[∞]JANE GRIMSON, WILLIAM GRIMSON, AND WILHELM HASSELBRING

The inability to share information across systems and between care organizations is just one of the major impediments in the health care business's progress toward efficiency and cost-effectiveness.

THE SI CHALLENGE IN HEALTH CARE

ealth care is an information-intensive business generating huge volumes of data from hospitals, primary care surgeries, clinics, and laboratories. Yet much of this data continues to be processed manually in spite of decades of experience in the success-

ful application of information technology (IT) in other information-intensive industries. There are a number of reasons for this state, including underinvestment in IT (especially clinical computing), lack of political will, fragmented markets with inadequate revenue streams to support development of new systems, and lack of standards or slow adoption of standards where they do exist. In addition, there are some specific challenges relating to the use of IT in health, such as the complexity of medical data, data entry problems, security and confidentiality concerns, the absence in many countries of a unique national patient identifier, and a general lack of awareness of the benefits-and risks-of IT. Historically, health care organizations have consisted of independent and autonomous units with little clinical benefit perceived for sharing of information, which in turn fostered a climate of independence in the use of IT. SI did not therefore have a high priority.

However, the pressure on the health care business to change is mounting. The gap between the demand for health care from an increasingly well-informed and expectant public, and the ability of the state and health care organizations to meet this demand is widening all the time. Efficiency and cost-effectiveness-balancing quality of care with cost containment-are two major driving forces behind this need to change as, for example, with managed care in the U.S. In addition, the single doctor-patient relationship is being replaced by one in which the patient is managed by a team of health care professionals each specializing in one aspect of care. Such seamless or shared care depends critically on the ability to share information easily between care providers. Indeed it is the present inability to share information across systems and between care organizations automatically, that represents one of the major impediments to progress toward shared care and cost containment. Paper records, no matter how useful to a single clinician in one location, do not really facilitate shared care. It is apparent, however, that a strong clinical motivation, something that's been missing in the past, to share information is beginning to emerge. Strategically there is need to take a more business process view of health care delivery and to identify the appropriate organizational and information infrastructures to support these processes. Business Process Reengineering [9] is central to ensuring that strategies are aligned through the clarification of needs, plans, and priorities, identifying opportunities and enabling even radical changes to existing processes to be undertaken. A key driver is the pressure to ensure that the care delivered to the patient is based on best-practice, so-called evidence-based medicine. Best-practice is increasingly being encoded in the form of clinical

guidelines and protocols [11] that drive the delivery of health care. This is an example of the business processes.

Not surprisingly, at the heart of the application of IT in health care is the Electronic Health Care Record (EHCR) in which medical data about the patient is recorded. Access and use of the EHCR is fundamental; the insertion of data into the record irrespective of its source (information system, manual entry, monitoring device, and so forth), maintaining and building it, and then making the record available to those who need it, presents obvious integration challenges. The linking of records to clinical guidelines and protocols is essential if best-practice is to be embedded as an integral part of the health care delivery process and if the problems associated with widespread variations in treatment costs and outcomes are to be addressed.

t a technical level, the health care industry, like other industries, has been faced with the challenge of moving from mainframes to L client-server computing with PCs, and more recently to thin clients. Problems are exacerbated by the highly heterogeneous and distributed nature of health care computing, generally. The requirements of the different user sectors, from the primary care physician with a few hundred patients to the large acute hospital with thousands of patients, from the X ray department to the intensive care unit, are so varied that it is inevitable a variety of computing solutions will have to be adopted. Integration of these diverse systems therefore remains problematic with a number of competing approaches, none of which alone represents the total solution. Messaging is widespread and relatively easy to implement, but can only provide loose coupling of systems, requires a high degree of interface engineering and support, and is generally not scaleable. Approaches to Enterprise Application Integration aim at alleviating the integration of specific Enterprise Resource Planning (ERP) systems within and among organizational boundaries. A single data repository (warehouse) approach is said to protect existing investment, but is essentially unidirectional with data uploaded from operational systems to a warehouse where it is generally only available in read-only mode. Federated database systems in principle support bi-directionality and also protect existing investment, but generic solutions adaptable for health care—in spite of more than two decades of research—have proved elusive. At the most ambitious level are distributed component-based architectures in which a set of common and domain-specific services, encapsulating both data and functions, support the business processes. This last approach requires a

high degree of standardization, while allowing a "mixand-match" environment in which vendor independence is achieved and existing investment protected.

