Viewpoint ■

(Bio)Medical Informatics in the Next Decade

THOMAS C. RINDFLEISCH, MS

Abstract Even though medical informatics is most often viewed from the perspective of its host disciplines in clinical and biologic medicine, it has an identity and agenda of its own. This paper is an attempt to promote discussion about the long-term role and agenda for medical informatics as a discipline into the next decade. The discussion has two main lines of argument, one about the "engineering" goals of informatics and the other about the "basic research" goals. These are, of course, influenced by ongoing developments in computing, communications, and software infrastructures, but informatics is now mature enough that many of its goals transcend these changes.

■ **JAMIA.** 1998;5:416–420.

In an earlier article we tried to project the priorities for medical informatics into the next decade from the perspective of the intersection and synthesis of informatics with clinical research and with research in the basic biomedical sciences.[1] The core argument in that paper is that clinical research and genomics research are on a "collision course" for merger and synthesis and that they must unavoidably share computer-based infrastructure, encoded and structured information methodologies, information-analysis and decision-support tools, and a communication-based collaborative framework. This paper is an attempt to view informatics a bit more expansively and to project future directions for medical informatics as a discipline unto itself. This sketch of the long-term role and agenda for informatics is not intended to be exhaustive in any sense. It is intended primarily to promote discussion. The presentation is broken into two main sections, one discussing the "engineering" goals of informatics for the next decade and one discussing the "basic research" goals of informatics. A concluding section attempts to project the future of informatics as a separate discipline, especially in relation to the expected broader deployment of information technologies through biomedicine.

Affiliation of the author: Stanford University, Stanford, California.

Correspondence and reprints: Thomas C. Rindfleisch, MS, Lane Medical Library, L-109 Rte 1, Stanford University Medical Center, Stanford, CA 94305-5123. e-mail: ⟨tcr@stanford.edu⟩.

Received for publication: 4/23/98; accepted for publication: 5/21/98.

An analysis such as this is unavoidably speculative (and even personal), because ten years is a long time in any modern technologic or scientific discipline. This is especially true for information-related fields drawing on the rapid advances in computing, communications, and software. Much evidence of the computer revolution of recent decades already exists in medical settings—computer screens and keyboards on almost every desk, high-speed network communication infrastructure throughout most campuses and workplaces, and World Wide Web–related Uniform Resource Locators (URLs) advertised everywhere for access to information from scholarly publications to entertainment. Many aspects of modern health care systems, including clinical laboratories, diagnostic radiology, nuclear medicine, cardiovascular testing, administrative and business functions, and genomics research, could not operate without computers.

The Context of the Past

As we think about projecting informatics forward ten years, it is helpful to remember that the current *practical* forms of computing and communications—cheap and powerful personal computers, fast and compact server machines, graphical user interfaces, powerful software, the Internet, and so on—have themselves been widely used for little more than ten years. The integrating technology of the Web is only four years old, and serious digital publication technologies are only two years old. This, of course, belies the fact that the underlying technologies were envisioned and invented in the "laboratory" 20 to 30 years ago—the first integrated circuit chips appeared in the early 1960s,[2] Engelbart[3] gave a demonstration of graphic computing with a mouse and Nelson's hypertext ideas[4] in 1968 (using a mainframe computer), testing of the 64-node ARPANET began in 1968–69 (the same era in which electronic mail, file transfer, and remote computing were invented),[5] and the first personal workstations and local area networks date from the mid-1970s.[6–8] So the context of this projection is that the technologic underpinnings of computing and communications will continue to evolve rapidly, but it will take a long time for many new ideas to diffuse into broad use. Grove's law will continue to apply to computing systems—that is, a factor-of-2 improvement in important parameters (like speed, storage capacity, screen quality, and cost) will occur roughly every 18 months. In the next 5 to 10 years, communication network speeds are projected to increase tenfold, and more powerful compact and portable systems will replace the simple hand-held computers and wireless communications of the late 1990s. Even older electromechanical devices have been heavily affected—for example, in the past six years, the average data storage capacity on disk drives has jumped 18-fold while the price per megabyte has fallen 50-fold, from more than $5 to just 10 cents. These changes will continue well into the next decade and will set the context for projecting how informatics will affect biomedicine.

