The Practice of Informatics

Application of Technology ■

Building National Electronic Medical Record Systems via the World Wide Web

ISAAC S. KOHANE, MD, PhD, PHILIP GREENSPUN, MS, JAMES FACKLER, MD, CHRISTOPHER CIMINO, MD, PETER SZOLOVITS, PhD

Abstract Electronic medical record systems (EMRSs) currently do not lend themselves easily to cross-institutional clinical care and research. Unique system designs coupled with a lack of standards have led to this difficulty. The authors have designed a preliminary EMRS architecture (W3-EMRS) that exploits the multiplatform, multiprotocol, client–server technology of the World Wide Web. The architecture abstracts the clinical information model and the visual presentation away from the underlying EMRS. As a result, computation upon data elements of the EMRS and their presentation are no longer tied to the underlying EMRS structures. The architecture is intended to enable implementation of programs that provide uniform access to multiple, heterogeneous legacy EMRSs. The authors have implemented an initial prototype of W3-EMRS that accesses the database of the Boston Children's Hospital Clinician's Workstation.

■ **JAMIA.** 1996;3:191–207.

General interest in deployment of electronic medical record systems (EMRSs) is increasing rapidly. Currently, there are more than 230 vendors of such systems serving acute care hospitals alone.[1] Studies suggest that these EMRSs may be of central importance in streamlining medical practice, reducing costs,[2] and improving quality of care.[3] Many commercially developed and widely deployed systems fall short of current needs, and show dangerous signs of losing additional ground to future demands. Many, though not all, of these systems are based on old architectural models, idiosyncratic local organization and coding styles, old implementation technologies, accretion of code and functions, and a closed-world style that makes their interoperation with other systems and the evolution of new functions very difficult to achieve.

The authors' preliminary research results are reported here, detailing a basis for constructing new, generic, extensible architectural frameworks for EMRS. Foremost among recent technologic changes enabling such work is the revolutionary growth of the so-called "Global Information Infrastructure," and in particular the World Wide Web (W3). The Web is rapidly incorporating institutions and individual users into a linked matrix of informational resources. A second enabling feature is the growing realization in most health care organizations that interoperation, and

Affiliations of the authors: Children's Hospital and Harvard Medical School, Boston, MA (ISK, JF); Laboratory for Computer Science, Massachusetts Institute of Technology, Boston, MA (PG, PS); and Albert Einstein College of Medicine, Boston, MA (CC).

Supported by the National Library of Medicine (U01 LM05877-01) and in part by the Oracle Corporation and the Charles Hood Foundation.

Correspondence and reprints: Isaac S. Kohane, MD, PhD, Children's Hospital Informatics Program, Children's Hospital, 300 Longwood Avenue, Boston, MA 02115. e-mail: kohane@a1.tch.harvard.edu

All identification data in the figures are fictitious.

Received for publication: 8/25/95; accepted for publication: 1/17/96.

therefore the adoption of new conventions and practices, is no longer optional. Third are the gradual improvement and increasing sophistication of standards and shared structures that support communication, comparability of data, etc. The momentum toward building new systems also encourages and enables institutions to adopt "state-of-the-practice" techniques. Institutions recognize that they must try to define capabilities in terms of architectures rather than particular programs, to lengthen the useful life of new designs.

Background

A number of developments motivate and make feasible the proposed architecture and its initial implementation.

Previous Efforts in Developing EMRSs

A number of pioneering institutions have implemented extensive EMRSs, among them the system at Beth Israel Hospital in Boston[4]; that at Massachusetts General Hospital[5]; the HELP System at LDS Hospital[6]; the Regenstrief system at the University of Indiana[7]; the Columbia-Presbyterian Medical Center system[8]; that at Boston Children's Hospital[9]; that at the Brigham and Women's Hospital[10]; the Duke Hospital Information System[11] and TMR[12]; the MARS project at University of Pittsburgh Medical Center[13]; and the Department of Veterans Affairs System (DHCP).[14] Their results suggest that many health care institutions eventually will be able to implement EMRSs that provide adequate breadth of clinically relevant machine-readable data. However, shortcomings discovered through work on existing systems include:

- Many existing systems record clinical information either as narrative text or, if coded at all, using locally developed nomenclatures that make comparison and sharing of data with others very difficult.[15]
- In very few places is it possible to retrieve a patient's history across the multiple EMRSs of the institutions at which the patient has received care.[16]
- Health services research across multiple institutions is difficult and expensive, even when limited to basic data such as diagnoses and costs, let alone detailed clinical histories.[17]
- Sharing data between disparate EMRSs requires building custom translators. For n EMRSs, order n^2 translators have to be produced, which is likely to be prohibitively expensive for national online data sharing.
- Growth in EMRS functionality is slow. Synergy among geographically diverse members of the informatics community is poor, in part because functions developed on one system are not readily transferable to another.[18,19]

There is no fundamental limitation that prevents existing systems from being reimplemented to overcome these problems, and many of the developers of the systems cited above are making substantial progress. However, the enormous costs of reimplementation, along with the tradition of parochial designs and the difficulty of adopting universal standards and conventions, generally make such reimplementation efforts unattractive to institutions that do not have large, in-house medical informatics groups.

Coding Standards for Medical Data

Present-day hospitals and other health care institutions currently vary widely in the kind and amount of clinically relevant data that are stored electronically. Information needed for billing is universally stored within institutional systems. Most large clinical facilities have computerized their clinical laboratories and pharmacies. Many sites capture electronic versions of some reports from diagnostic services such as radiology, but so far relatively few capture records of the history and physical, doctors' notes, nursing reports, bedside impressions, etc. The absence of a sufficient breadth of machine-readable data in such EMRSs makes it impossible to serve the full range of functionality required in an ideal computer-based medical record.[20]

The simplest method for storing patient information is to allow users to enter descriptions in narrative text. While use of uncontrolled text simplifies the entry of information, it complicates retrieval. Attempts have been made to develop algorithms that can retrieve uncontrolled text with accuracy that approaches that for retrieving material indexed with a controlled vocabulary.[21,22] However, in the patient care domain, the highest accuracy possible is needed. Currently, this can be ensured only by controlling what terms are used in the content of the record. Thus, most institutions face the dilemma of how to record data in a standardized, coded manner that does not encumber the busy clinical care provider. Few institutions have developed locally successful approaches, and none has developed a generalized solution to the problem.

Numerous controlled vocabularies have been developed in the health care domain (MeSH,[23] ICD9-CM,[24] SNOMED III,[25] Read,[26] Gabrielli,[27] NANDA,[28] each for a specific purpose. None of these vocabularies has been developed to encompass all the purposes of an EMRS.[20] For example, no existing controlled vocabulary is considered satisfactory for the coding of clinical problem lists.[29] To help bridge differences among various "standard" structured vocabularies, the National Library of Medicine's Unified Medical Language Systems (UMLS) project is building a Metathesaurus that identifies synonymy and similarity relationships among concepts in different source vocabularies.[30,31] The UMLS also contains information on concept term sources, definitions, antonyms, and other relations among the vocabularies. For medical resources using terms that are not included in the Metathesaurus, the UMLS nevertheless includes a broad enough sample of terms in many domains so that a majority of new terms can be expected to match to existing terms.[29,32-36] Structuring the knowledge of medicine requires more than synonym relationships among different expressions of concepts. The UMLS project is also developing a Semantic Network[31,37,38] that indicates relationships such as "is part of" and "is affected by." Such relationships augment the terminologic data by providing information about how different terms may relate to each other.[39-46]

Database-independent Transactions

Patient records are typically spread across multiple databases. The task of accessing information in a variety of databases is complicated by the fact that each database may be implemented in a distinct technology (e.g., flat file, hierarchic, network, relational or object-oriented), may come from a distinct vendor, and may use completely distinct interfaces and query languages. Various approaches have been taken to overcome this problem.[47-51] The most obvious is to define a set of communication protocols that identifies the format and content of data as they are being interchanged. The internal structure of any database is, then, isolated behind the translation functions that create the messages to be communicated from the database and that decode such messages and store their content or respond to the requests they contain. This is just the approach taken by HL7,[52] which is the most widespread set of conventions for communicating medical data. The principal challenge of this approach is that the language of messages must be designed to accommodate all useful data that may need to be exchanged. The principal disadvantage is that translation software must be written for each database system.

Other, more sophisticated schemes for client–server access to heterogeneous databases rely on having a model of the content and native interface of each database and synthesize appropriate database commands as needed.[53,54] One of these efforts is SIMS,[55] a knowledge-based system that supports a semantic model of the problem domain and uses this model to reformulate uniform queries (specified in the LOOM language[56]) as database-specific queries. In SIMS, associated databases are treated as information servers, whereas the semantic model and the query reformulation methods form the client. Other efforts have been recently announced by other researchers.[47,57] In the MARS[13] project, data from incompatible systems are integrated through configurable parsers that load heterogeneous textual records into one database where they are fully indexed for subsequent searches and study. In the MARS project, some knowledge of the information model of each database is embedded in their respective parser. Reddy et al.[48] generated a survey of earlier efforts in managing heterogeneous databases.

The World Wide Web

The W3 is a collection of Internet protocols that support easy access to a huge variety of information.[58] The use of W3 mechanisms provides a rich toolkit for useful functions that can support EMRSs.[59] HyperText Transfer Protocol (HTTP) supports a universal naming scheme for information on all computers accessible through the Internet.[60] HyperText Markup Language (HTML) is a relatively simple markup language that allows formatted and multimedia documents to be displayed in a way that is device-independent (i.e., the document appears approximately the same on a Macintosh, under Windows, or on a UNIX workstation) and that supports embedding hyperlinks that connect to other documents. Use of such mechanisms in EMRS development addresses and averts, in part, the otherwise high cost of developing special-purpose systems for medical use "from scratch."

