Library and information science and biomedical informatics: converging disciplines

Prudence W. Dalrymple and Nancy K. Roderer
Division of Health Sciences Informatics, School of Medicine, Johns Hopkins University, Baltimore, MD

Abstract
The disciplinary boundaries of library and information science and of biomedical informatics are remarkably similar. Both disciplines deal with data, information, and knowledge and with their storage, retrieval, and use in the service of society, yet each is rooted in its own unique sociocultural and historical context. While it is undeniable that computer technology has substantially influenced both fields, it is the fundamental principles of information and knowledge organization, storage, retrieval, and use that provide a common foundation for research and practice in the two fields. In this comparative survey, representative models and principles are provided to illustrate each field, and the methods used in the two fields are compared. Training, professional organizations, and accreditation processes in each field are described. In all of the areas reviewed here, there seems to be a convergence of the two disciplines and the likelihood of more to come.
Definitions and scope

The disciplines of information science and biomedical informatics have followed somewhat different developmental paths over the years but have a great deal in common. Both deal with information, technology, and people: information science across all domains, and biomedical informatics within the narrower domain of biomedical activity. Both make heavy use of information technology, whose ongoing advances have made it easier to retrieve and combine disparate types of information in disparate subject areas, thus blurring the lines that have distinguished the two fields.

In order to appreciate the interrelatedness of the two disciplines as well as the unique contexts from which they have emerged, it is important to understand how the terminology used to define these fields has reflected the underlying process of evolution in the respective disciplines.

Library science, Library and information science, and Information science

It would be difficult to define the scope and nature of library information science without also considering the related terms library science (or librarianship) and information science. Considerable overlap exists among these terms, and the distinctions between them have not been universally agreed upon.

According to the American Library Association's Standards for Accreditation of Master's Programs in Library and Information Studies (1992),

The phrase "library and information studies" is understood to be concerned with recordable information and knowledge and the services and technologies to facilitate their management and use. Library and information studies encompasses information and knowledge creation, communication, identification, selection, acquisition, organization and description, storage and retrieval, preservation, analysis, interpretation, evaluation, synthesis, dissemination, and management.

In his cogent discussion of librarianship and information science, Buckland (1983) observed more than 20 years ago that our rhetoric concerning these terms is dominated by references to a well-understood institution, the library. However, he also pointed out that librarianship is in fact a subset of information science, sharing many of the same concerns and approaches to information. These two terms came into use together during the era in which computerized information systems began to appear and many of the concepts of library science came to be seen as more broadly applicable to the organization, storage, and retrieval of information in general. From this perspective, information science can therefore be seen as a broader superset because it calls attention to all phases of the information life cycle.

Although the term library and information science is widely used (as for example in the title of this conference), the authors of this paper share Buckland's view (1983) that library science is a subset of information science and will therefore consider the more general field of information science in the present paper.

An early definition of information science was provided by Borko in 1968:

Information science is that discipline that investigates the properties and behaviour of information, the forces governing the flow of information, and the means of processing information for optimum accessibility and usability. It is concerned with that body of
knowledge relating to the origination, collection, organization, storage, retrieval, interpretation, transmission, transformation, and utilization of information.

Bates (1999) has further explicated the field of information science by pointing to some of its interesting "below the water line" features, those that are less consciously and explicitly addressed. She has noted that information science, like education and journalism, is a meta-field; that is, it cuts across the conventional academic and content disciplines, offering a particular overarching perspective. Thus, it is interdisciplinary, with ties to a number of other disciplines, including communication and computer science. Key concerns of information science, according to Bates, are the organization of information as well as information seeking and searching processes.

In a recent series of articles, Zins (2007) has reported on a Delphi study entitled "Knowledge Map of Information Science." This study, conducted in 2003-5, sought input from 57 internationally recognized experts in the field as a means of exploring the foundations of information science. Zins began with an identification of data, information, knowledge, and message as the explored phenomena in information science, then identified six possible models or perspectives (hi-tech, technology, culture/society, human world, living world, and physical world). He reported that most of the panelists associated information science with the culture/society model, suggesting that this field is focused primarily on the mediating aspects of data, information, knowledge, and messages as they are implemented in human societies.

