

# business integration

JOURNAL

## Integrating Electronic Health Records Using HL7 Clinical Document Architecture By Jim Klein

**ALSO IN THIS ISSUE:**

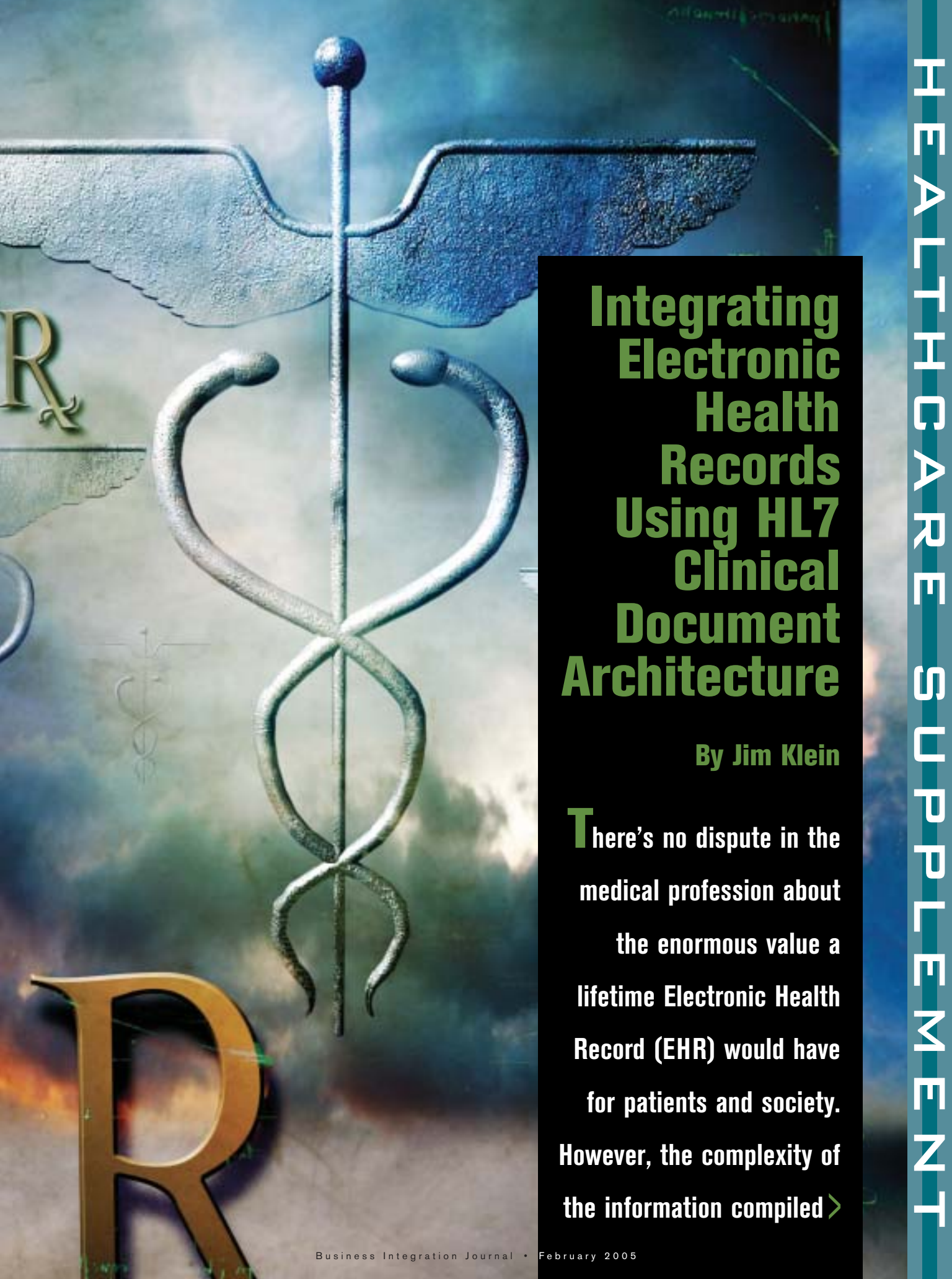
Healthcare Supplement

Hitting the Integration Sweet Spot

Integration Audit Methodology

Extending the SOA

Successful Dashboards  
for Healthcare



# Integrating Electronic Health Records Using HL7 Clinical Document Architecture

By Jim Klein

**T**here's no dispute in the medical profession about the enormous value a lifetime Electronic Health Record (EHR) would have for patients and society. However, the complexity of the information compiled >

by even a single provider is staggering. Achieving a lifetime EHR system requires overcoming significant application integration challenges.