Research laboratories and commercial vendors have rushed to embrace W3 mechanisms and now provide free or inexpensive servers that can transmit stored documents or create them on the fly in response to specific requests.[61] The same vendors have also distributed free or low-cost clients that allow even relatively computer-naive users to explore "cyberspace" by navigating through what appears to be a single hypertext document that encompasses all information sources on W3. Although these capabilities are quite new and still undergoing rapid development, W3 protocols already support formatted text, sound, still images, and digital video. The use of client applications such as Mosaic and Netscape is growing at astonishing rates, and computer vendors are now selling introductory level personal computers with these clients preinstalled to attract the novice computer buyer by offering "the world of information at the user's fingertips."

Almost monthly, new features and capabilities are adjoined to the basic functions of W3 servers and clients. During just the past year, we have seen added features for authenticating clients and servers, two competing mechanisms for encrypting communication between client and server, nearly real-time updating of information through an open channel from server to client, increased sophistication of possible layouts, and the ability to run "applets" locally at the client to support interfaces that are more dynamic and sophisticated.[62] Additional capabilities already on the horizon include improved means for naming resources, better caching schemes to improve network efficiency, additional capabilities to support cooperative work, and (probably several) micropayment schemes that allow efficient billing of very small amounts to cover the incremental costs of services.[63,64] Longer-term work is also under way, among other things, to support better content-based discovery of relevant material in the vast reaches of the Web and to enable declarative descriptions of W3 information content so that it can be manipulated by program, not simply retrieved by human users.[65] The simplicity of the basic means of access—click and follow links—and the universality of access—through nearly free server and client software—have made the Web very popular, and this popularity has formed a positive feedback loop, where almost every major software developer is working hard to create even more attractive capabilities that will attract even more users.

User interface design is, of course, a broader topic than simply what can be accomplished using the mechanisms of W3.[66] The design of the user interface for the EMRS is critical for clinician acceptance and use, and part of the reluctance of health care providers to adopt EMRSs, particularly for data entry, stems from the awkwardness of the user interfaces available and their heterogeneity across different applications, even within a single institution.[67] Ethnographic studies have been used to identify user needs, work patterns, and environmental context when designing systems,[68] and the need for better understanding of human–computer interactions is evident in the many recent studies in this.[69–72] Nonetheless, the use of such methodologies has not been widespread among EMRS vendors. Even if these ethnographic studies were performed, their results would be difficult and expensive to apply to the wide variety of nonstandard EMRS databases, functions, and user interfaces. If W3 can serve to make interfaces more standard, then the results of ethnographic studies (and other user interface research) can be applied more broadly.

Methods

An EMRS Architecture Based on Common Medical Records and W3 Mechanisms

As part of the National Library of Medicine's EMR Collaborative effort that began in 1994, the authors report preliminary work on an architecture, W3-EMRS, that responds to the above-identified needs. It is based on four major design components:

1. Definition of a Common Medical Record (CMR) that represents an evolving consensus of what information should be present in an EMRS, and in what form. The CMR provides a standard abstract set of database structures and transactions that are independent of the particular structures and transactions of the local database. This will allow construction of programs that do not have to become mired in the idiosyncrasies of each local EMRS and will be able, nonetheless, to access and, if authorized, to modify the contents of the local EMRS. The authors view the CMR as an empirical and evolving design object rather than as a strictly defined standard, though experience and widespread adoption could lead to its standardization.

2. Conventional mechanisms to access data from existing clinical data repositories and convert the data to match CMR specifications and formats.

3. The conventions and technology of the W3. W3 technology addresses a number of needs for an ideal EMRS, including universal, multiplatform availability; standardized communication protocols; available security mechanisms; simple formatting and multimedia presentation capabilities; and at least minimal interoperability with other information services on the Web.

4. A presentation abstraction layer that includes generic methods for presenting data that serve as bridges between the CMR and the formatting capabilities of Web browsers.

While the authors intend the proposed architecture to meet the design criteria listed above, and to support widespread interoperability of EMRS eventually, the work reported here focuses on a more limited goal: to show that a prototype can be constructed that captures clinical data and reports results in a manner consistent with the proposed architecture. While one of the most difficult tasks in realizing the planned architecture is step 2 above (retrieving data from existing "legacy" repositories and converting it to the CMR standard), this component is accomplished in the preliminary model through the use of a single clinical database, the clinical data repository used at Boston's Children's Hospital. This paper describes how components of the ideal architecture have been initially incorporated into a prototypic system for the Children's Hospital informational setting.

The W3-EMRS implementation parallels its architecture. Four components, the CMR, the CMR–Legacy System Gateway, the Screen Management Layer, and the Visual Presentation Abstraction. As illustrated in Figure 1, several processes (drawn as spheres and cylinders spanning one or more abstractions) mediate transformations of data and user actions between these layers. The architecture deliberately leaves open many implementation choices, which must be made for any particular EMRS.

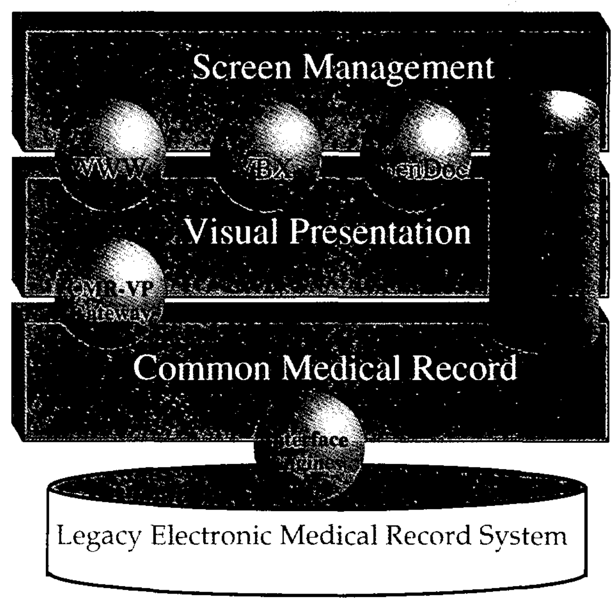


Figure 1 The W3-EMRS architecture. The three dark gray boxes represent the three abstraction layers of W3-EMRS: The Screen Management abstraction, the Visual Presentation (VP) abstraction, and the Common Medical Record (CMR) abstraction. At the bottom is the legacy EMRS. The spheres and cylinders represent processes that translate between these various layers. At the bottom, there are the interface engines that implement the CMR–Legacy gateway. The CMR–VP gateway translates data from the CMR into the Visual Presentation layer and translates Visual Presentation actions into queries against the CMR. WHAM! is an authoring tool that directly links HyperText Markup Language (HTML) forms to data elements in the CMR. The three remaining spheres represent translations of visual elements from the Visual Presentation abstraction into a specific user interface technology: World Wide Web (WWW), Visual Basic (VBX), and one of the emerging document component technologies (OpenDoc).

The Common Medical Record

If W3-EMRS viewers and editors are to be truly independent of site and local database design, then a standardized representation of the information model contained in every legacy EMRS is required. This representation is provided by the CMR. The CMR defines what concepts and relationships will be expressible for users of the W3-EMRS.

The authors began by defining a CMR based on the clinical data repository in place at Children's Hospital. Currently, the CMR information model includes patients, providers, visits, problems, laboratory data (subcategorized to fine detail), clinical measurements, medications, immunizations, notes, and letters, and some image data are about to be added. The definition of the CMR continues to be refined iteratively as the authors study the internal structures of other hospital information systems, plan experiments to integrate the presentation of data from multiple institutions, and adopt those parts of currently accepted standards that have developed a consensus following among developers of EMRSs (e.g., portions of the HL7 standard[52]).

The CMR also defines a set of terminologies in which all corresponding terms from the legacy EMRS may be expressed. For example, in the Children's Hospital Clinician's Workstation, each clinic has its own terminology for its problem list. If data across the entire CMR are to be coordinated, it is important to define translations among such terminologies. Architecturally, the system provides for such a translation mechanism as part of the CMR–legacy gateway, described below. The authors are not committed to any particular translation mechanism, however. For example, translations could be based on UMLS relationships, other semantic matching methods, or natural language processing techniques.[35,41,73–84]

Figure 2 shows the problem list page from the original Clinician's Workstation (CWS) interface. Figure 3 shows the combined patient identification and problem list page of W3-EMRS. Of note, the W3-EMRS system enables the user, with a single mouse click, to retrieve articles from a MEDLINE database or the Online Mendelian Inheritance in Man (OMIM)[85] database related to specified combinations of the patient's problems (Discussed below).

