it is one of the most widely discussed phenomena in many disciplines. At the same time, there is a need to accept that there is a science of information that rigorously aims to describe, predict, and explain all the characteristics of information, broadly construed. Theoretical information science clearly focuses on information, while applied information science is a much more difficult discipline to describe. When someone focuses on applying what one has learned from theoretical information science to practical situations and existing processes, one is practicing applied information science. Clearly, information is being produced in these specific processes, and information scientists should study the processes in order to understand the information that is being produced. This may also help one to understand better the nature of information in this problem area.

While field-specific studies of the application of information principles may serve as a part of applied information science and are clearly useful to the members of the field being studied, the interests of that specific discipline should not be seen as the primary determinant of what is useful to those studying the intersection of information science and another discipline, where one needs knowledge and skills in both the application field and in information science. When one studies a problem domain and almost exclusively examines phenomena using the models and definitions of that discipline, with an occasional mention of words like *information* or *informatics* or *information technology*, one really is not conducting applied information science. Specialists in other domains best understand the processes that occur in that domain, and most studies of the processes that occur in fields like psychology, physics, business, or computer science are often best conducted by specialists with degrees in those fields. Discipline specialists have a strength in the understanding of how processes in their domain use information. Trained in the various modeling and evaluative techniques of a discipline, these specialists have an expertise in the application of these models and measures and can estimate accurately what will be useful.

Differentiating between information science and another discipline becomes more problematic when dealing with disciplines such as journalism or communication that seem on one level themselves to focus on information. However, in many studies in these disciplines, emphases are placed on how to produce an article or a web page or production techniques or the politics of persuasive communication. Clearly there are informational aspects to all of these, but the disciplines in which a study might best take place depends on whether the focus is on information-as-information and the producing processes, or whether the focus is so narrow and so deep as to look almost exclusively at the problems of that specific field with relatively little interest in the general problems associated with information.

The analysis of existing informative processes often takes place among researchers, systems analysts, and other more specialized disciplines. Analysts often study ways that systems can be improved, with or without the use of computers, often maximizing various quantitative factors. Others have taken more qualitative approaches to the study of informative processes. By using standard modeling techniques, languages, and graphical methods for describing systems, analysts have developed as a profession that often overlaps with the techniques and knowledge of information professionals. The expertise of information specialists is usually different from that of systems analysts who focus primarily on automating or modifying systems, often in areas in which they have significant professional expertise. Information scientists have an appreciation for how information is produced or can be produced across a range of application domains.

There are a range of basic types of processes that are studied theoretically and that can be applied across a range of disciplines. One of the most fundamental aspects of human descriptions of informative outputs from processes is as a form of informational representation. A representation at the output of a process is a modified form of the input to the process, as well as being a partial representation of the process itself. Cameras capture representations of what is in front of the camera, as do the retinas in many species of animals. Computers may store representations of data entered or obtained from a network. All these informational representations take a form that supports a further use of the information. The eye produces an image on the retina so that the image may be sensed and processed. Other representations may exist but we often do not think of them as interesting. The light hitting a fingernail may contain the same information that arrives on the human retina, but one representation is clearly more useful for many human endeavors than the other.

Inputs to processes may be represented many ways, referred to as encodings, and an appreciation for the various types of encodings is one of the major skills of an information scientist. Besides the general notion of encodings being useful, one finds encodings designed to support the detection and correction of errors. Representations that support error detection have capabilities built into the representation that allow for the detection of some errors that may occur. Error correcting representations not only provide error detecting characteristics but also allow for the correction of problems. The problems that may be detected or corrected may range from hardware errors to data entry errors to higher level conceptual problems.

Informational representations may also take on more specialized aspects. Input information may be compressed by a process to save space or to facilitate processing. A baby's cry can be digitized into hundreds of thousands of bits, but its mother can often quickly guess what the cry means, digesting many bits down to a state of the baby's feelings that may be represented by a few bits. Similarly, encryption may occur to keep information secret, to respect privacy or to control information. Encryption usually requires an encrypting process with a matching, inverse process that decrypts.

Metainformation is an often used aspect of representation, where both a representation is provided, along with information about the production of the information. For example, the reader of an enjoyable work of fiction might want to find other works by the same author, or other works about the same culture, or books written in the same genre. Metainformation about each work might allow the individual to locate desired works. As an increasing amount of work appears in digital form on networks, the presence of useful metainformation often provides an access point that allows the work to be located and used, when it otherwise might not be located with a reasonable amount of time or effort.

Organizing information is, like metainformation, a higher form of information representation. While traditional metainformation may carry information about the topicality of an informative entity, organizing a group of entities may add value to a collection of entities, allowing for a searcher to find items more rapidly. Browsing, seeking items for use, depends on an arrangement that supports the browsing needs of either the browser or the collection manager (126; 123). Libraries often arrange material by topic for ease of subject browsing (121), whereas retail stores often arrange materials to encourage shoppers to navigate through as much of the store as possible, often attempting to maximize the store's sales.

Finding informative entities is often accomplished through searching, in which a single item is retrieved, or possibly members in a list are ranked by the probability that the informational entities are relevant to the user. For example, using a search engine often results in the system computing the similarity between the representation for the query and the document or web page's representation. The same techniques that support retrieving information may be used to filter email, video material, or other information. This has the effect of allowing the user to spend time accepting the information that is most interesting or most relevant and ignoring the information that is least likely to be interesting, minimizing information overload.

Information may often be improved upon by the execution of processes to produce new information. Processes themselves may change over time through direct efforts, such as through logical or statistical operations, or as companies produce more sophisticated computers. Improvement may occur through the execution of evolutionary processes, where the environmental rewards associated with some processes surviving result in those processes reproducing and thus generating more of their type of information, while other processes cease to exist and thus produce less of their type of information. New information is produced in many cases when information producing processes change.

Information producing processes must exist in order for individuals to communicate with one another and generally, for sentient beings to gain information about the world outside them or, conversely, for the beings to influence the world. Perceptual and observational processes produce information about the world outside the perceiver or observer. Languages are used to provide information from one entity to another entity. Terms in natural languages often represent aspects of the real world, serving as informational representations.

Information is used by decision makers, and understanding the nature of the information is essential to the understanding of decision making. For example, the value of information may be determined by comparing the value of the decision that would be made, given certain information, and the value of the decision that would be made without the information. Similarly, using a representation that may be chosen to communicate information, an information structure, may be compared to using a different representation, allowing one to determine the economic value of a representation. An informative process may similarly be viewed by studying how it improves cooperation. Processes that hide past information may increase one's economic state, and processes that determine past information about someone with whom one is considering making a contract can help avoid economic loss for those on the other side of such a contract. Failing to disclose information or actions may be to one's advantage. Signaling processes communicate information about an individual or organization, and knowing the benefit associated with sending the signal or receiving a signal may prove beneficial to decision makers.

The study of information is a discipline with both breadth and significant depth. While many people nibble at the edges of information science, with a few taking big bites out of the discipline, it will be necessary, in the future, for many to have both a full appreciation for the nature of processes, the precise nature of the information produced at the output of these processes, and the types of processes that occur in various disciplines, as well as the methods to study information in those disciplines.