This article describes Health Level Seven's (HL7's) Clinical Document Architecture (CDA), an XML standard for integrating healthcare delivery through the exchange of signed, electronic medical documents, and the formal Unified Modeling Language (UML)-based methodology behind it. The article examines the suitability of CDA documents as the atomic unit of information in an individual's lifetime EHR.

### The Challenge

In the U.S., when a consumer changes doctors, sees a different doctor for a second opinion, visits an emergency room, or is referred by their primary care physician to a specialist, it's rare for any electronic information about the patient's health status to be exchanged between the physicians involved. If there's any communication between physicians, it's almost certainly paper-based and either hand-carried or faxed. In ambulatory settings, including emergency rooms, new patients are routinely treated with no information on the patient's history other than what comes via patient interview.

Consequently, viable options aren't considered, tests are needlessly repeated, interventions are less effective, treatment is needlessly prolonged, and avoidable medical errors occur. Productivity is reduced, costs escalate, quality of care suffers, everyone is inconvenienced, and sometimes people die.

### The President's Vision

In his 2004 State of the Union address, President Bush acknowledged that, "By computerizing health records, we can avoid dangerous medical mistakes, reduce costs, and improve care." Then, on April 24, 2004, in a major policy speech, he pronounced, "Within 10 years, every American must have a personal Electronic Medical Record [EMR]." In

Despite its popularity within hospitals for simple data synchronization and exchange, HL7 V2 has proved to be an inadequate foundation for exchange of medical records information between different healthcare organizations.

that same speech, he announced the creation of a sub-cabinet-level position to coordinate a national initiative to achieve that goal; he described information standards as the cornerstone of that initiative.

The President's stated goal is semantic interoperability between the various EMRs (computer-based records about a patient within the scope of a single-provider organization) and ancillary systems (e.g., laboratories or pharmacies) that contain personal medical records information about an individual over the course of the person's lifetime (i.e., the EHR for that person).

### The HL7 Framework

HL7 is a not-for-profit, ANSI-accredited, standards development organization. HL7 develops standards for the exchange, management, and integration of data that supports clinical patient care and the management, delivery, and evaluation of healthcare services. HL7 has more than 2,500 members worldwide and 27 international affiliate organizations. Its more than 1,300 corporate members include the largest information systems vendors serving healthcare delivery organizations.

The popularity of HL7 standards burgeoned with the release of its Version 2 (V2) messaging standards in 1990. HL7 V2 messages are widely used within hospitals to synchronize and exchange infor-

mation between the ancillary departmental systems of the hospital to create a simple EMR. More than 93 percent of hospitals in the U.S. with healthcare IT systems reported using HL7 V2 standards in 2002.

Despite its popularity within hospitals for simple data synchronization and exchange, HL7 V2 has proved to be an inadequate foundation for exchange of medical records information between different healthcare organizations.

HL7 V2's shortcomings stem from the ad hoc nature of HL7 V2 message development. With no information model underlying HL7 V2 message definitions, there are ambiguities between messages and even between segments of the same message. Many message segments and elements are optional and HL7 V2 messages don't specify coded terminologies as value sets for message elements.

Vendor interpretations of a "standard" HL7 V2 interface vary considerably. HL7 V2-based interfaces are hardly "plug-and-play." There are many issues to be worked out at implementation time; significant customization and the use of integration middleware are almost always required. HL7 V2 is clearly not up to the task of enabling the semantic interoperability necessary to build a lifetime EHR.