Patient Records

There are a variety of users in health care with varying needs for how data is represented, stored, and used, and even though the focus of this article is on supporting the clinical process—the fundamental one in health care—most of the comments are applicable to other non-clinical (such as administrative and financial) processes. When considering patient records, the starting point is the paper record since it has many advantages. It is familiar, portable, and can be easily browsed or scanned. However, in the climate of modern health care delivery, it has a number of major shortcomings.

The idea of computerizing the health care record has been around since the early 1960s when hospitals first started using computers. Initially, computing systems in hospitals and elsewhere throughout the health care system focused on supporting financial processes. Hence there was a need to record basic data about the patient in order to ensure they were correctly billed for the treatment they received. Since these systems already stored such basic patient data, it was natural to seek to extend them to include more clinically relevant data. At the same time, hospital laboratories were becoming increasingly computerized, which meant test results were available in electronic form and could be integrated with the basic demographic information. However integrating the basic patient data gathered for financial purposes and test result data produced for the efficient operation of the clinical laboratory is not the appropriate way to achieve the computerized record. Health care records, whether manual or electronic, are much more than simply arbitrary collections of patient data. First, the data in a record is structured, for example, chronologically, by source (i.e., all laboratory reports are together), by problem, or some combination of these [12]. They serve a variety of purposes, including a single access point for relevant, active data about the patient, an informal workspace for recording ideas and impressions, a medico-legal archival record of what has happened to the patient, and a means of communication between health care professionals. The advantages of the electronic health care record over its paper-based counterpart are clear-it is always available, information can be transferred, and it can support different views of the record for nurses, doctors, physiotherapists, and other users. In principle, the electronic record can also be linked to evidence-based, best-practice guidelines to provide

decision support. However, even with a computerized record approach agreed, the problem of populating the record remains; this is where SI comes into focus. The key objective of SI in health care is to provide access by authorized users to the relevant health care data about a patient in the form of a record regardless of the source of that data or how it was inserted into the record.

The data, which feeds or populates the EHCR, already resides in a variety of highly heterogeneous and autonomous information systems. Simply integrating this data does not result in a valid EHCR—it is a user-defined constrained aggregation of data. Effectively what is required is a federated databasetype approach in which the individual participants in the federation (the local systems) are self-contained autonomous systems, but together form part of a wider picture—the federation. How then does one capture this concept of the Federated Health Care Record?

Achieving Federation

Current practice shows there are in essence three approaches: message-based, data warehousing, and common architecture.

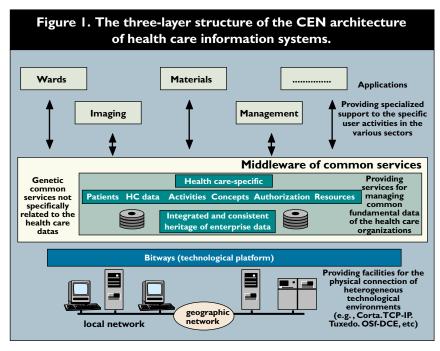
The message-based approach, by far the most prevalent in the health care community today, relies on defining a set of standard messages that allow different health care information systems to exchange messages carrying data (www.HL7.org). A number of health software vendors have developed so-called integration or *interface engines* that allow heterogeneous health information systems within a hospital or region to exchange information via standardized messages [3]. Such integration engines provide a useful way of solving the basic communication problems between systems, but they do nothing to address true interoperability and integration of information. This approach has worked well and has been effective, but when the number of possible interactions between systems increases, such as happens with shared care, then the limitations of scalability become apparent. Also, such an approach cannot be regarded as SI but rather as inter-system communication.

Data warehousing offers an alternative that allows the data from individual systems to be integrated and homogenized in a data warehouse. This has proved in health as in other domains to be a useful way of bringing disparate data together in a single repository. Moreover, since the data is integrated it is possible to use the data as the basis for an integrated EHCR. However, a major drawback of this approach is that the data in the participating information systems is duplicated in the warehouse. It is natural; if the warehouse is used to support the EHCR and to deliver relevant patient data in an integrated fashion to the clinician's desktop, users will want to update the data and insert new data. But data warehousing is not designed to support operational systems, and such an integrated EHCR system available through the warehouse effectively soon becomes an operational system. While data in the warehouse may be enriched in a variety of ways, the technology is basically intended to support applications in which read-only access to the data is required. When the warehouse becomes an operational system, the typical problems of keeping data that is replicated within the hospital consistent arise [5].