We cannot know when the next "World Wide Web" phenomenon will take place, however. In fact, such rapid and explosive changes are very rare and most often represent opportunistic syntheses of earlier ideas. Progress is generally much slower—a fact that underscores the need for a steady, long-term program (with financial support) for informatics research. It is perhaps most sensible to forecast that computing in 2008 will be much like today, but more so—e.g., universal high-performance communications and computing; powerful tools for information authoring, searching, and sharing; convenient interpersonal communications; online services like commerce, education, and entertainment; reliable security and cryptographic technologies; and increased robustness and reliability. Still, examples of technologies that, if perfected, could have a profound influence (like the Web) might be fast and reliable continuous speech-recognition systems and natural-language processing systems.

A Structure for Projecting Informatics Goals

Many disciplines, including informatics, have a basic science part and an applied or engineering part—as do physics, biology, chemistry, mathematics, and computer science. The "basic science" part elucidates the structure, principles, and base technologies of the discipline. The "engineering" part seeks to synthesize and utilize these principles in real-world applications. So one way to frame the long-term goals of medical informatics is in terms of its science and engineering goals. But medical informatics sits at the interface of computer science and its applications in medicine. So by "science" and "engineering" goals do we mean relative to computer science (with its subdisciplines of distributed systems, databases, human–computer interfaces, software engineering, intelligent systems, theory of computation, and so on) or medicine (with its subdisciplines of the various facets of clinical care, clinical research, biologic science research, and education)?

The answer is both, and perhaps we need a matrix, in which one axis represents medical applications and the other computer science technologies. At each intersection we would elaborate the application needs, the application impact, the relevant technical ideas and opportunities, and the technical risk. With such a matrix, we would look for important informatics "engineering" work at the intersections with strong need, high impact, existing technologies and opportunities, and relatively low risk. We would seek important informatics "science" work at the intersections with strong need, high impact, potential technical ideas and opportunities, and relatively high risk. (The other intersections might be interesting from either the medical application or computer science points of view but probably would not be fundable or would not get enough management attention to make a difference.) It might be argued that informatics work that represents a solution in one matrix intersection and has broad impact in others should be regarded as even more important. Some examples of such impactful work might be defining workable information-exchange standards; building a practical distributed software architecture that facilitates software reuse and effective validation; developing feature-matching algorithms that are useful for comparing both DNA sequences and clinical cases; and developing reliable and workable systems for speech input or natural-language understanding. What follows falls far short of a full analysis of this framework for organizing long-term informatics goals. Rather, it gives the author's views about the most important engineering and science agenda items for informatics—drawn largely from the context of the intersection of clinical and genomic research.

An "Engineering" Agenda for Informatics

From a medical informatics ''engineering'' perspective, many efforts are already underway that will probably not be resolved in ten years but that deserve continued effort. Among these are:

- Fully deploying electronic medical record (EMR) systems that contain information encoded by means of shared ontologies, interoperate seamlessly, are integrated deeply into health care settings, are reliable and secure, and have architectures designed to be scalable and extensible to meet future information needs in any health care setting.
- Facilitating the development of increasingly structured research databases derived from EMR repositories to support clinical trial studies, meta-analyses, outcome analyses, quality-of-care evaluations, and other health services research studies. As with EMRs, these data repositories must be designed to be scalable and extensible to meet future performance and information needs in any health care setting.
- Making collaboration tools and resources broadly available to facilitate the increasing interactions among genomic scientists, clinicians, and informaticians.
- Facilitating the integration of new diagnostic and therapeutic technologies from genomics research, including instrumentation, information types, and decision-support tools.
- Providing better user interfaces, and extending the powerful integration and navigation ideas of Web-based designs with mobile systems, voice input and output, images, natural language, and so on.
- Developing and refining necessary open standards for information acquisition, encoding, and interchange, based on shared ontologies and leading to more effective industry-supported systems.
- Putting in place open, robust, and effective security and privacy measures.
- Facilitating education at all levels for health care providers, with the goal especially of supporting broad lifelong learning.
- Accommodating better patient access to information and care advice from all settings—home, clinic, and hospital.
- Evaluating clinical effectiveness, cost-effectiveness, and productivity measures for informatics systems.
- Continually integrating newly evolving communication, computing, and software technologies as they become practical.