HL7 formally began work in 1994 on a new generation of standards to support semantic interoperability. The principle behind HL7's new family of standards, designated Version 3 (V3), is that complete semantic understanding of a data exchange can be achieved only if the sender and receiver share a common model of the data that represents the domain of communication and if the sender and receiver use common sets of terms (codes) drawn from comprehensive, fully defined terminologies.

V3 standards are intended to provide a:

- Common information model across all domains of interest
- Robust data type specification
- Vocabulary specification, including a method for binding to terms from con-

## business integration journal takeaways

### BUSINESS

- Addressing business culture issues is as important as choosing the right technology in designing cross-enterprise integration solutions.
- Delivering some immediately tangible business benefit is critical to catalyzing adoption.
- The U.S. government's enormous purchasing power and regulatory authority can drive the promotion of application integration standards.

### TECHNOLOGY

- The use of XML to both markup text and serialize information objects lets CDA provide immediate value to the under-automated healthcare industry and enables semantic interoperability between electronic medical records systems.
- Semantic interoperability requires a common information model, a robust data type specification, and a method of binding to concept-based terminologies, all of which are provided by the HL7 development framework underlying CDA.

cept-based terminologies

- Formally defined process for designing specific structures to be exchanged between machines to meet a set of business requirements.

The foundation of V3 is the HL7 Reference Information Model (RIM), a static model of healthcare information as viewed within the scope of HL7 standards development activities (see Figure 1). The RIM is the ultimate source from which the information-related content of all V3 standards is derived; it's modeled using the Object Management Group's (OMG's) UML.

HL7 data types define the structure and constrain the allowable values of attributes in the RIM and derived information models. The HL7 data type specification declares the set of data types, identifies components of complex types, and defines relationships between data types. The HL7 data type implementa-

tion technology specification defines constraints and operations for data types and describes how data types are to be represented in XML. Data types are reusable atomic building blocks used to create all V3 information structures.

The V3 vocabulary specification defines vocabulary domains used in RIM-coded attributes and coded data type properties. A vocabulary domain is an abstract collection of interrelated concepts. It describes the purpose or intent of a coded attribute. A value set is a collection of coded concepts drawn from a single code system. A coded concept has a concept code assigned by the coding system and a concept designation that names the referenced concept.

The vocabulary specification includes the rules for linking a RIM-coded attribute, or HL7-coded data type, to an external terminology standard such as the College of American Pathologists Systematized Nomenclature of Medicine—Clinical

Terms (SNOMED CT), the most widely used clinical terminology system.

The HL7 Development Framework (HDF) is the formal methodology by which an HL7 V3 standard for a particular healthcare domain is developed. It's a highly structured process that makes heavy use of OMG's UML, including information models (the RIM and domain-specific information models derived from it), use cases, state machines, and interactions. Figure 2 shows a graphical summary of HDF.

**The Challenge of Physician Adoption**

Any proposed EHR solution must consider the extremely low rate of adoption by group practices and individual physicians. Only about 7 percent of U.S. physicians' practices have implemented an EMR system. There were 1 billion ambulatory visits in the U.S. in 2002, compared to 8 million hospitalizations. Most of the information to populate a lifetime EHR is trapped in paper documents in the many doctors' offices that a person visits. A lifetime EHR must meet the needs of physician practices to meet the needs of individual patients and the nation. A standard that supports semantic interoperability is necessary but not sufficient to build a lifetime EHR.

What's also required is a medical information representation standard powerful enough to record complex, attested (i.e., signed) medical information about individual patients, yet simple enough to be used by physicians who haven't implemented an EMR. It would be a bonus if the standard, while not requiring an EMR, encouraged physicians to install one.

Although few physicians have implemented an EMR, most have one or more Internet-connected PCs with browsers and e-mail accounts in their offices. These PCs are often running billing and claims submission applications.

Documents are the atomic elements from which medical records have been constructed. Clinical documents record subjective statements, objective findings, measurements, assessments, actions, and plans from the diagnostic and care-giving processes. Physicians generally sign clinical documents before they're considered part of the patient's permanent medical record.