The third alternative is to use federated database management system technology adapted to the health care domain. Even loosely coupled Federated Database Systems depend on a common canonical model into which the underlying data models can be mapped in order to present a uniform view of the data at the federation layer. Effectively, this means agreeing a common domain-specific EHCR data model so that data from the participating information systems has a predefined place in the overall model or architecture of the EHCR. It is reasonable to base those common data models on approved health care standards.

A number of attempts have been made to develop such a common model, but widespread adoption across a range of clinical disciplines has proved slow. For example, the Good European Health Record [6] concentrated on providing a comprehensive model that captures all the rich ethico-legal semantics of the EHCR, while the W3-EMRS [8] placed more emphasis on the ability to share records securely between institutions and agreeing on a common data set for the shared record. However, achieving consensus is difficult at a local level let alone nationally or internationally. This is perhaps not surprising, as it is inconceivable that all departments in a single hospital (never mind all hospitals in a region) would agree to a single structure for the record. An alternative and potentially more flexible approach is to present the users with a set of constructs from which they can build their own record. These building blocks are carefully defined so that the context in which the data was originally generated is preserved.

To date it has proved difficult to reach agreement on the precise definition of a health care record architecture that can lead to realizable record systems. The Synapses project [4], funded under the EU's 4th Framework Program, has addressed this problem in a pragmatic manner, building on and contributing to the European standards work of CEN TC251 (Comitié Européen de Normalisation



Technical Committee 251; www.centc251.org). The Synapses approach is to base the sharing of data (records) on a common data model, the Synapses Object Model (SynOM). In essence the SynOM provides a set of abstract "building blocks" that can be used to construct the shared or Federated Health Care Record (FHCR). The SynOM provides an aggregation mechanism for the record, in which the aggregation hierarchy contains the object classes in the model. In addition to their use in the exchange format, the classes of the SynOM form the basis for the Synapses Object Dictionary (SynOD). The SynOD is an active data dictionary/directory containing definitions (and locations) of shareable userdefined clinical objects that conform to the rules and constructs of the SynOM; put another way, the SynOD object definitions are expressed using only instances of classes derived from those found in the SynOM. The records thus constructed are indeed aggregations of information available from the underlying source information systems, but are constrained aggregations in accordance with the requirements of the record architecture.

Integration

From an IT point of view, delivery of health care can be supported by generic middleware components like DCOM and CORBA, which provide solutions for the technological interconnection of distributed systems. A middleware approach is an attractive proposition since a health care organization can be considered as a set of disparate users, performing diverse tasks, but all needing to rely on and share a common data set and using a common set of business services. Such common data and facilities should be accessible to applications through standard interfaces. From the clinical point of view it is essential to identify the constraints which need to be put in place to provide those elements of the common data that various users require without making every piece of data available to everyone. Some control in addition to "authorization" needs to be in place. And furthermore, there is the need to be able to construct meaningful aggregations of data to support the EHCR. These issues are directly addressed by the two middleware approaches, COR-

BAmed and the Healthcare Information Systems Architecture (HISA).

The overall mission of CORBAmed is "to improve the quality of care and reduce costs by use of CORBA technologies for interoperability throughglobal health care community" the out (www.acl.lanl.gov/OMG/CORBAmed). In addition to domain-independent services specified by the Object Management Group (OMG), a set of health care domain-specific services have been defined including Person Identification Services, Lexicon Query Services, Clinical Observation Access Service, Clinical Image Access Service, and Health care Resource Access Control. However, while there is considerable interest in the work of CORBAmed, the rate of progress is slow, especially in Europe.

The CEN ENV 12967-1 standard HISA also provides a middleware solution (www.centc251.org). In this standard the information system is modeled as three cooperative layers (applications, middleware, and bitways), each individually responsible for addressing specific design and operational aspects of the information system, as shown in Figure 1. The bitways refer to the technical infrastructure providing distribution and connectivity in the generally heterogeneous environment and are not specific to the health care domain. The applications refer to the software components identified and associated directly with the user activities, and in very general terms include data operations, the use of embedded knowledge and presentation of information on computer screens and other devices. In addition, other components offering generic support, such as security and an EHCR record service, are ideally placed in the middleware layer.