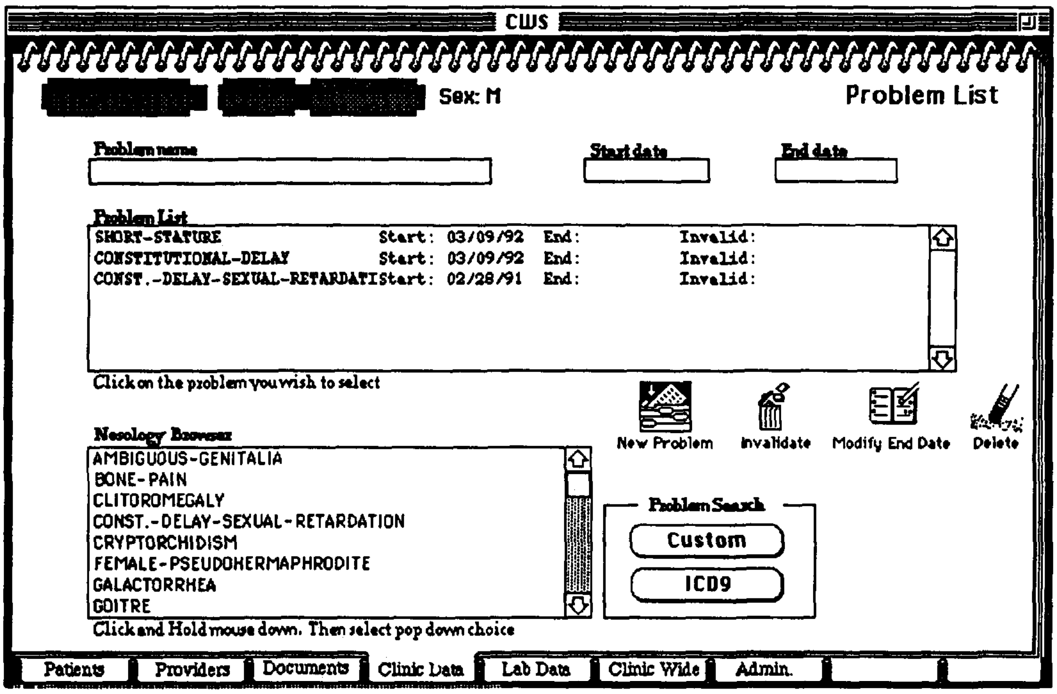


Figure 2 Problem list in original Clinician's Workstation (CWS) interface. Shown is one of the screens of the original CWS interface. Running along the bottom of the figure are tabs that represent different clinical rubrics. This screen, the problem list selection and editing screen, is under the Clinical Data rubric. It allows the user to select problems by navigating a problem nosology in the lower left scrolling field. The user can also select one of the buttons on the lower right to perform a substring search on terms from the ICD-9 vocabulary or from a vocabulary customized for each clinic using the CWS. The top scrolling field shows the problems assigned to the selected patient. Below this scrolling field are icons of buttons that allow modifications to this problem list.

The CMR–Legacy System and CMR–Visual Presentation Gateways

The CMR–Legacy gateway responds to query requests issued from the CMR–Visual Presentation gateway, issues the query to the legacy database and then returns the data in a format intelligible to the CMR–Visual Presentation Gateway. The authors have experimented using Structured Query Language (SQL) and HL7 as alternatives for the querying-and-response formats. If SQL is used, then the CMR–Legacy gateway has to translate SQL queries that refer to the CMR information model into queries in the local database manipulation language, which in the case of the CWS is also SQL. If HL7 is used, then the CMR–Legacy gateway performs the same translation process after parsing the query in the HL7 message. In both cases, the CMR–Legacy gateway returns a tagged stream of data, which is sent to the CMR–Visual Presentation gateway. If HL7 is the specified response format, then the tagged data stream is in HL7 format.

The gateway between the Visual Presentation abstraction and the CMR, the CMR–VP gateway, evaluates messages attached to predefined manipulations of the data elements displayed to the user. For instance, a time-ordered flowsheet in the Visual Presentation abstraction might attach a `showDetail` message to the user *selection* action for each element of the flowsheet. When the user selects the data element (e.g., the value of a laboratory result), the `showDetail` message is sent to the CMR–VP gateway, which then issues a CMR query that is sent to the CMR–Legacy gateway. If, as the authors believe will be typical, the CMR–Legacy gateway is implemented at the legacy EMRS site, then on receipt of the CMR query, the CMR–Legacy gateway generates the appropriate query in the legacy database's data manipulation language (DML) to obtain the detail (e.g., the reference range on a laboratory result or annotations by the laboratory technician).

All functions of the VP–CMR gateway and the CMR–Legacy gateway are implemented in the Oraperl[86,87] scripting language. Oraperl is based on the Perl scripting language, which is a popular, flexible language with good built-in operators to manipulate text and simple extensions that make other features of the computing environment, such as most UNIX operating system functions, conveniently available. Oraperl adds a set of functions that permit connection to and manipulation of an Oracle database. Although other scripting languages offer at least equivalent capability on other hardware platforms (e.g., Applescript on the Macintosh[88] and Telescript, being developed by General Magic[89]), both the existing W3 (HTTP) server and the CWS's database run under UNIX, so use of Perl was convenient. When Perl scripts were inadequate or slow (e.g., for parsing HL7), project members used compiled C code.

The Screen Management Layer

A very large number of competing technologies implement programmable user interfaces on one or more hardware platforms.[90] Congruent with the goal of leveraging existing technologies, W3-EMRS is designed to work with any of these user interfaces so long as they support the visual presentation abstraction. However, the choice of user interface technology will be driven by the particular task application for which W3-EMRS is used. For project purposes, the ability to provide the user interface on multiple hardware platforms, to provide access across the Internet, and to

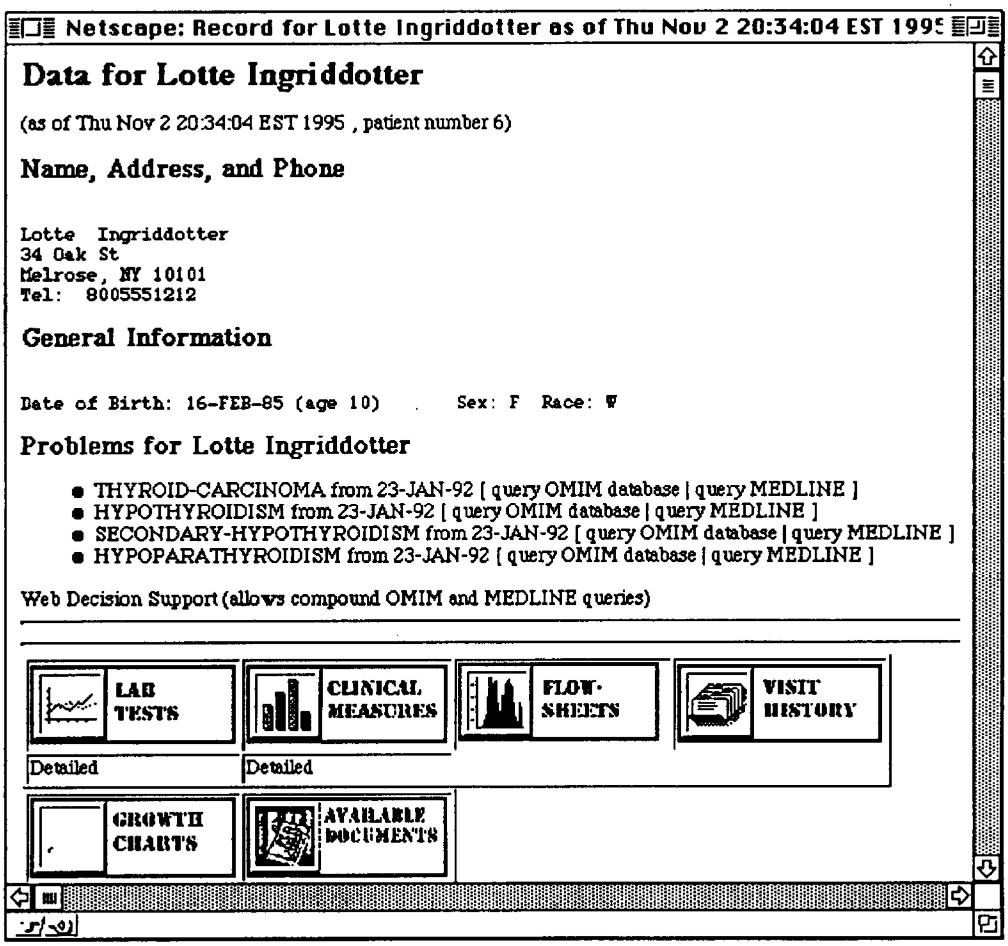


Figure 3 Patient identification/"face sheet" and problem list in W3-EMRS. The top of the page contains patient demographics followed by the problem list. Selecting a problem name generates a list of all patients in the legacy electronic medical record system (EMRS) who share the same problem and lists associated problems. Selection of the Online Mendelian Inheritance in Man (OMIM) or MEDLINE link to the right of each problem triggers a search of the corresponding database using that problem as a search key. The six icons at the bottom of the page lead to the following functions 1) Lab tests: Simple view of laboratory results against time; 2) Clinical measures: measurements (e.g., height, weight, heart rate, or blood pressure made at the bedside or in the clinic; 3) Flowsheets: spreadsheets of results and clinical measurements grouped by topical relevance; 4) Visit history: History of patient visit with provider information and billing and procedure summaries; 5) Growth charts: Tables of patient heights and weights with sex-specific calculations of Z-scores of height and weight for age (each table has a corresponding graphic plot); and 6) Available documents: Narrative (noncoded) text notes from clinicians in the clinics and in ancillary departments.