These directions, of course, must evolve in the context of the growing emphasis on evidence- and guideline-based medicine using the results of clinical research, the growing impact of molecular biology and genomics research on clinical practice, the business of health care, and the increasingly global nature of public health issues. We can expect the need for powerful information- and situation-analysis resources only to increase. This agenda is clearly a full plate.

The "Science" Agenda for Informatics

From an informatics (or computer) ''science'' point of view, we continue to seek ever more capable computer systems that can help complement human intelligence in coping with the complexities of professional problem solving and the ongoing rapid growth of knowledge. As with the engineering component of informatics, the science side is also working on many of the same (difficult) issues that were recognized during the last two decades. Among these are:

- Developing flexible ways to represent, acquire, structure, and use knowledge of various sorts in computer systems for task-oriented applications. These methodologies must bridge specific applications so that the representations and the knowledge they encode can be shared widely and reused.
- Facilitating human collaboration with computers—that is, accommodating human speech and natural language and facilitating views of knowledge and information useful to human reasoning processes. These capabilities need to support human reasoning with problem-solving aids, augmenting human memory and helping in the retrieval of relevant information, taking advantage of human perceptual capabilities, and allowing the automation of tedious and lengthy processes beyond human stamina.
- Facilitating computer collaboration with computers—that is, accommodating different representations of information and nomenclature among machines (including some legacy systems) and facilitating the sharing of information so that meaning is (declar-

atively) represented and preserved in support of distributed applications. This work must anticipate the full miniaturization of computers so that they may become ubiquitous—that is, associated with almost all objects in our environments.

- Facilitating human collaboration with humans—that is, accommodating human social needs and conventions for synchronous and asynchronous interactions, facilitating memory, and managing shared representations of information and knowledge.
- Supporting human retrieval, analysis, and use of complex data types such as patient cases, clinical studies, biomolecular structures, images, sound recordings, and software modules, so that essential features of each data type are represented explicitly, extracted automatically, and used to relate the similarity of one data element to another at a conceptual level comparable with human understanding.
- Integrating computer-based information and problem-solving systems seamlessly into human work environments so that there are minimal impediment and maximal support of human goals and workflows.
- Learning to use new genomic, biochemical, and structural knowledge to construct more and more accurate models of biologic systems and subsystems (e.g., models of cells, populations of cells, organs, organ systems, and organisms) that can be used for research, education, clinical diagnosis, drug design, therapy planning, and such. Display of such models would probably take advantage of advanced graphic representations, including virtual reality environments.
- Learning to build, maintain, and deploy computer-based information systems that will support lifelong learning in ways that accommodate individual human styles and needs for learning.
- Developing methodologies to use computer-based understanding of the contents of information resources to better protect the privacy, integrity, and proprietary nature of sensitive information.
- Facilitating human efforts in the design, engineering, validation, and maintenance of increasingly complex, distributed, and interoperating software systems needed for biomedical applications. Essential goals of this work would include enabling broad reuse of software modules and facilitating construction of scalable high-performance and high-integrity systems.
- Developing standards and tools to ensure that archive-quality information and knowledge will be shareable, will be preserved, and will remain accessible over time.

Informatics as a Discipline

In this future view, we may ask what will happen to medical informatics as a discipline. As computer-based information management systems take a broader hold in medicine—particularly clinical medicine, which has been slower to respond—we can expect the locus of much computer work to shift to the individual domains of application. We see this process occurring in other complex fields such as engineering, in which the informatics work of mechanical engineering or electrical engineering, for example, is done largely in those domains rather than in a core computer-science area. This process is already under way in medicine as well, where most of the computer-related work in molecular biology, genetics, radiology, cardiology, clinical laboratory studies, and pathology is done in the application area rather than in a central informatics group. Thus, the identity of medical informatics, especially in its engineering track, can be expected to diffuse with the broader adoption of computing and information methodologies. We can expect successes in medical informatics to show up as broadly adopted technologies and methodologies whose applications in new domains are no longer associated with informatics.