The middleware layer represents the central element of the system, providing a holistic infrastructure where all applications may be connected. This layer, through its services, facilitates the management of the information common to the whole organization through stable and technology-independent interfaces. The intended evolution of the HISA standard focuses mainly on the common services. This will guarantee the support and the consistency of the overall information heritage of the health care organization, by means of direct access to the elementary data, together with the more complex "business objects" capable of implementing complete functions. In this respect, the EHCR is handled through a set of user-defined health care objects, that are mapped to the data representations of those objects found in the "common heritage" middleware layer using the SynOD or its equivalent, the COAS in CORBAmed [3].

Messaging

In terms of the interoperability envisaged in both CORBAmed and HISA, it can be noted that the exchange of messages is not sufficient, since this approach can only lead to a loosely coupled interworking among the different sectors in an organization, without meeting the requirements of the health care organization as a whole. Nevertheless the use of messages has made a very significant impact in bringing IT support to health care [10].

The HL7 (www.HL7.org) integration approach is pragmatic and achieves data integration, which in turn depends on the acceptance of a detailed data model. As already stated this does not provide interoperability, but it achieves much through providing acceptable integration costs. The widespread adoption of HL7 throughout the U.S. in particular attests to the benefits that are achievable.

There is growing interest in the use of Extensible Markup Language (XML) (www.w3.org) which allows the structure and content of the record to be described in the form of a locally specified Document Type Definition (DTD). Thus each participating system has its own local DTD aligned with whatever local EHCR system is in use. It is then technically possible to map data/records from one EHCR system to another through DTD conversion, thereby allowing records or fragments of records to be exchanged between heterogeneous systems. This approach is not offering true interoperability but it has the potential to unblock the path to progress. Indeed once systems start communicating through the use of XML DTDs it can be asserted that the motivation to standardize the underlying record architectures will increase.

An example of the effective use of XML to support the meaningful exchange of patient records has been developed in the SynEx project (www.gesi.it/synex), in which a standard DTD-SynXML DTD-has been defined, fulfilling the need for a common syntactical platform that enables the exchange of structured patient data between diverse sites. The approach adopted is to define a DTD that describes the classes and aggregation of the SynOM, which is then used to build site-specific SynODs. The documents that are exchanged on request use the Hypertext Transfer Protocol (HTTP). The need to allow some level of sitespecific specialization is inevitable in the short-term. However a detailed and predefined reference DTD is still required to fulfill the role of the SynOD in the federated record server. Site-specific mapping to and from this reference structure will be the basis to manage the flexible exchange and automatic interpretation of the records [7].

Enterprise Application Integration

Hospitals are complex organizations. As such they have Enterprise Resource Planning (ERP) systems deployed to manage the hospital's data and processes. ERP systems promise to provide an integrated application environment with seamless access to a single unified database. Different departments within a hospital may have deployed different ERP systems that should be integrated in order to support the business processes adequately. For supporting shared care, health care organizations need to integrate their ERP systems horizontally to constitute some kind of networked, virtual health organization. Integration of existing legacy systems, packaged software (ERP systems), and newly developed systems is far from trivial.

National legislation often has a significant influence on the design and use of EPR systems. At German University Hospitals, for instance, SAP R/3 (www.sap.com) is the market leader for ERP systems, while in the Netherlands, HISCOM/Baan (www. hiscom.nl) dominates. In the U.S. no real market leader can be identified, but major players include Cerner (www.cerner.com), Lawson (www.lawson.com), People-Soft (www.peoplesoft.com), and SAP.

As an example of Enterprise Application Integration (EAI), consider the situation in a typical German University Hospital. The SAP R/3 module IS-H is specifically designed to support patient data management and some clinical functions. In addition, R/3 modules for material management are used for pharmacy, and modules for financing and investing are employed in administration. Integration with other R/3 modules can be achieved within the central database according to the general principle of SAP "everything in one database." Thus, integration is solved by deploying a single integrated system. This approach corresponds to the origin in the system of coordinates for the integration problem dimensions discussed in the introduction to this special section on Information SI (see Figure 3 in Hasselbring's "Introduction" to this section). However, in all German University Hospitals, we also find enterprise information systems from other vendors, for instance in radiology, and operating theatres. Usually, it is not possible or desirable to replace these systems with R/3 modules for a variety of reasons. So, other mechanisms are required to integrate such systems. The situation becomes even more challenging when the integration of enterprise applications from different health care organizations is required. A tight integration within one ERP system is not feasible in such a setting, because the local ERP systems must remain autonomous to a great extent, reflecting the autonomy of the corresponding organizational units.