**CDA**

HL7's CDA defines XML-based document markup that standardizes the structure and content of clinical documents for exchange. CDA's principal goal is to facilitate the continuity of

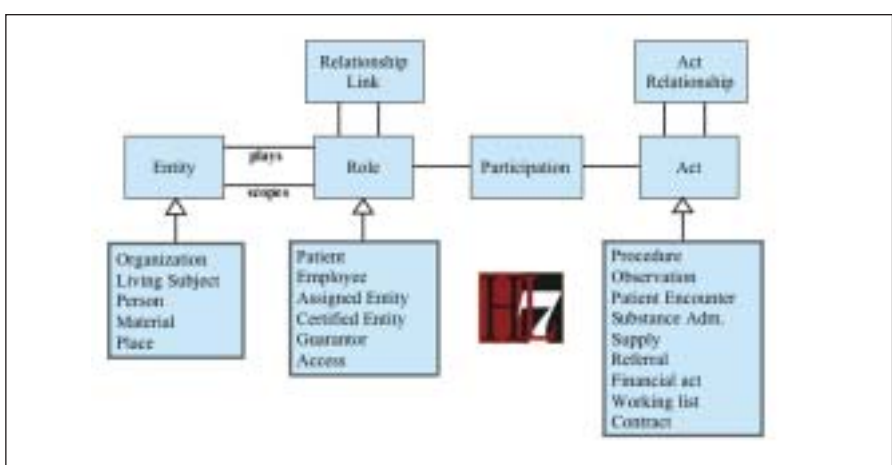


Figure 1: Core Classes of the HL7 Reference Information Model (RIM) HL7 Standards Development Methodology

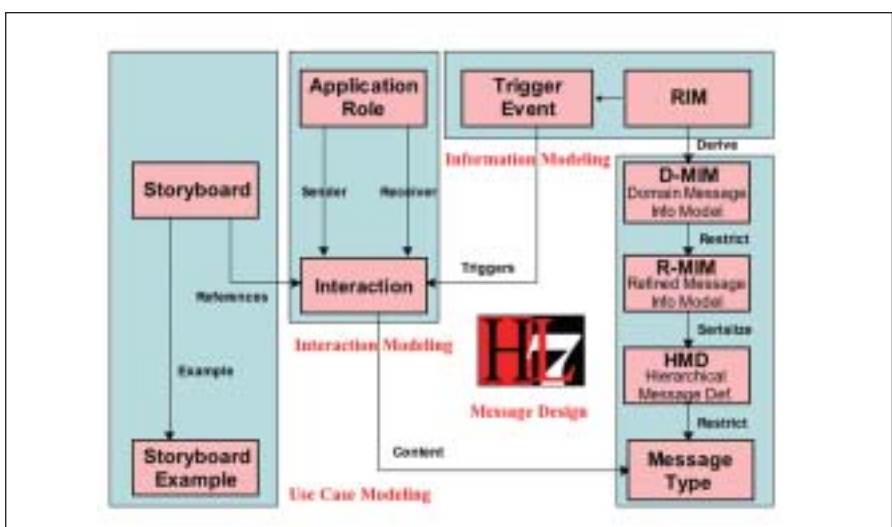


Figure 2: A High-Level Overview of HL7 V3 Development Methodology

patient care across disparate organizations by creating a human-readable, machine- and application-independent format for the exchange of electronic clinical documents. This promotes the longevity of clinical records and minimizes barriers to creation and use.

The HL7 CDA specification, designed to meet the historic and legal requirements of a clinical document, describes a clinical document as having these characteristics:

- **Persistence:** A clinical document continues to exist in an unaltered state, for a period defined by local and regulatory requirements.
- **Stewardship:** A clinical document is maintained by an organization entrusted with its care.
- **Potential for authentication:** A clinical document is an assemblage of information that's intended to be legally authenticated.
- **Context:** A clinical document establishes the default context for its contents.
- **Wholeness:** Authentication of a clinical document applies to the whole and not portions of the document.
- **Human readability:** A clinical document is human-readable.