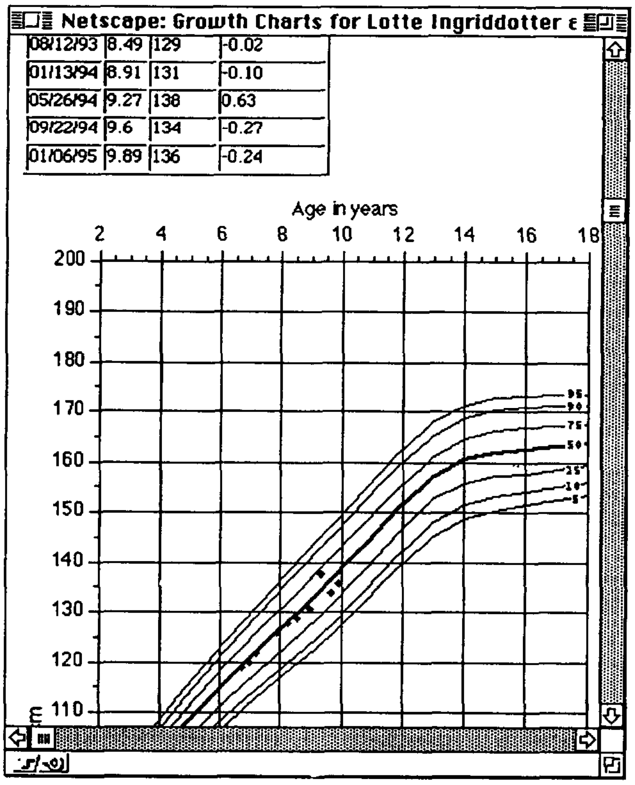


Figure 4 Tabulation and graphing of bedside measurements in W3-EMRS. Shown is the tabulation of date, age, height, and height Z-score and a corresponding plot of the heights below the table. The plot displays the standard growth centiles for North American girls. Not shown in this figure, but included in W3-EMRS, are the weight tables and weight plots.

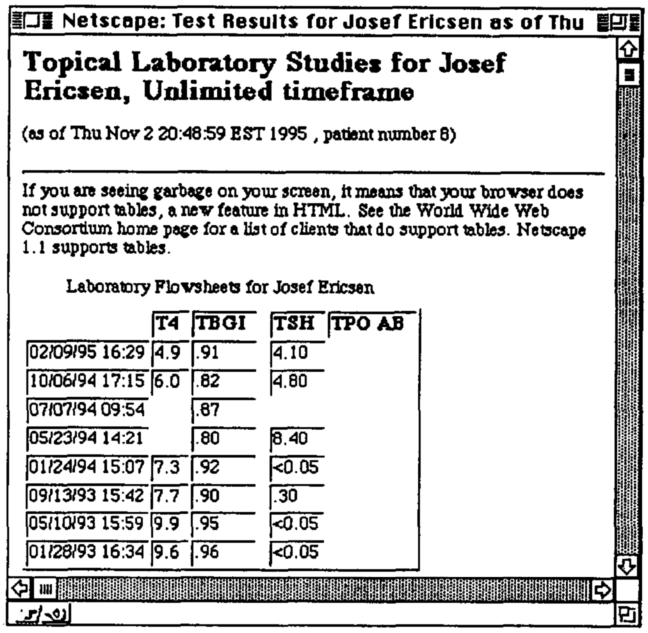
provide state-of-the-art cryptographic protection were paramount criteria. Therefore, the popularity of the W3 protocols and the low-cost, ubiquitous W3 browsers made choice of screen management technology fairly straightforward.

The Visual Presentation Layer

Clinical data displays deliver information that describes some part of the patient's status, present or past. They also encode the functions that support user interactions with the delivered information, for example, responding to the selection of the date of a clinical visit by displaying the full clinical note for that visit. The Visual Presentation abstraction describes both functions. It includes a representation of the visual layout of clinical data elements and the interactions with the CMR that the user can access through these visual layouts. Layouts include archetypal clinical presentations such as time-ordered flowsheets, graphs of related sets of time-varying data, annotated images, multipart narrative text documents, and value-restricted or coded fields. Supported user actions include selections, modifications, or deletions of presentation objects and insertions into these objects.

The principal motivation in creating the visual presentation abstraction is to separate explicitly the design of visual layout and allowed user responses from the programs that implement the user interfaces on the machines of the users of W3-EMRS. That is, for each element that is visible to the user, there may be one or more attached tags that describe messages that should be evaluated upon completion of defined user actions. As a result, the W3-EMRS architecture can accommodate a variety of client–user interface implementations, including W3 client programs, Visual Basic programs, or OpenDoc parts. Detailed appearance of a VP abstraction (e.g., a flowsheet) can then match the visual style of the implementation and interact closely with its native capabilities. User actions may also be performed by different "gestures" in different systems; e.g., selection can be a mouse click in one, typing the initial letter of its target's name in another, or speaking it in a third.

Figure 4 demonstrates that the Visual Presentation layer can direct the generation of graphics such as


Netscape: Test Results for Josef Ericsen as of Thu

Topical Laboratory Studies for Josef Ericsen, Unlimited timeframe

(as of Thu Nov 2 20:48:59 EST 1995 , patient number 8)

If you are seeing garbage on your screen, it means that your browser does not support tables, a new feature in HTML. See the World Wide Web Consortium home page for a list of clients that do support tables. Netscape 1.1 supports tables.

Laboratory Flowsheets for Josef Ericsen

	T4	TBGI	TSH	TPO AB
02/09/95 16:29	4.9	.91	4.10	
10/06/94 17:15	6.0	.82	4.80	
07/07/94 09:54		.87		
05/23/94 14:21		.80	8.40	
01/24/94 15:07	7.3	.92	<0.05	
09/13/93 15:42	7.7	.90	.30	
05/10/93 15:59	9.9	.95	<0.05	
01/28/93 16:34	9.6	.96	<0.05	

Figure 5 Topical flowsheet of laboratory data. This flowsheet is for thyroid function tests commonly used in the authors' clinics. Selecting any value in the flowsheet will trigger the generation of graph of the particular analyte over time. The abbreviations used in this flowsheet are: T4 = thyroxine, TBGI = thyroxine-binding globulin index, TSH = thyroid-stimulating hormone, and TPO AB = thyroid peroxidase antibody.

data plots "on the fly." Shown is the bottom section of a table in displaying age, height, and height Z-score, followed by a plot of the heights of the patient against the NCHS height centile standards.[91] Figure 5 shows laboratory studies, collated in a topically organized flowsheet. Shown is the thyroid function flowsheet; there are now more than 20 such specialized flowsheets implemented. Figure 6 illustrates how users can edit clinical data in the database. Other CWS functions implemented not illustrated here include: user customization of flowsheets, review of full narrative text documents, and review of visit history with annotations of diagnoses, procedures, and attending physicians for each visit.

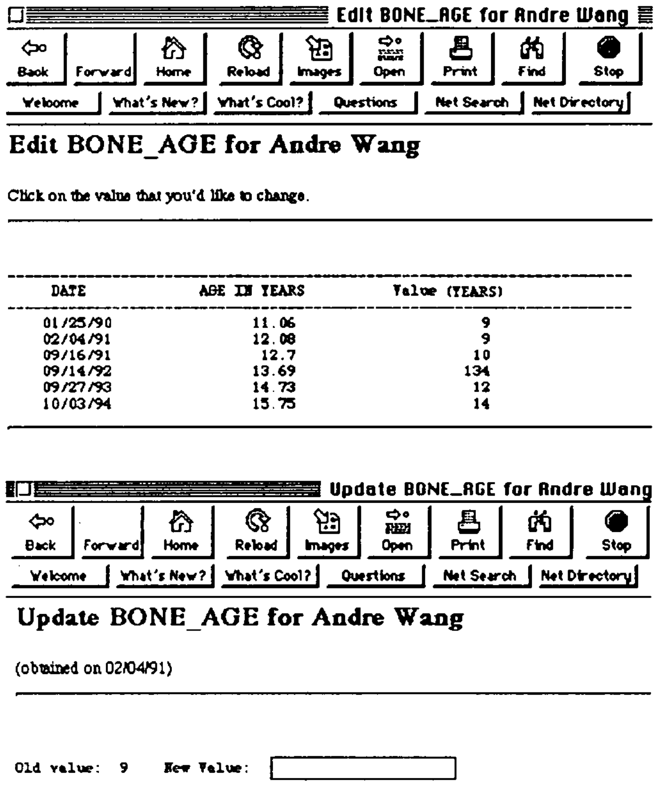


Figure 6 Editing clinical data measurements in W3-EMRS. Shown above are two W3-EMRS pages that enable user editing of data in the underlying database. The topmost page shows a list of bone age estimates (including several obvious errors). Selection of one of the items in the list brings the user to the second page, where he or she has the option of entering a new value, and thereby updating, the selected bone age measurement.

Exceptions to Design

Not all implementations of the W3-EMRS architecture have to use all the layers described. If, for example, the goal is to accrete clinical data in uniform format in a central database, then only the CMR abstraction layer is required to provide a consistent query interface to each legacy EMRS. Also, if one wanted to give the user direct broadly customizable control of the visual presentation of clinical data, one might bypass the Visual Presentation abstraction layer and directly encode the presentation in the Screen Management programs.

Although most users of the W3-EMRS may be satisfied with the visual presentation and user interactions available to them in a particular W3-EMRS implementation, some may wish to design their own special-purpose displays. For this reason the authors implemented a program called WYSIWYG HTML Authoring for Medicine (WHAM!) (reported elsewhere[92]). WHAM! allows users to generate W3 forms using a palette of standard HTML visual components (e.g., text fields, lines and buttons) and CMR abstractions (e.g., problem lists, physical examination measurements) using solely a "drag-and-drop" interface. This enables users to construct ad hoc queries and visual presentations without having to perform any text-based programming or without knowledge of HTML, CMR details, or the legacy EMRS structures. WHAM! bypasses the Visual Presentation layer to directly link CMR transactions to HTML constructs.