From his research, Zins proposes a ten-faceted hierarchical model, where the facets are foundations, resources, knowledge workers, contents, applications, operations and processes, technologies, environments, organizations and users. The foundations category is described as the metaknowledge of the field and is further described at the theoretical and conceptual levels. The other categories are considered the essential body of knowledge in the explored phenomena, with associated issues and types. Overall, Zins' model suggests the structure of knowledge in information science and the conceptual relations among the parts of the field.

Characterizations of information science by the American Society for Information Science and Technology (ASIS&T) and the recently established United States I-Schools Project underscore the importance of this mediating aspect of information science: ASIS&T defines its membership as "information professionals leading the search for new and better theories, techniques, and technologies to improve access to information." (http://www.asis.org/) This focus on information and technology is expanded in the characterization of the I-Schools Project as being made up of schools "interested in the relationship between information, technology, and people." (http://www.ischools.org/oc/web site)

Biomedical informatics

The beginning of biomedical informatics as a discipline has been associated with the 1959 publication in Science of an article by Ledley and Lusted (1959) concerning the reasoning foundations of medical diagnosis. This focus on medical decision making has continued and evolved in bioinformatics, in tandem with computer science and information science.

The discipline of bioinformatics was initially concerned with computer and information technology that focused on medical problems. By the 1980s, however, a more integrated view of medical computing had emerged. Scott Blois (1984), a pioneer in the field, described computing as a "novel research tool" that would do more than merely help us manage information: It would also help us understand the nature of that information and provide support for decision-making in the clinic and laboratory. Originally, the field of biomedical informatics was most often referred to as
medical informatics. More recently, subfields of this discipline have emerged in the clinical arena, including subfields of medicine, such as primary care and oncology, together with nursing, dental, and veterinary informatics. In addition, informatics concepts have also come to be applied to other types of health-related data, particularly basic science- and public health-related information. While unanimous agreement has not been achieved with regard to an appropriate name for the overarching discipline, current practice seems to suggest a convergence on the term biomedical informatics, which is the term we will use in this paper.

As has been noted by Perry et al. (2005), definitions of health-related informatics over time have reflected the evolving nature of the field:

- 1977: Medical informatics is the application of computer technology to all fields of medicine – medical care, medical teaching, and medical research. (Collen, 1977)
- 1984: Medical informatics comprises the theoretical and practical aspects of information processing and communication, based on knowledge and experience derived from process in medical and health care. (van Bemmel, 1984)
- 1990: We define medical informatics as the rapidly developing scientific field that deals with the storage, retrieval and use of biomedical information, data, and knowledge for problem solving and decision making. (Blois and Shortliffe, 1990)
- 2006: Biomedical informatics is the scientific field that deals with biomedical information, data and knowledge – their storage, retrieval and optimal use for problem solving and decision making. (Shortliffe and Cimino, 2006)

Another definition of biomedical informatics put forth by Stead in 1998 as a simplification, but still capturing the essence of the field, is "the science that deals with health information, its structure, acquisition, and use." This definition clearly underscores the fundamental principles that bioinformatics shares with information science.

Shortliffe and Cimino's Biomedical Informatics, the core textbook in this field and the source of the 2006 definition above, goes on to note that the field is "probably best viewed as a basic biomedical science, with a wide variety of potential areas of application. The analogy with other basic sciences is that biomedical informatics uses the results of past experience to understand, structure and encode objective and subjective biomedical findings and thus to make them suitable for processing. This approach supports the integration of the findings and their analysis. In turn, the selective distribution of newly created knowledge can aid patient care, health planning and basic biomedical research.