Still, one can argue that there exists a core medical informatics research agenda that is drawn from the various disciplines of biomedical applications and is fueled by the technologic base of computer science. This core agenda rightfully deserves to be labeled "informatics." One might include in this unique informatics agenda topics such as research and software frameworks for knowledge representation and (re)use, user-interface design principles, computer-based collaboration infrastructure, software-engineering and complex-system validation, and other topics from the list of science goals given earlier.

If this core agenda proved to be identical with the agenda of computer science, it might be argued that as computers come into wider and wider use in medicine, then medical informatics would disappear as a separate discipline. Indeed, we do not have any very clear existing examples to use as guides—that is, there are no departments of business informatics, or legal informatics, or social science informatics that light the way. The nearest analogy might be the relationship between the departments of statistics and biostatistics. Even though they are deeply connected, the problems of statistical analysis in biomedicine are so important and so difficult that a special research and service department (biostatistics) is needed. It is not likely that this argument could be made for all of what is now included under the rubric of medical informatics, but there are strong arguments for the core parts. The development of many systems and processes—biomedical knowledge-based decision-support systems, image processing and analysis systems, biomolecular and biologic modeling systems, perhaps aspects of human–computer interfaces and collaborative systems, data structure/encoding/interchange standards, and some aspects of infrastructure such as security framework integration—all seem to justify the existence of a field of medical informatics because of the particular difficulty and complexity of their biomedical forms. As with biostatistics, medical informatics also rightfully has a persistent service role in educating medical students and graduate students about information sciences and advising on applications work.

In this attempt at a medical informatics vision the author benefited greatly from talks with Russ Altman, Bruce Buchanan, Milt Corn, Kaz Kulikowski, Mark Musen, and Pete Szolovits. They are not to blame for what I did with our discussions.

References ■

1. Rindfleisch T, Brutlag D. Directions for clinical research and genomic research into the next decade: implications for informatics. J Am Med Inform Assoc. 1998;5:404–11.
2. Kilby J. The history of the integrated circuit [Plenary Lecture]. Proceedings of the 44th National Symposium of the American Vacuum Society, 1997. Available at: http://www.vacuum.org/symposium/plenary.html. Accessed July 13, 1998.
3. Engelbart D. Augmenting Human Intellect: A Conceptual Framework. Menlo Park, Calif.: Stanford Research Institute, 1962. Summary Report AFOSR-3223 under contract AF49(638)-1024, SRI Project 3578, for Air Force Office of Scientific Research. Also available at: http://www.histech.rwth-aachen.de/www/quellen/engelbart/ahi62index.html.
4. Nelson T. Dream Machines: New Freedoms Through Computer Screens: A Minority Report. Chicago, Ill.: Hugo's Book Service, 1974. Republished by Microsoft Press, 1988.
5. Leiner B, Cerf V, Clark D, et al. A brief history of the internet. Version 3.1. Last revised Feb 20, 1998. Internet Society Web site. Available at: http://www.isoc.org/internet/history/brief.html. Accessed July 13, 1998.
6. Metcalfe R, Boggs D. Ethernet: distributed packet switching for local computer networks. Commun ACM. 1976;19(7):395–404.
7. Thacker CP, McCreight E, Lampson B, Sproull R, Boggs D. Alto: A Personal Computer. Palo Alto, Calif.: Xerox Corporation, 1979. Xerox Palo Alto Research Center Technical Report CSL-79-1.
8. Wadlow T. Fifteen years ago in BYTE: the Xerox Alto computer. BYTE [online magazine]. September 1996. Available at: http://www.byte.com/art/9609/sec4/art3.htm.