A CDA document is a contextually complete information object that can include text, images, sounds, and other multi-media content and can stand alone, outside the environment in which it was created or communicated.

Human readability is a critical element in lowering the technical and workplace cultural barriers to the adoption of CDA. All CDA documents can be rendered any XML-aware browser using the same generic stylesheet.

CDA includes the capability to optionally include encoded medical information when the source application can support it. CDA leverages the enhanced semantic interoperability conferred by the HL7 V3 methodology to ensure that, in the broadest possible context, machines can process CDA-compliant documents. CDA-compliant documents use V3 data types and the V3 terminology-binding framework and derive their meaning from the HL7 RIM.

Figure 3 summarizes the differences between electronic documents and messages, the traditional medium of loosely coupled application integration.

#### CDA Document Structure

CDA documents are XML documents that consist of a header and a body (see Figure 4).

The schema for the header is derived from the HL7 RIM in the manner prescribed in the V3 development methodology. The header's purpose is to enable clinical document exchange across and within institutions; facilitate document management; and facilitate compilation of an individual patient's clinical documents into a lifetime electronic patient record.

The header classifies the document, uniquely identifies the instance of the document, and defines its relationship to any previous instances that may exist. A sample of information in the header includes:

- Identification of the subject referenced in the document (usually, but not always a patient)
- Author of the document
- Participants in activities described in the document
- Participants in the preparation of the document, including the person who has attested to or "signed" the document.

The body consists of one or more sections. Sections may be nested within sections (see Figure 4). Each section has a mandatory narrative block (bounded by <Text> tags) that contains the human-readable text to be rendered. Any section may also have zero or more structured entries. These optional structured entries encode content for processing by applications. A structured entry encodes content from the narrative block of the same section in which that structured entry appears.

The markup was designed to be easily processed by common scripting languages to render the narrative block on standard browsers and printers of limited capabilities. Pieces of the narrative block surrounded by <content></content> tag sets become content elements. The value assigned to an optional "ID" attribute in the opening <content> tag allows CDA structured entries to refer to content elements.

The structure and meaning of a CDA structured entry is derived from the RIM and encoded in conformance with the V3 data type and vocabulary specifications. This contributes to the longevity of the encoded information and its predictable use. Structured entries can reference content elements in the narrative block (see Figure 5). CDA doesn't specify the manner or timing of the creation of the narrative block and any related structured entries. Many possibilities exist.

### CDA documents::HL7 messages

	Documents	Messages
Lifetime	persistent	temporary
Communication	human-to-human	system-to-system
Relation to caregivers	care-givers are trained to create documents ...	... not messages
Legal aspects	have legal standing	signed? legally accepted?
Source	defined by precedent	designed per use case
Context	document as a whole	must be defined in each segment

Copyright IBM, Lisa Rishler

Figure 3: Comparison of CDA Documents and HL7 Messages

```

<clinical document>
  < header> ... </header>
  <body>
    <section>
      <text>...</text>
      <structured entry>...</structured entry>
      <structured entry>...</structured entry>
    </section>
    <section>
      <section>...</section>
    </section>
  </body>
</clinical document>

```

Figure 4: A Simplified View of Major Components of a CDA Document

With the right supporting applications software, a human expert could add structured entries to a section after the creation of the narrative block, or the narrative block could be automatically generated from structured entries. Structured entries could even be generated by a medical natural language-processing system that's HL7 V3-aware. The referencing mechanism supports all these scenarios.

Multi-media elements in the narrative block can reference structured entries. This has many uses, including annotating images.

#### CDA: The Foundation for EHR

The simplicity of CDA, based on its document paradigm, facilitates physician use. This makes CDA a candidate for the foundational technology of a national EHR. The case becomes even stronger because CDA can leverage the V3 methodology to support semantic interoperability between advanced healthcare applications.