**Similarities and differences between information science and biomedical informatics**

Although the definitions of both fields have evolved over the years, information science and biomedical informatics share a focus on information, technology, and people and on the functions associated with information. Both define themselves as a science, and both consider themselves to be multidisciplinary fields. One clear difference between the two disciplines is that biomedical informatics focuses on the particular domain of biomedicine, while information science is broader. Other types of informatics are beginning to emerge, and we expect to see the emergence of informatics in other domain areas; nevertheless, in this paper we will restrict our discussion to biomedical informatics.

**Models and methods**

In the view of the authors, the field of information science has been more explicit than biomedical informatics with regard to identifying models
and theories within the domain. Given the extensive literature in information science that addresses discipline-specific theory, we will focus first on the major areas of study included in biomedical informatics, and then comment on the parallels to information science.

In their textbook, Shortliffe and Cimino (2006) provide a comprehensive description of biomedical informatics and then use a structure based on recurrent themes to further explicate this field. Their themes are:

- the nature and organization of biomedical data
- acquisition, storage and use of biomedical data
- biomedical decision making
- use of biomedical data from a cognitive perspective
- computers in biomedical informatics
- standards in biomedical informatics
- system design and engineering in biomedical informatics
- issues of images in biomedical informatics
- natural language and text processing of biomedical information for information retrieval

These themes recur as additional chapters in the book, which describe various types of biomedical information systems and their applications. For example, in the chapter on public health informatics, Yasnoff et al. (2006) suggest that the themes remain similar in this area, but they are set in the specific context of public health; that is, the themes relate to the health of populations rather than that of individuals and to the support of the core public health functions of assessment, policy development, and assurance.

The nature and types of information involved

There is an extensive literature in information science that considers that most basic of concepts, the nature of information. Much of the early work in this field is nicely summarized by Blois (1984), who describes the issues of both defining information as a concept and identifying what information does. He notes that concepts of information are employed in many disciplines and in each case relate to their field of origin.

He suggests that the concept of information viewed from a communication perspective must include elements of a sender, an information channel, a receiver and a decision maker. He notes Shannon and Weaver's (1949) early identification of information as a message, or measurable sequence of signals, an approach that excludes meaning, and the more commonly considered view of information as transmitting meaning. In discussing what information does, Blois distinguishes between information and knowledge and the process of human perception that translates information received into one's personal perceptions. Blois goes on to discuss Whittemore and Yovits' (1974) view of information as data of value in decision making, a concept that finds a parallel in the fundamental interest of biomedical informatics in medical decision making.

The concept of differentiating between data, information, and knowledge is prevalent in both fields. Shortliffe and Cimino (2006), for example, identify a datum as any single observation of fact, and they define knowledge as relationships derived through the analysis of facts. Information is seen as a more generic term that also reflects the organization or analysis of data. Both information science and biomedical informatics concern themselves with a variety of forms of information, including data files, text, and images. Both use the concept of recorded knowledge to define their main area of concern, with biomedical informatics concentrating on patient records, genetic descriptions, public health data sets, and other
entities specific to biomedicine. Images are an area of particular interest in biomedical informatics, but they are also increasingly of interest to those in information science.

**Organization of information**

From library cataloging and classification to indexing and thesauri and on to metadata, the organization of information has been a persistent theme throughout the history of information science. Similarly, a large number of vocabularies have been developed within biomedicine that are directly relevant to biomedical informatics, including vocabularies for the description of medical research (such as MeSH) and vocabularies describing medical conditions and treatments (such as SNOMED). The National Library of Medicine (NLM) has been a leader in the development of medical vocabularies, beginning with Medical Subject Headings (MeSH), first published in 1954 and then called the Subject Heading Authority List. More recently, the NLM has created the innovative [Unified Medical Language System (UMLS)](http://www.nlm.nih.gov/research/umls/). Within the UMLS, a metathesaurus serves as a large, multi-purpose, and multi-lingual vocabulary database that incorporates a very large number of the medical vocabularies in use around the world today. Other components of the UMLS are the [Semantic Network](http://www.nlm.nih.gov/research/umls/) and the [SPECIALIST Lexicon](http://www.nlm.nih.gov/research/umls/specialist/).