Security and Confidentiality

In any EMRS, security and confidentiality are a primary concern. Not surprisingly, this is also a concern of commercial developers of W3 applications (e.g., to communicate credit card numbers securely in an online shopping application). Therefore, protocols for secure authentication and end-to-end encryption for use on the W3 have already been implemented (e.g., the Secure Sockets Layer protocol[93] and s-HTTP[94]). The advantage of end-to-end encryption is that no matter how insecure are any intermediate computers or network components, the privacy of complete messages can be guaranteed secure against any but the most determined and costly attacks. Several holes in the security protocols of various W3 clients applications have recently been uncovered, including those of Netscape browser. The authors have no doubt that in the course of the very widespread use of these browsers in many industries, more problems will be discovered. However, the authors also trust that just because of the widespread use of these capabilities, their security will improve rapidly. In contrast, encrypted transmission and secure authentication are very rare in commercial EMRS products. Passwords are often transmitted unencrypted to database servers and clinical data are transmitted unencrypted even to outlying clinics, thereby making the EMRS very vulnerable to breaches of security. Therefore, especially as the W3 security protocols undergo large-scale testing and debugging, the choice of W3 for the Screen Management layer for W3-EMRS is likely to provide more protection of privacy than is currently available in commercial EMRS.

Secure communication and authentication are, however, only part of a comprehensive security solution for any EMRS. Other equally important components such as role-specific access to segments of the data in the EMRS or auditing mechanisms are *not* addressed by the use of W3 protocols alone. For example, role-specific access to different segments of an underlying database must be defined with respect to the CMR so that the same restrictions will hold across multiple EMRSs.

Results: Pilot Implementation

In October 1994, the authors began implementation of a prototype clinical information system based on the W3-EMRS architecture. An important asset in this effort was the experience gained in developing the CWS[95,96] and the availability of the rich data set provided by the CWS against which to test W3-EMRS functionality. The CWS is an integrated set of software designed to provide clinicians with convenient access to the large volume of patient data stored in the Children's Integrated Hospital Information System (IHIS),[9] and to support additional functions such as the semiautomatic generation of letters to referring physicians to document each outpatient clinic visit. The CWS also allows its users to browse data from inpatient admissions. The CWS maintains data in addition to that of the IHIS, such as physical examination measurements, problem lists, and medications, but stores these in the same Oracle repository.

The current version of the CWS is implemented on Macintosh computers networked to the IHIS.[9] The IHIS has as its centralized data repository an Oracle database stored on several Digital Equipment Corporation VAX computers (the "VAX Cluster"). This repository receives real-time data updates from several departmental applications. The CWS retrieves and displays all pertinent administrative, financial, and clinical data residing on the VAX Cluster. These data include: demographics; visit history, with associated procedure and diagnostic codes; inpatient pharmacy orders; and inpatient laboratory studies, which are entered into the IHIS through other departmental applications (e.g., the Cerner laboratory system). Users of the CWS enter additional clinical documentation into the IHIS through the electronic forms within the CWS interface. These data include: problem lists, patient–provider relationships, bedside measurements, outpatient medications, history, past medical history, family history, review of systems, and other components of clinic notes or letters to referring physicians. Access to this information is controlled by assigning data access/modification privileges to various provider roles. The CWS serves to maintain *all* clinical data/documentation of patients seen by *all* clinicians in each participating clinic. Data displays are designed to follow the metaphor of the paper chart when possible but employ other metaphors where appropriate.

Since its first deployment in July 1991, the CWS database has accumulated the records of more than 5,800 patients (i.e., 100% of the patients seen in each implemented clinic; note, however, that there are more than one million patients in the Children's IHIS). Excluding reports generated by other departmental applications (e.g., radiology and pathology, which are accessible through the same CWS interface), 15,500 electronic visit forms were completed. In the process, 100,000 individually coded clinical measurements were automatically entered into the database, as well as 6,700 problems (using the clinics' controlled problem list vocabularies). Because the number of clinics using the CWS has grown recently (it now includes the endocrinology, nephrology, nuclear medicine, and rheumatology clinics), the authors anticipate rapid growth in these numbers in the future.

For the purpose of development of the W3-EMRS, the records pertaining to 275 patients were exported from the CWS. After replacing all identifiers of referring providers and patients ("scrubbing"), these records were imported into an Oracle server running on a Sun workstation under the SunOS UNIX operating system. The first prototype of W3-EMRS was implemented in November 1994 using a subset of this scrubbed database. Shortly thereafter, the W3-EMRS software was modified to access the full CWS database within the hospital "firewall." Even though it is protected from access from outside the hospital, the W3-EMRS software within the hospital requires electronic provider authentication, and implements role-specific access restriction to data. As described above, the CWS database contains a large amount of detailed and coded data for 5,800 patients and sparse data sets for the more than one million patients in the Children's IHIS.

In the prototypes so far, the CMR implementation has followed the planned architectural design of an abstraction between the VP–CMR gateway and the CMR–Legacy gateway. User actions that request or update information in the visual presentation layer are translated into DML statements in the underlying IHIS, and the returned results are translated into visual presentations. In the short term, the authors have taken the information model of the CWS database as a "straw man" for the CMR to test the technology. The CMR currently includes coded data items (e.g., medications and problem lists) as well as narrative text (e.g., radiology reports, clinic visit summaries).

With the exception of the initial introductory page of the W3-EMRS, none of the "pages" that are seen by users of the W3-EMRS prototype are "canned" or static files on the W3 server. They are generated dynamically upon receipt of the tagged data stream from the CMR–Legacy gateway, which then sends these data to a Perl script (part of the VP–CMR gateway) that converts the tagged data into a still abstract visual presentation object such as a flowsheet or a list. These are, in turn, translated to graphic elements that are supported by the HTML 3.0 specification (e.g., a field in a form, a button, static text, a table). Finally, these are assembled into an HTML data stream that is sent to the W3 client browser.

Many of the visual presentation objects define actions that are to take place when one of their elements is selected by the user. For example, on a general laboratory examination flowsheet, each laboratory result has an attached action that requests a new flowsheet and graph, specifically for that laboratory measurement. Such actions are encoded as hypertext links in HTML, and cause the HTTP common gateway interface (CGI) mechanism to run the appropriate Oraperl script (part of the VP–CMR gateway) for each action. That script issues a request for the appropriate data to the CMR–Legacy gateway, and the above-described process of data access and output generation recurs. Therefore, the actions defined for elements of the visual presentation abstraction determine the flow with which users will browse or update the medical record.

The ability to enter data, to correct it if incorrectly entered, and to record orders requires the ability not only to browse but also to put in new information. The project has, so far, implemented very limited versions of these capabilities, specifically for correcting data entry errors. Such data entry and correction are sanctioned by the architecture, and the current implementation uses W3 browsers' "forms" capabilities to implement it. This requires that the legacy database recognize and permit update transactions triggered by update messages from the VP–CMR gateway. In the instance of the W3-EMRS prototype within the hospital firewall, the user's login name and password are used to verify whether update privileges have been granted to that user for the specified data elements in the CWS Oracle database.

The authors have taken advantage of W3's easy ability to link to Internet-wide resources to define visual presentation elements whose actions consult well-known and clinically important information sources on the Web. The implemented system currently links presentation of the patient's problem list to MEDLINE, where relevant articles can be retrieved, and to OMIM, a vast textual compendium covering known human heritable diseases. A standard HTML "form" is constructed dynamically when a clinician selects one or more items of the problem list. This form is dispatched to the W3 medical resource server (e.g., OMIM), which then responds with a list of matching syndromes formatted in HTML. A weakness in the current linkage is that it works only to the degree that the EMRS problem list and the W3 resource share vocabularies. Vocabulary translation services from the EMRS problem list to the W3 resource vocabulary would help, but these were not implemented in the first W3-EMRS prototype. Through specially designed Oraperl scripts that implement new functions, project members have also developed potentially useful capabilities that link clinical care to past clinical experience. For example, one of the actions implemented and associated with a patient's problem is the ability to find the list of other problems that co-occur in the CWS database, how often, and which other patients have this combination. Thus, selecting a patient's problem will yield a frequency-ordered list of problems that co-occur with it, selecting one of those will reveal a list of patients who have this combination, and selecting one of those patients brings up his or her record, assuming the user has suitable access authority.

Perhaps the strongest evidence for the utility of the W3-EMRS architecture is the rapidity with which the project team has produced the first working prototypes. After the design work, within 14 person-days of effort, access to all the CWS datatypes, without any graphic elements, tabular formatting, or editing capabilities, was provided using a scrubbed subset of the CWS database. It required an additional three months to refine the user interface, add editing capabilities, and provide tables and graphs of clinical data. With this relatively short development effort, users can browse, edit, and enter data in the CWS relational database management system (RDBMS) us-

Table 1 ■

Task-specific Benchmark Times for CWS and W3-EMRS*

Task	CWS (Sec)	W3-EMRS (Sec)
Find all patients with a specific name in the IHIS†	6.0	7.0
Obtain 25 clinic notes for a specific patient	4.5	4.1
Generate a collated flowsheet showing all electrolytes and ACTH‡ levels for a patient with 1,220 laboratory results	26	14

*CWS = Clinician's Workstation; W3-EMRS = the preliminary electronic medical record system using the World Wide Web technology.
†IHIS = Integrated Hospital Information System.
‡ACTH = adrenocorticotropic hormone.

ing standard W3 clients. This in itself is a notable result. Previously, the CWS could only be accessed on a Macintosh via a small number of network protocols. It now can be accessed by W3 clients on Windows, DOS, UNIX, and Macintosh on any network supported by the Internet.