One of the key challenges with EAI is to gain access to the ERP systems. ERP systems are usually built as monolithic systems not intended to cooperate with other systems. Wrapping is an established technique, for instance, by providing CORBA IDL interfaces to ERP data and services, but it is not as simple as it might at first appear. The strength of EAI approaches comes with offering (generic) adapters for specific ERP systems. These adapters are only usable for these specific ERP systems, but for integration purposes these customized adapters are very useful. A typical vendor of software for EAI is Tibco (www.tibco.com). Adapters for standard ERP packages such as SAP R/3 and PeopleSoft are offered together with message queuing services that connect to CICS, COM, and CORBA. Process coordination is supported through cross-system transaction monitoring and publish/subscribe notification services. However, the main strengths of EAI vendors lie in the adapters that encode much of the internal structures and use techniques specific to the ERP systems, such as interfacing R/3 by means of SAP's Business API.

EAI also aims at integrating ERP systems across enterprises. Thus, the integration is essentially message-oriented, and XML is used as the standard for encoding messages. To allow independent communication partners to refer to the same DTD, some companies and consortia offer central repositories of XML DTDs to which communicating ERP systems can refer. Examples of those repositories are BizTalk.org, mySAP.com, and XML.org. These centrally managed DTDs then become domain-specific communication standards. For shared care, the DTDs that address the health domain are the relevant standards.

The techniques for EAI can be related to the three layers of the HISA. The common middleware layer of HISA is, to a great extent, replaced in EAI by a message-based communication structure. Thus, many of the problems associated with messaging discussed earlier are not really solved with EAI. However, in order to deploy the HISA middleware architecture to achieve inter-organization SI, all participating organizations must use this specific middleware system, which could be viewed as requiring somewhat tight integration. Agreeing on common standards for general messages may be easier in such a setting, but is likely to be problematic when records or record fragments must be shared between organizations. Therefore, a loose coupling with EAI techniques may be the preferred approach for SI across organizations, and a tight coupling with the HISA middleware may be a good choice for SI within organizations.

In addition, EAI also supports the integration of the user interfaces from dissimilar applications into a graphical user interface on the desktop. Traditional screen scraping tools allow parallel display of several application interfaces on the same computer screen, but does not achieve real integration. Modern Web techniques allow for more customized integration within a single graphical user interface—the Web browser. The mySAP Workplace, for instance, aims at providing a single, Web-enabled window to different ERP systems and other applications. Role-based customization facilitates the adaption of the graphical interfaces to a specific user's needs.

Security

Guaranteeing the integrity, confidentiality and security of sensitive patient data is essential if patients and clinicians are to have confidence in the application of IT in health care as a whole, as well as the implementation of EHCRs in particular. Technically it is possible to ensure the security of data in an open distributed system environment as can be seen from their widespread use in the financial sector. However, there is no doubt the health environment is more heterogeneous and complex. A number of security concerns must be addressed including confidentiality (identification, authentication and authorization), integrity (authorized modification of information) and availability (access at the right time) [2]. If an error is detected in a data item, it must be corrected (not by overwriting the erroneous value, but simply indicating that it has been superseded) and just as importantly ensuring that everyone that has accessed the data item is informed of the correction. This implies that it is necessary as a minimum to maintain a full disclosure log, as well as the conventional audit trail. The health care environment is characterized by the open nature of access to clinics and hospitals, which leaves them particularly exposed. Moreover, especially in the case of large university hospitals, there is a large turnover of staff as cohorts of students, doctors and nurses undertake rotations in each of the different departments and units. User identification, authentication and authorization become burdensome in such situations. The consequence of unauthorized disclosure of information may seriously affect a patient's health, social standing, and employment prospects. However, one of the main problems is the fact that the present culture and climate throughout the health sector is not conducive to the implementation of appropriate security measures. The major challenge is not so much technical as cultural. For example, in many countries in Europe the pace is being set by the introduction of Data Protection Legislation and by Freedom of Information Legislation.