**The information life cycle**

Both fields make considerable use of the information life cycle as a model for their area of concern. In Shortliffe and Cimino (2006), for example, the life cycle of patient records is portrayed with an eye to the complexity that points to the limitations of paper-based record systems. An example of a life cycle in information science is that described in 1981 by King et al., who identified the system participants as authors, publishers, libraries and information center, abstracting and indexing services, and users and listed the system functions as research and information generation, composition, recording, reproduction, distribution, acquisition and storage, organization and control, identification and location, physical access, and assimilation by the user. While these functions were articulated more than 25 years ago, it is noteworthy that they continue to represent today's information transfer model.

**Technology**

An interest in computer and networking technology, computer software, and the standards needed for computing and communication is a predominant characteristic of both information science and biomedical informatics. Indeed, both fields have grown to a considerable extent out of the development of computer technology, and both have benefited from increases in the types of computer technology available and the ever-growing use of technology for communication.

**Information retrieval**

A considerable amount of work has been done by information science researchers on information retrieval: that is, on the methods for obtaining the text or documents most relevant to a particular query. This work has involved the exploration of a variety of structures for databases, approaches to describing documents, and mathematical approaches to selection. In biomedical informatics, similar approaches have been applied to biomedicine, for example, as part of the effort to find ways to retrieve information related to disease. The Text REtreival Conference (TREC),
which supports information retrieval research by providing the necessary infrastructure for large-scale evaluation, includes a genomic track that looks at retrieval within the broadly construed domain of genomics.

**Information use, including decision making**

Within information science, there has been a considerable emphasis on models and theories of information behaviour. The recent ASIST monograph, *Theories of Information Behavior*, provides an overview of the many the models and theories in information science (Fisher et al., 2005).

Models such as Taylor's information use environment (Palmquist, 2005) and Dervin's sense-making (Tidline, 2005) provide different perspectives on how individuals identify and seek to meet their information needs. Taylor, whose image of the information use environment recognized that the user's environment or situation has a critical effect on the nature of information needed, has developed some of the earliest models of information use and was a precursor of a number of models focusing on the user's cognitive state. In the case of sense-making, a paradigm that focuses on the user of information and the relationship of communication, information, and meaning for that user has an associated suite of data collection and analysis techniques that can be employed to yield a better understanding of user activities. The latter is an example of how work from the field of communications has also contributed to information science.

While there are relatively few principles, or verified theories, in either field, the area of information use has spawned a few such principles. The principle of least effort, articulated by Zipf (Case, 2005) suggests that "in performing tasks, individuals adopt a course of action that will expend the probable least average of their work," or that "people invest little in seeking information, preferring easy-to-use, accessible sources to sources of known high quality that are less easy to use and/or less accessible." (Bates, 2005) This principle has been found to hold true for many people in many different environments.

There are long traditions of user studies in both biomedical informatics and information sciences. Many of these studies have been based on, or have formulated, the kinds of information use models described above. According to Herbert Poole (1985), the first study of information use was Charles Eliot's article about the used and unused portions of a library's collection. Early studies in general concentrated on channels of information use, especially libraries.

With the growth of science in the 1940s, attention began to turn toward improving dissemination of information from research. Many studies of information use in science were conducted in the 1950s and 1960s, and described both in original articles and in reviews such as those of Menzel (1960) and Paisley (1965). In the latter parts of this period, researchers brought in communication-based models and methods, and more attention came to be paid to the information needs of different demographic groups, such as urban residents (Dervin, 1973). Studies from each of these perspectives - library use, occupational groupings, and demographic groupings – continue to be conducted, with the American Society for Information Science and Technology's Annual Review of Information Science and Technology providing useful summaries, most recently by Case (2006). Case's *Information Use* (2002) provides a detailed review of many of the user studies carried out in the information science field and the methods they have employed.