One early concern of the authors was whether the implemented mechanisms would be fast enough to support useful clinical access. Extensive measurements have not been made, but response time for access to small scrubbed databases appears to be only on the order of a few seconds, even though the prototype is running on old equipment (a Sun SparcStation 2). Unexpectedly, the implementation running on the large CWS database at Children's Hospital appears noticeably faster in generating the more complex displays than was the original CWS implementation. The CWS remains faster for simple query functions.

The authors have begun to generate benchmarks of the aggregate performance of our W3-EMRS prototype implementation, running on a Sun SparcStation 20 and using the hospital's Oracle server under typical user loads. Each benchmark (Table 1) reports a time (in seconds) averaged over 50 trials. These times aggregate the performance of the networks, various W3-EMRS gateways and translation processes, and the Oracle RDBMS. The tasks measured are also only a small subset of all that the total program does. Therefore, the only conclusion that can be drawn is that the performance does not seem to be significantly worse than that of the existing CWS system. This partially allays the initial concerns about performance.

Clinicians who have used the W3-EMRS prototype uniformly gave positive reviews of its ease of use compared with either the CWS application or vendor-specific departmental applications. These clinicians have been using the W3-EMRS prototype, within the Children's firewall, as a clinical tool in their daily practice. They include a subset (eight clinicians) of the clinicians who use the CWS and other IHIS applications in their practices. The reviews were obtained during unstructured interviews designed to help guide further development of W3-EMRS and cannot substitute for formal user-driven evaluations. The predominant comment was that the user interface is always simple and predictable, compared with other applications. The improved ergonomics may be explained by the fact that HTML defines a small, consistent, and useful set of interaction metaphors that usefully constrain what could be designed and built into the visual presentation layer. Further, this constrained design appears to fit the tasks, and it avoids the possibility of overly great complexity that is possible to build with most general-purpose client software authoring systems. Formal ethnographic studies are required to substantiate and warrant these initial results.

As mentioned before, the W3 Screen Manager layer also provides the capability for encrypted data transmissions, which is unavailable in the CWS implementation. However, the project is not currently using a commercially secure HTTP server for the Internet-wide demonstration prototype, because that uses a scrubbed database and is intended to be accessible by anyone.

Discussion

Other Architectures for Accessing Medical Information via W3

Many working prototypes have been implemented that access "legacy" EMRS via W3 technology.[97-99] In addition to a high degree of user acceptance,[100] early experience with W3-based browsers has shown promise of significant improvement in clinical performance (e.g., reduction of clinical interpretation errors[99]). Applications are also proliferating in W3 use for education[101] and automated decision support.[102] These initial applications are likely to represent only a small fraction of the biomedical applications of W3 technology.[103]

In some systems, the information retrieved from legacy databases is directly encoded as a stream of text with HTML markup tags for formatting and support of hyperlinks. This approach seems architecturally impoverished, because HTML tags encode only formatting information and do not adequately support the semantics of the CMR. In other systems where the legacy EMRS supports output of data as HL7 messages, those messages are taken to play the role of the CMR. In our design, we have concluded that the flexibility of HL7, and therefore the variability in the ways in which information can be encoded, is a liability, especially if one plans to interchange data among institutions. For example, in the Columbia-Presbyterian Medical Center (CPMC) system, an HL7 message returns information about glucose results that is coded distinctly for different procedures (e.g., Chem-7, Chem-20).[97] This information is clearly useful at the CPMC and might also be of interest across several institutions, but only if the institutions sharing the data have a consensus model of how a particular laboratory result relates to a procedure.

Most of these systems also omit the VP layer of the W3-EMRS, translating the results of database queries (whether they are returned as HL7 messages[97] or SQL response strings[99]) directly into W3's HTML. The W3-EMRS architecture's visual presentation layer defines a useful intermediate target for translation and an abstraction that provides flexibility. For example, the visual presentation layer representation of a laboratory flowsheet can be used to drive, through object linking and embedding, a presentation in a Microsoft Excel spreadsheet or an HTML 3.0 table displayed on a W3 client. In view of the limitations of W3 as an EMRS interface noted by several developers of W3 interfaces to medical information systems,[97,99,103] the flexibility to use alternate presentation mechanisms is appealing. A similar approach appears to have been taken in the CERC ARTEMIS project,[50,104] which includes an abstraction of EMRS services in the Model-based Information Directory ("MIND") and which also implements interfaces to each distinct data repository. The Common Object Request Broker Architecture[105] and the W3 protocols provide server-level and client-level access to the services of MIND. The content of the demonstration W3 site for CERC ARTEMIS suggests that an important focus of the architecture is the distribution of pictures of scanned clinical documents (e.g., the printed copy of a urinalysis).

The World Wide Web also has many limitations as a client–server architecture, not the least of which is the statelessness of the transactions, as noted by others.[100,103,104] That is, every single transaction initiates a new connection to the HTTP server without any standard way of linking consecutive accesses by a single user. This might require users to authenticate themselves repeatedly if they have to complete several transactions. A typical workaround has been to embed state information as hidden fields in the HTML document and additional arguments in CGI calls. This remains an inelegant, and unsafe, solution. As this problem pervades the use of W3 technology in many industries, the authors anticipate that potential solutions to this problem will be offered in the near future by competing W3 browser vendors or the W3 Consortium.

Lessons Learned during Preliminary Implementation

In retrospect, the development of the prototype of the W3-EMRS benefitted significantly from prior decisions made at Children's Hospital in other information systems projects. Perhaps the most significant of these was the choice of commercial, standard technologies to implement the Children's IHIS,[9] particularly for the clinical data repository. For example, because the data repository is implemented in an SQL[106] compliant RDBMS, many commercial and public tools (e.g., Oraperl[87]) could be immediately used to access and format the data for distribution via W3. Also, the prior implementation of a hospital-wide high-speed network running standard ethernet protocols signified that, once implemented, W3-EMRS was immediately available to all clinicians with connected desktop computers throughout the hospital. In contrast, the CWS problem list vocabulary is nonstandard, and therefore the effort required to translate problem list terms to MeSH (e.g., to increase the reliability of the linkage between W3-EMRS problems and MEDLINE articles) or other standard vocabularies is significant. This experience only adds to the growing list of reasons that EMRS developers might benefit from adherence to standards.[107]

The W3-EMRS prototype would appear less successful and useful if the process of clinician data entry had not been addressed, supported, and improved over the course of the last four years of operation of the CWS project. That is, successful implementation of a W3-based EMRS still requires completion of the task that McDonald et al.[7] termed "the difficult side of medical record systems," namely, data acquisition and in particular acquisition of data from clinicians. Sustainable integration of clinician data entry into daily clinical workflow involves social engineering, clear institutional mandates, and significant financial investments, in addition to any software architectural decision. This is likely to be the reason why early adopters of W3-based EMRSs[97-99,104] are also those who have made significant prior investments in the data integration and data entry process.

Future Plans for Additional Development

Iterative Definition of CMR

The current implementation of the W3-EMRS implicitly uses the CWS DBMS and data structures as the CMR. This was born of expedience rather than any deep claim of the universality of this particular CWS design. Now that the project has working prototypes of the W3-EMRS, it will be able to observe experimentally how different representations and implementations of the CMR will affect the performance and utility of W3-EMRS.

To make W3-EMRS actually support the goal of cross-institutional data access, the authors have been working with groups at the Beth Israel Hospital and the Massachusetts General Hospital in Boston to define a consensus model of what data are to be shared and in what form they are to be requested and retrieved. To drive this consensus process, the authors have chosen the following task: define the data elements that are required when a patient from one hospital appears for treatment in the emergency department of another hospital, and define the standardized vocabularies that are sufficient to represent the terms encoded at each hospital site. It is assumed that the required data are stored in a heterogeneous variety of independent hospital information systems, and that those data are to be gathered as needed rather than stored in a large central repository that would effectively build a single information system to hold the data of many institutions. Project members are working with members of the HL7 community to ensure that both requests for information and the content of information being returned can be suitably encoded in that standard, to avoid creating yet another medical data standard. Experiments are planned to determine the degree to which data from multiple institutions can be usefully aggregated, for example, to produce longitudinal data plots of laboratory values collected at several institutions and (with greater difficulty) to create a "best-guess" aggregate of what medications a patient is taking. Project members do not know yet whether the obvious potential semantic problems of such an approach will prevent such aggregation and force us to merely report the data in a common format.

In any multi-institutional setting, it is safe to assume that translation from the common communication standard used by a broad-area W3-EMRS to and from local legacy database systems will take place locally. Therefore, implementation requires a potentially custom translation program to mediate transactions between the CMR and each local EMRS. Whenever possible, we prefer to use existing commercial solutions for such a task, and we are investigating several commercial products that serve as gateways between heterogeneous DBMSs. For example, Oracle Corporation (Redwood Shores, CA) has a suite of generic gateway products. In the medical domain the Software Technologies Corporation's (Arcadia, CA) DataGate and Cerner Corporation's (Kansas City, MO) OpenEngine interface engines are engineered for such access to many commercial EMRS databases. These products are only partial solutions; they do not embody all of the structure of medical databases and therefore still require significant engineering effort for each local EMRS. Other important issues not addressed by these translation engines include: 1) varied granularity of data types in different EMRSs (e.g., whether laboratory results are stored individually or as part of textual laboratory reports); and 2) differences in meaning that will not be overcome simply by translations to a standard vocabulary (e.g., different age-specific reference ranges for laboratory results of the same type in different institutions).