Browsers can always render a CDA document using a simple stylesheet that can be applied to any CDA document. The signing of a CDA document applies to the human-readable narrative block, so structured entries, if present, may be safely ignored by the document receiver. This means that an EHR based on CDA documents would be useful to any physician with a browser. In fact, CDA documents can be printed and hand-delivered or faxed to physicians without PCs. However, two important questions must be addressed concerning a lifetime EHR based on CDA documents:

- How will healthcare provider organizations and ancillary medical service organizations be induced to create CDA documents?
- Where will the CDA documents that constitute a person's EHR be stored

and how will they be accessed to ensure protection of patient privacy?

#### Bootstrapping CDA Document Ubiquity

Several federal government initiatives are already focused on CDA as the mechanism for exchanging semantically interoperable healthcare information. The forthcoming clinical attachments regulation from the U.S. Department of Health and Human Services (HHS) will specify CDA as the standard for sending clinical information in support of a claim submitted to a payer, such as Medicare or a private health insurance company. The National Electronic Disease Surveillance System (NEDSS) specifies HL7 V3 messages for reporting laboratory results to public health authorities wherever feasible. Converting these messages into CDA documents would be a simple matter of producing a human-readable narrative from the completely structured and encoded information in the V3 message payload.

These government initiatives and others that reflect governmental economic clout or regulatory authority can help establish CDA as the syntax and semantics for communicating healthcare information to the federal government. It's likely that the aggressive pursuit of such a strategy by the U.S. government would prompt laboratories, pharmacies and community hospitals to support CDA as the path of least resistance for communicating patient medical records information to physicians.

#### Storing and Accessing EHR Content

The Office of the National Coordinator for Health Information Technology (ONCHIT), created by the President in April 2004 and charged with planning for the implementation of the President's EHR vision, favors the creation of private sector entities called Regional Health Information Organizations (RHIOs) to support regional EHR reposi-

tories, which would be subsequently federated to deliver a national EHR solution.

These EHR repositories might be "virtual" (i.e., there need not be a central database server), but rather the clinical content could remain with its originators. In that case, the RHIO would maintain a directory of the patients in its region and a registry of the information available on each patient and its location. CDA documents would work well as the standard for exchanging information, regardless of whether the RHIO's repository is physical or virtual.

If RHIOs generate political opposition from privacy advocates and are never widely implemented, the point-to-point exchange of persistent, attested, human-readable CDA documents between physicians, labs, pharmacies and hospitals would still provide an enormous boost to the continuity of patient care. CDA documents, encapsulated as Adobe Portable Document Format (PDF) documents or Microsoft Word or InfoPath documents, could be easily collected, retained and communicated to physicians by patients themselves.

#### The Bottom Line

CDA does simple things simply; it provides XML markup for the exchange of human-readable clinical documents that support physician use with little or no computerization. CDA also supports the semantically interoperable exchange of complex medical information between healthcare applications by virtue of its adherence to the HL7 V3 development methodology. This combination of simplicity and power make CDA a good candidate for the healthcare information representation and exchange standard that provides a basis for a lifetime EHR for U.S. citizens. **bij**

#### About the Author



Jim Klein is director, Healthcare Technology, at InterSystems Corp., a leading provider of database technology in healthcare. InterSystems develops and markets the Ensemble universal integration platform.

Previously, he was vice president and research director at Gartner, Inc., where he was the lead analyst on the interpretation of HIPAA privacy and security rules. He has more than 25 years of experience in healthcare technology.  
e-Mail: [jklein@intersystems.com](mailto:jklein@intersystems.com)  
Website: [www.intersystems.com](http://www.intersystems.com)

```
<section>
  <text>
    There is a history of <content ID="a1">Asthma</content>
  </text>
  <entry>
    <Observation>
      <code code="48318009" codeSystem="&SNOMEDCT;"
        displayName="Prior dx"/>
      <value xsi:type="CD" code="195967001"
        codeSystem="&SNOMEDCT;" displayName="Asthma">
        <originalText><reference
value="#a1"/></originalText>
        </value>
      </Observation>
    </entry>
  </section>
```

Figure 5: Example of a Reference to Narrative Block From a Structured Entry