Among the earliest of the user studies in biomedical informatics is that of Covell et al. (1985) who used surveys and observation to look at the kinds of information resources employed by physicians in their offices. A striking finding from this study was that only 30% of physicians'
information needs were met during the patient's visit. A group of later studies of physicians by Forsythe et al. (1992) used an ethnographic approach to directly observe communication about information needs in a university teaching hospital. Gorman (2001) studied the information needs of primary care physicians. Recently studies have focused on the cognitive aspects of user behaviour, as in Kushniruk and Patel's (2004) work on the impact of a computer-based patient record on physician's information gathering and reasoning strategies.

**Methods**

Bates (2005) has described the information science approach to methods: "The fundamental methodological stance of information science can be described as socio-technical. The two most important methodological traditions that we draw on are the social sciences and the engineering sciences." She goes on to note that information science addresses three large-scale questions:

- What are the features and laws of the recorded-information universe?
- How do people relate to, seek, and use information?
- How can access to recorded information be made most rapid and effective?

The methods most commonly used in addressing these question are statistical and other, philosophically based, analytic techniques for the first question, social science techniques for the second, and engineering techniques for the third.

Both fields make considerable use of systems analysis techniques in the conduct of their research and development activities. In recent years there has been an increasing emphasis on an expansive view of the information system, including producers, distributors, and users of information, and on a holistic view of activity within this system, taking into account all factors. Evidence of this general approach in biomedical informatics can be found in Friedman and Wyatt (2006) and in Brender (2006). A noteworthy study in the area is the 7-year qualitative study by Ash et al. (2005) regarding computerized physician order entry implementation at successful sites.

Another methodology frequently called upon in both fields is the usability study, which seeks to determine the ease with which people can use a particular tool, particularly a computerized tool, to achieve a goal. Here, the early work was done primarily within information science, by researchers such as Shneiderman (2004). More recently, a number of usability studies have been done in biomedical informatics, as for example the study by Starren et al. (2006) concerning clinical trial management systems. Certainly the importance of such studies is likely to increase in both fields, as information tasks become more complex and the computer techniques available to increase usability become increasingly sophisticated.

Evaluation methods in biomedical informatics have grown out of the medical tradition of evolutionary assessment, as in the case of drug trials, which proceed from case studies to rigorous quantitative analysis and then randomized controlled trials (Friedman, 2006). In contrast, evaluation methods in information science have taken a more social science-oriented approach, using exploratory techniques such as case studies, observation, and interviews.

**Training in information science and biomedical informatics**

In essence, training in both fields began with the apprentice model, as has been true for many other professions. Early training of librarians took
place in libraries but moved to the university setting over a hundred years ago. As library science expanded to the broader field of information science, many of the same schools began to offer programmes in both library and information science or a combination of the two. Other cognate fields also embraced the emerging discipline of information science, with expansions of programmes in their areas, such as computer science, communications, and business. Today a wide array of degrees are available in library information science and in information science, with an emphasis on the master's level degree.

Early training in biomedical informatics took place in the laboratories or clinical settings of the first workers in the field, who were predominantly physicians. Generally the model followed was the post-doctoral or fellowship approach that is characteristic of advanced training in many areas of medicine. As this field grew, and as more basic principles and methods were identified, training came to include more formal coursework. Degrees were added, again focusing on the master's degree.

Schools providing training in information science in the United States include a number of schools that began with a focus on librarianship. Many, but not all, of these have programmes in librarianship accredited by the American Library Association's Committee on Accreditation. A nexus for programmes that identify their primary focus as information is the recently formed I Schools caucus (http://www.ischools.org/). Several of these schools have their origins in programmes of computer science, but at least two – Indiana University's School of Information and Pennsylvania State University's College of Information Science and Technology – have recently been created to address the broad area of information science.