The authors believe, as well, that there will be a need for widespread adoption of standard ways to annotate predictable but problematic situations in medical records. For example, when data are entered and approved in error, the legal status of the medical record makes it impossible simply to correct the error subsequently. Instead, a separate notation must be made that overrides the erroneous value but without compromising the integrity of the original record. Subsequent access to the corrected records should show the corrected values, but with an indication that those values have been corrected. Uniformity in such conventions is important to make sure that data from multiple systems can be presented compatibly.

At this point in the project, treating the CMR as an adaptable, flexible set of conventions helps to support rapid experimentation. As project participants agree on new conventions, they can quickly try them out. It is likely that this will lead to convergence on a useful standard.

Vocabulary Translation

The project is currently investigating several technologies to enable both static translations of vocabularies of local EMRSs to standard vocabularies as well as context-sensitive translations. The first goal is to determine how useful the translations obtained using existing thesauri such as the UMLS Metathesaurus can be. However, the W3-EMRS architecture is neutral with respect to the choice of any particular thesaurus and will be driven in this respect by other standardization efforts. The first target in the CMR for such translations will be the problem lists. Currently each of the parties involved in the CMR definition use a different problem list vocabulary, so the translation task will be appropriately challenging.

Performance Tuning

Project members recognize that performance of W3 services can be poor, especially at sites with high user loads. Although no performance degradation has occurred with the small number of users of the W3-EMRS prototypes, experience with other heavily used W3 servers suggests that extensive use of W3-EMRS will quickly tax the current capabilities of the W3. While the W3 continues to improve the quality of its services and particularly to increase available bandwidth, the project is considering several approaches to minimize the impact of the limitations of the current W3. These include a variety of buffering techniques and also task-specific filtering of the CMR so that only small fractions of the patient record have to be viewed at one time.

Conclusion

The authors designed an architecture for client–server access to EMRSs that takes advantage of the multiplatform and multiprotocol support of the W3. The project has also implemented working prototypes that use this architecture to access and modify the databases generated by deployed EMRSs. The principal feature of the W3-EMRS architecture is its use of multiple abstraction layers to provide independence from the information models of any particular legacy EMRS. The authors' principal goal is to work on refining these abstractions with other groups developing local EMRSs who are interested in cross-institutional applications. The project will involve the working groups and organizations currently working on EMRS standards. The authors believe that the project's original prototype implementation has been successful.

The authors thank Dr. Randolph Miller and two anonymous reviewers for invaluable and substantial editorial efforts.

References ■

1. Dorenfest S. Creating a "top 100" HIS firm: the lessons of history. Healthcare Informatics. 1994;11(6):49–72.
2. Tierney WM, Miller ME, Overhage JM, McDonald CJ. Physician inpatient order writing on microcomputer workstations. JAMA. 1993;269:379–83.
3. Ornstein SM, Garr DR, Jenkins RG, Rust PF, Arnon A. Computer-generated physician and patient reminders. J Fam Pract. 1991;32:82–90.
4. Bleich H, Beckley R, Horwitz G, et al. Clinical computing in a teaching hospital. N Engl J Med. 1985;312:756–64.
5. Barnett GO. Computer-Stored Ambulatory Record (COSTAR). Bethesda, MD: DHEW, 1976.
6. Pryor TA, Gradner RM, Clayton PD, Warner HR. The HELP Syst. J Med Syst. 1983;7:87–102.
7. McDonald CJ, Tierney WM, Overhage JM, Martin DK, Wilson GA. The Regenstrief Medical Record System: 20 years of experience in hospitals, clinics and neighborhood health centers. MD Comput. 1992;9:206–17.
8. Friedman C, Hripcsak G, Johnson SB, Cimino JJ, Clayton PD. A generalized relational schema for an integrated clinical patient database. SCAMC Proc. 1990:335–9.
9. Margulies D, McCallie DP, Elkowitz A, Ribitsky R. An integrated hospital information system at Children's Hospital. SCAMC Proc. 1990:699–703.
10. Teich JM. Advanced clinical systems in a hospital environment: the Brigham Integrated Computing System (BICS). SCAMC Proc. 1995:1024.
11. Kirby JD, Walker LP, Aaron WH, Whitesall JJ, Stead WJ. Planning and implementing a disaster recovery capability for a mainframe-based hospital information system: Duke University Medical Center's experience. SCAMC Proc. 1983:872–81.
12. Stead WW, Hammond WE. Computer-based medical records: the centerpiece of TMR. MD Comput. 1988;5(5):48–62.
13. Yount RJ, Vries JK, Council CD. The Medical Archival System: an information retrieval system based on distributed parallel processing. Information Processing and Management. 1991;27:1–11.
14. Dayhoff R, Maloney D. Exchange of Veterans Affairs medical data using national and local networks. Ann NY Acad Sci. 1992;670:50–66.
15. Zielstorff RD. Capturing and using clinical outcome data: implications for information systems design. JAMIA. 1995; 2:191–6.
16. Kahn MG. The desktop database dilemma. Acad Med. 1993;68:34–7.
17. Whitman ED, Frisse ME, Kahn MG. The impact of data sharing on data quality. SCAMC Proc. 1995:952.
18. Frawley SJ. Building a database of data sets for health services research. SCAMC Proc. 1994:377–81.
19. Sujansky W, Altman R. Towards a standard query model for sharing decision-support applications. SCAMC Proc. 1994:325–31.
20. Dick RS, Steen EB (eds). The Computer-based Patient Record: An Essential Technology for Health Care. Washington, DC: National Academy Press, 1991.
21. Salton G. Development in automatic text retrieval. Science. 1991;253:974–80.
22. Hersh W. Evaluation of Meta-1 for a concept-based approach to automated indexing and retrieval of bibliographic and full-text databases. Med Decis Making. 1991; 11(4 suppl):S120–S124.
23. McCarn DB. MEDLINE: an introduction to on-line searching. J Am Soc Inform Sci. 1980;31:181–92.
24. International Classification of Diseases, Ninth Revision, with Clinical Modifications. Washington, DC: U.S. National Center for Health Statistics, 1980.
25. Côté RA, Robboy S. Progress in medical information management: Systematized Nomenclature of Medicine (SNOMED). JAMA. 1980;243:756–62.
26. Read Codes File Structure Version 3. Overview and Technical Description. Woodgate Leicester, UK: NHS Centre for