Conclusions

That advances have been made in providing wellfunctioning systems are not in doubt. An editorial in a special issue on legacy systems [10] asks the question "are these systems legacies to be discarded and replaced with the best of breed vendor approach?" For the hospitals directly involved, the response to the question is perhaps clear, but for the rest of the market the question is still valid, and the answer cannot be straightforward. The transferability of solutions-well established after many years of effort at one site-to another site is still problematic. The cost of transferring solutions will remain high unless benefits of scale are realized and this means choosing a limited number of general approaches supported by standards. It could be imagined that market forces would eventually provide the answer to the question, but the real challenge for the health care informatics community is to attempt to short-circuit this prolonged process and make rational decisions based on technology assessment and achieve more effective spending of budgets [1]. Increased globalization is likely and will come about more readily, if a better attitude to standardization is adopted. The non-health care market will dictate the preferred engineering and technology options, due to the relative sizes of the market, but the information aspect is entirely within the control of the health care community. The catalyst for progress must surely be the shareability of patient records and hence the efforts

worldwide on devising record architectures. Progress has been slow and earlier optimism for true interoperability on a wide scale was misplaced. Nevertheless there are valid reasons for the doubling of effort on standardization and ISO TC 215 is likely to be crucial to bringing order to an area that demands international co-operation. Finally, it should be mentioned that Web technology is having and can be expected to have an ever-growing impact on the delivery of information across all domains. While not entirely relevant to this article it is noted that existing health care systems are being wrapped and Web-enabled. In the short term this has the advantage of giving clinicians the opportunity to assess the clinical impact of having relatively easy access to patient data. This in turn will provide impetus toward more scaleable and generic approaches to full information sharing and SI.

WILLIAM GRIMSON (william.grimson@dit.ie) is assistant head of the School of Control Systems and Electrical Engineering at the Dublin Institute of Technology, Dublin, Ireland.

WILHELM HASSELBRING (hasselbring@acm.org) is an assistant professor at the Infolab in the Department of Information Management and Computer Science, Tilburg University, the Netherlands.

References

- 1. Bakker, A., Leguit, F. Evolution of an integrated HIS in the Netherlands. *Int. J. Med. Info. 54*, 3 (Mar. 1999).
- 2. Barber, B., Treacher, A., and Louwerse, K., Eds. Towards security in medical telematics: Legal and technical Aspects. In *Technology and Informatics Series No. 27*, IOS Press, 1996.
- 3. Ferrara, F., Grimson, W., and Sottile, P.A., The holistic architectural approach to integrating the healthcare record in the overall information system. In *Proceedings MIE '99*, IOS Press, pp. 847–852.
- Grimson, J. et al. A CORBA-based integration of distributed electronic healthcare records using the synapses approach. BioMedicine 2, 3 (Sept. 1998), 124–138.
- 5. Hasselbring, W. Federated integration of replicated information within hospitals. *Int. J. Dig. Lib. 1*, 3 (Nov. 1997), 192–208.
- Ingram D., The Good European Health Record, In Health in the new communication age, MF Laires, MF Ladeira and JP Christensen (Eds), IOS, 1995, 66-74.
- 7. Jung, B., Grimson, J., and Grimson, W., The EHCR-If not now, when? In Proceedings TEHRE '99, pp. 391-397.
- Kohane I.S., Greenspun, P., Fackler J., Cimino, C., and Szolovits, P. Building national electronic medical record systems via the World Wide Web. JAMIA 3, 3 (1996), 191–207.
- 9. Peppard J. and Rowland P. *The Essence of Business Process Re-engineering*. Prentice Hall, 1995.
- 10. Safran, C., Ed. Special issue on legacy systems. Int. J. Med. Info. 54, 3 (Jun. 1999).
- Thorsen T. and Makela M. Professional Practice: Theory and Practice of Clinical Guidelines Implementation. Copenhagen, DSI, 1999.
- 12. Van Bemmel, J.H. and Musen, M.A., Eds. Handbook of Medical Informatics. Springer, 1997.

© 2000 ACM 0002-0782/00/0600 \$5.00

JANE GRIMSON (jane.grimson@tcd.ie) is an associate professor of Computer Science and Chair of the Center for Health Informatics at Trinity College, Dublin, Ireland.