Training in biomedical informatics has benefited greatly from funding from the NLM, which has funded a number of programmes for over 20 years. The largest of these programmes includes 20 schools that provide training either across biomedical informatics or in more specific subareas, such as imaging informatics. There are a number of educational programmes in nursing informatics, and certification as an informatics nurse is available through the American Nurses Association. The Robert Wood Johnson Foundation has recently partnered with NLM to fund four institutions to create training programmes in public health informatics. There seems to be an increasing amount of education in the bioinformatics area, with a number of academic institutions announcing new degree programmes in this area.

Is there convergence in this array of educational programmes in the two fields? At this point, the most observable activity is the increased coordination of separate programmes in universities having multiple offerings, for example those at the University of Pittsburgh and the University of Washington. The formation of broader-based programmes, such as that at the University of Indiana (http://www.informatics.indiana.edu/), may be a harbinger of more to come.

Professional organizations and accreditation agencies

As in other disciplines, there are a number of professional organizations that represent the individuals and organizations involved in information science and biomedical informatics. Key organizations in the US in the general information science field include the American Library Association (http://www.ala.org/) and the American Society for Information Science and Technology (ASIST). Educational associations include the Association for Library and Information Science Education and the I-Schools consortium. In biomedical informatics, the primary professional organization for academics is the American Medical Informatics Association (AMIA), with the Healthcare Information and Management Systems Society being the professional home for many practitioners in the field. AMIA has recently established an academic forum to promote the development of biomedical informatics as an academic discipline by identifying best practices and serving as a locus for collaborative activities.
Accreditation of educational programmes in a professional discipline reflects an agreed-upon set of standards for the training of individuals in the field, including specification of the values of the discipline and the basic competencies that identify a professional in the field. Such processes are generally carried out through the professional association in the field.

Accreditation of information science programmes is currently carried out by the American Library Association, with the scope of the accreditation activities evolving to some extent as schools of library science have evolved into schools of library and information science. The process remains mainly focused on programmes leading to degrees in librarianship or combined library and information science degrees. A few schools that provide separate library and information science degrees also submit their information science programmes for accreditation. In the context of a world view in which library science is a subset of information science, this approach amounts to a broadening of the field and the scope of ALA's accreditation programme. Alternately, when information science is viewed as separate from library science, ALA's accreditation process is inadequate to define education in the field.

Accreditation processes in medicine are well established and cover a variety of subdisciplines, but there is no programme for accreditation of biomedical informatics programmes. AMIA's recent creation of an academic forum for the exchange of information among programmes offering degrees in biomedical informatics suggests a maturing of the field, and one of the items of interest for the academic forum is the consideration of an accreditation programme.

The parallelism here is that both biomedical informatics and information science, broadly defined, can be said to be at a developmental stage at which it is appropriate to consider accreditation programmes as well as other signposts of an established discipline.

Conclusions

The notable convergence of biomedical informatics with library and information science in recent years is a direct reflection of the common foundation of these two disciplines, both of which are focused on the organization, storage, retrieval, and application of data, information, and knowledge. Despite the fact that these information-focused areas have emerged from different sociocultural and historical backgrounds, the boundaries between the two fields continue to blur, and there is clearly potential for beneficial cross-fertilization in the future.

Both biomedical informatics and information science are multidisciplinary, drawing on many of the same domains of knowledge, including philosophy, psychology, computer science, and communications. Both fields are also meta-fields, cutting across conventional academic or content disciplines from the information perspective.

These two fields clearly exhibit a considerable number of parallels, particularly with regard to their development, the approaches followed, and the discoveries that have been made. There are also some parallels in their approaches to training new professionals in their respective fields, and some indication of collaboration in this arena.

We suggest that a continued exploration of the parallels between these two information-based fields, begun in the present paper, will allow each
field to benefit from the other. Particularly useful areas of collaboration include the sharing of research findings and methods, identification of basic theories and models, and delineation of competencies and training approaches.
References


---

**How to cite this paper**


---

**Find other papers on this subject**