Coding and Classification, 1994.
27. Gabrielli ER. A new electronic medical nomenclature. J Med Syst. 1989;13:355–73.
28. NANDA Nursing Diagnosis: Definitions and Classifications. Philadelphia: North American Nursing Diagnosis Association, 1992.
29. Campbell JR, Payne TH. A comparison of four schemes for codification of problem lists. SCAMC Proc. 1994:201–5.
30. Humphreys B, Lindberg D. Building the Unified Medical Language System. SCAMC Proc. 1989:475–80.
31. Lindberg D, Humphreys B. The UMLS knowledge sources: tools for building better user interfaces. SCAMC Proc. 1990:121–5.
32. Lamiell JM, Wojcik ZM, Isaacks J. Computer auditing of surgical operative reports written in English. SCAMC Proc. 1993:269–73.
33. Miller PL, Frawley SJ, Wright L, Roderer NK, Powsner SM. Lessons learned from a pilot implementation of the UMLS information sources map. JAMIA. 1995;2:102–15.
34. Rosenberg KM, Coultas DB. Acceptability of Unified Medical Language System terms as substitute for natural language general medicine clinic diagnoses. SCAMC Proc. 1994:193–7.
35. Sheretz D, Tuttle M, Blois M, Erlbaum M. Intervocabulary mapping within the UMLS: the role of lexical matching. SCAMC Proc. 1988:201–6.
36. Zielstorff R, Cimino C, Barnett G, Hassan L, Blewett D. Representation of nursing terminology in the UMLS Metathesaurus: a pilot study. SCAMC Proc. 1992:392–6.
37. McCray A. The UMLS semantic network. SCAMC Proc. 1989:503–7.
38. McCray A, Hole W. The scope and structure of the first version of the UMLS semantic network. SCAMC Proc. 1990:126–30.
39. Campbell JR, Kallenberg GA, Sherrick RC. The clinical utility of META: an analysis for hypertension. SCAMC Proc. 1992:397–401.
40. Cimino C, Barnett G, Hassan L, Blewett D, Piggins J. The interactive query workstations: standardizing access to computer-based medical resources. Comput Methods Programs Biomed. 1991;35:293–9.
41. Cimino JJ. Representation of clinical laboratory terminology in the Unified Medical Language System. SCAMC Proc. 1991:199–203.
42. Cimino JJ, Aguirre A, Johnson SB, Peng P. Generic queries for meeting clinical information needs. Bull Med Libr Assoc. 1993;81:195–206.
43. Nelson SJ, Fuller LF, Erlbaum MS, Tuttle MS, Sherertz DD, Olson NE. The semantic structure of the UMLS Metathesaurus. SCAMC Proc. 1992:649–53.
44. Peng P, Aguirre A, Johnson SB, Cimino JJ. Generating MEDLINE search strategies using a librarian knowledge-based system. SCAMC Proc. 1993:596–600.
45. Rindflesch TC, Aronson AR. Ambiguity resolution while mapping free text to the UMLS Metathesaurus. SCAMC Proc. 1994:240–4.
46. Sperzel WD, Abarbanel RM, Nelson SJ, et al. Biomedical database interconnectivity: an experiment linking MIM, GENBANK, and META-1 via MEDLINE. SCAMC Proc. 1991:190–3.
47. Sujansky W, Altman R. Bridging the representational heterogeneity of clinical databases. Working notes of the AAAI Spring Symposium on Artificial Intelligence in Medicine: Interpreting Clinical Data (Palo Alto). Menlo Park, CA: AAAI Press, 1994.
48. Reddy MP, Prasad BE, Reddy PG. Query processing in heterogeneous distributed database management systems. In: Gupta A (ed). Integration of Information Systems: Bridging Heterogeneous Databases. New York: IEEE Press, 1989; 264–77.
49. Scherrer JR, Lovis C, Borst F. Diogene 2, a distributed hospital information system with an emphasis on its medical information content. In: Bemmel JHv, McCray AT (eds). Yearbook of Medical Informatics. Stuttgart: Schattauer, 1995;86–97.
50. Reddy R, Jagannathan V, Srinivas K, et al. ARTEMIS: a collaborative framework for healthcare. SCAMC Proc. 1993:559–63.
51. Lowe HJ, Walker WK, Vries JK. Using agent-based technology to create a cost-effective integrated, multimedia view of the electronic medical record. SCAMC Proc. 1995: 441–4.
52. Health Level Seven: an application protocol for electronic data exchange in healthcare environments; version 2.2. Chicago, IL: Health Level Seven, 1990.
53. Annevelink J, Young CY. Heterogeneous database integration in a physician workstation. SCAMC Proc. 1992:368–72.
54. Sheth AP, Larson JA. Federated database systems for managing distributed, heterogenous, and autonomous databases. ACM Comput Surv. 1990;22:183–236.
55. Arens Y, Chee CY, Hsu C-N, Knoblock CA. Retrieving and integrating data from multiple information sources. Int J Intelligent Cooperative Information Syst. 1993;2(2):127–58.
56. MacGregor R. A deductive pattern matcher. St. Paul, MN: Proceedings of the annual meeting of the AAAI. 1988; 403–8.
57. Papazoglou MP. An organizational framework for cooperating intelligent information systems. Int J Intelligent Cooperative Information Syst. 1992;1(1):169–202.
58. Berners-Lee TJ, Cailliau R, Groff J-F, Pollerman B. World-Wide Web: The Information Universe. Electronic Networking: Research, Applications and Policy (vol 2). Westport, CT: Meckler Publishing, 1992;52–8.
59. Liu C, Peek J, Jones R, Buus B, Nye A. Managing internet information services: World Wide Web, Gopher, FTP and more. Sebastopol, CA: O'Reilly and Associates, 1994.
60. Berners-Lee T, Masinter L, McCahill M. Uniform Resource Locators. Internet Engineering Task Force (Reston, VA: Corporation for National Research Initiatives). Network Working Group, 1994.
61. Baran N. The greatest show on earth. BYTE. 1995;20(7): 69–87.
62. Haff Av, Shaio S, Starbuck O. Hooked on Java. Reading, MA: Addison-Wesley, 1996.
63. Kalakota R, Whinston AB. Frontiers of Electronic Commerce. Reading, MA: Addison-Wesley, 1996.
64. Hallam-Baker PM. Micro Payment Transfer Protocol (MTP), version 0.1. World Wide Web Consortium (W3C), 1995.
65. Krauskopf T, Miller J, Resnick P. Treese W. Label syntax and communication protocils. World Wide Web Consortium, 1995.
66. Marchionini G. Information seeking in electronic environments. Cambridge, England: Cambridge University Press, 1995. Cambridge Series on Human–Computer Interaction.
67. Hammond WE, Stead WW. Adopting TMR for physician/nurse use. SCAMC Proc. 1991:833–7.
68. Forsythe D. Using ethnography to build a working system: rethinking basic design assumptions. SCAMC Proc. 1992: 505–9.

69. Wilton R. User-specific interfaces for clinical data-management systems: an object-based approach. SCAMC Proc. 1992:265-9.
70. Shaw D, Czaja R. User interactions with the PDQ cancer information system. Bull Med Libr Assoc. 1992;80:29-35.
71. Green C, Gilhooly K, Logie R, Ross D. Human factors and computerization in intensive care units: a review. Int J Clin Monit Comput. 1991;8(3):167-78.
72. Beaudouin-Lafon M. An overview of human-computer interaction. Biochimie. 1993;75:321-9.
73. Payne TH, Martin DR. How useful is the UMLS Metathesaurus in developing a controlled vocabulary for an automated problem list? SCAMC Proc. 1993:705-9.
74. Cimino JJ. Data storage and knowledge representation for clinical workstations. Int J Biomed Comput. 1994;34(1-4): 185-94.
75. Burgun A, Delamarre D, Botti G, et al. Designing a subset of the UMLS knowledge base applied to a clinical domain: methods and evaluation. SCAMC Proc. 1994:968.
76. Cimino J, Hripcsak G, Johnson S, Clayton P. Designing an introspective, controlled medical vocabulary. SCAMC Proc. 1989:513-8.
77. Cimino J, Hripcsak G, Johnson S, Friedman C, Fink D, Clayton P. UMLS as knowledge-base—a rule-based expert system approach to controlled medical vocabulary management. SCAMC Proc. 1990:175-9.
78. Cimino JJ, Johnson SB, Peng P, Aguirre A. From ICD9-CM to MeSH using the UMLS: a how-to guide. SCAMC Proc. 1993:730-4.
79. Cimino JJ. Controlled medical vocabulary construction: methods from the Canon Group [editorial]. JAMIA. 1994; 1:296-7.
80. Hersh WR, Hickam DH, Haynes RB, McKibbon KA. A performance and failure analysis of SAPHIRE with a MEDLINE test collection. JAMIA. 1994;1:51-60.
81. Huff SM, Rocha RA, Bray BE, Warner HR, Haug PJ. An event model of medical information representation. JAMIA. 1995;2:116-34.
82. Rocha RA, Huff SM, Haug PJ, Warner HR. Designing a controlled medical vocabulary server: the VOSER project. Comput Biomed Res. 1994;27:472-507.
83. Sager N, Lyman M, Bucknall C, Nhan N, Tick LJ. Natural language processing and the representation of clinical data. JAMIA. 1994;1:142-60.
84. Schuyler PL, Hole WT, Tuttle MS, Sherertz DD. The UMLS Metathesaurus: representing different views of biomedical concepts. Bull Med Libr Assoc. 1993;81:217-22.
85. Pearson P, Francomano C, Foster P, Bocchini C, Li P, McKusick V. The status of Online Mendelian Inheritance in Man (OMIM) medio 1994. Nucleic Acids Res. 1994;22: 3470-3.
86. Wall L, Schwartz R. Programming Perl. Sebastopol, CA: O'Reilly and Associates, 1992.
87. Stock K. Oraperl, version 2.004. 1994.
88. Applescript Language Guide. Reading, MA: Addison-Wesley, 1994.
89. Telescript Technology: The Foundation for the Electronic Marketplace. Sunnyvale, CA: General Magic, 1994.
90. Linthicum DS. The end of programming. BYTE. 1995;20(8): 69-72.
91. Hamill PVV, Drizd TA, Johnson CL, Reed RR, Roche AF. NCHS growth charts, 1976 (vol. 25). Rockville, MD: Vital Statistics Report, 1976.
92. Hinds A, Greenspun P, Kohane I. WHAM! A forms constructor for medical record access via the World Wide Web. SCAMC Proc. 1995:116-20.
93. Hickman KEB, Elgamal T. The SSL Protocol. Internet Draft. Mountain View, CA: Netscape Communications Corporation, 1995.
94. Rescorla E, Schiffman A. The Secure Hypertext Transfer Protocol. Internet Draft. Menlo Park, CA: Enterprise Integration Technologies, 1994.
95. Kohane IS, David P, McCallie J. A dynamically reconfigurable clinician's workstation with transparent access to remote and local databases. Snowbird, UT: First Annual American Medical Informatics Conference, 1990.
96. Kohane IS. Getting the data in: three-year experience with a pediatric electronic medical record system. SCAMC Proc. 1994:457-61.
97. Cimino JJ, Socratous SA, Clayton PD. Internet as clinical information system: application development using the World Wide Web. JAMIA. 1995;2:273-84.
98. McDonald CJ, Overhage JM, Tierney WM, et al. The Regenstrief medical record system: cross-institutional usage, note writing, and MOSAIC/HTML. SCAMC Proc. 1995: 1029.
99. Willard KE, Hallgren JH, Sielaff B, Connelly DP. The deployment of a World Wide Web (W3) based medical information system. SCAMC Proc. 1995:771-5.
100. Cimino JJ, Socratous SA, Grewal R. The informatics superhighway: prototyping on the World Wide Web. SCAMC Proc. 1995:111-5.
101. Galvin JR, D'Alessandro M, Erkonen WE, et al. The virtual hospital: an IAIMS on the Internet focused on education via integration of continuing education into the workflow. SCAMC Proc. 1995:1028.
102. Widman LE, Tong DA. EINTHOVEN on the World Wide Web: a tool for the analysis of cardiac arrythmias. SCAMC Proc. 1995:968.
103. Lowe HJ, Lomax EC, Poonkey SE. The World Wide Web: a review of an emerging Internet-based technology for the distribution of biomedical information. JAMIA. 1996;3:1-14.
104. Jagannathan V, Reddy YV, Srinivas K, et al. An overview of the CERC ARTEMIS project. SCAMC Proc. 1995:12-6.
105. Mowbray TJ, Zahavi R. The Essential CORBA Systems Integration Using Distributed Objects. New York: John Wiley & Sons, 1995.
106. Newcomer LR. Select SQL: The Relational Database Language. Upper Saddle River, NJ: Prentice-Hall, 1991.
107. Standards for medical identifiers, codes, and messages needed to create an efficient computer-stored medical record. JAMIA. 1994;1:1-7.