

Metadata and Meaningful Use

[Save to myBoK](#)

By Allison Viola, MBA, RHIA, and Shefali Mookencherry, MPH, MSMIS, RHIA

The healthcare industry will find out soon if ONC intends to include metadata requirements in stage 2 of the meaningful use program. Many feel it is too soon. But given metadata's potential to support health information exchange, the expanded and standardized use of metadata tagging in healthcare is ultimately a matter of when, not if.

It should come as no surprise that when in August 2011 the Office of the National Coordinator for Health IT (ONC) issued an advance notice of proposed rulemaking on the use of metadata standards the health IT community expressed a collective gasp. Despite so many health IT and healthcare reform initiatives under way, ONC indicated it was considering recommendations to add the use of metadata-tagged data elements in stage 2 of the meaningful use program.

ONC issued the notice in response to a December 2010 report released by the President's Council of Advisors on Science and Technology. In that report PCAST outlined strong recommendations to speed the growth of health information exchange through the use of metadata tags.

The PCAST report called for a universal language for exchanging health data, an extensible markup language (such as a variation of XML) where health data would be separated into the smallest individual pieces that make sense to exchange. These data elements would be accompanied by mandatory metadata tags or minimal standards that describe the data and the patient's preferences for the data's uses, security, and privacy protections. Such a solution would enable healthcare providers to share health information reliably and effectively.

What's more, PCAST urged the federal government not to wait. ONC "should signal now," PCAST wrote, that EHR systems must have the ability to exchange health data in a universal manner based on metadata-tagged data elements by 2013 in order to qualify for use in the meaningful use program.

In principle, the healthcare industry agrees on the value of greater metadata use. A range of entities have recognized that metadata tagging has the potential to increase the usefulness and integrity of data for health information exchange by better describing the information being shared. However, PCAST's recommendation that ONC "move boldly" is too bold for many. In responding to ONC's advance notice, commenters cited multiple barriers to the quick implementation of metadata requirements.

The release of the proposed rule on stage 2, expected this month, will be the first test of how quickly ONC intends to include metadata in the meaningful use program. But ultimately the question of applying metadata to discrete pieces of health information is a matter of when, not if.

What Are Metadata?

Metadata can be envisioned as words in a dictionary-words that describe words. Metadata provide information about another piece of data, such as:

- The time and date of its creation
- Its creator or author
- The means of its creation
- Its purpose

Metadata can describe the standards used to create a piece of data or the location on a computer network where it was created.

For example, a digital photograph may include metadata that describe the photograph's size, color depth, image resolution, and the date and time it was created. A text document's metadata may contain information about the document's size, who created it, when it was written, and a short summary of its contents.

Metadata have long been used in various forms as a means of cataloging archived information. The Dewey Decimal System employed for the classification of library materials is an early example of metadata use. Library catalogs used 3 x 5" cards to display a book's title, author, subject matter, and brief description; an abbreviated alphanumeric identification system indicated the physical location of the book within the library. Such data helped identify, classify, and retrieve books.

Metadata can describe data, processes, and systems. They can be technical, business, or operational in nature. They can relate database schemas, data definitions, key performance indicator definitions, interface specifications, report layouts, source-to-target mappings, and process execution statistics.

Definitions

Metadata-structured information that describes, explains, locates, and otherwise makes it easier to retrieve and use an information resource.

Administrative metadata-metadata related to the use, management, and encoding processes of digital objects over a period of time.

Descriptive metadata-metadata that describes a work for purposes of discovery and identification, such as creator, title, and subject.

Structural metadata-metadata that indicates how compound objects are structured, provided to support use of the objects.

Scheme (schema)-a metadata element set and rules for using it.

Semantics-the names and meanings of metadata elements.

Syntax-rules for how metadata elements and their content are encoded.

Structuring Metadata

Metadata **schemes** (also called schema) are sets of metadata elements designed for a specific purpose, such as describing a particular type of information resource. The definition or meaning of the elements themselves is known as the **semantics** of the scheme. The **values** given to metadata elements are the content. Metadata schemes generally specify names of elements and their semantics.

There may also be **syntax** rules for how the elements and their content should be encoded. Metadata can be encoded in any definable syntax. Many current metadata schemes use SGML (Standard Generalized Mark-up Language) or XML (Extensible Mark-up Language).

Excerpted from: National Information Standards Organization. "Understanding Metadata." 2004. www.niso.org.

Metadata's Uses in Healthcare

Metadata have application in healthcare at any point where they may provide useful information about a particular piece of health data. This includes indexing information for storage and retrieval as well as describing data when exchanging it beyond the facility walls.

The best way to manage data for advanced data mining, PCAST wrote, is to "break it down into the smallest individual pieces that make sense to exchange or aggregate." The authors call these pieces "tagged data elements," because "each unit of data

is accompanied by a mandatory 'metadata tag' that describes the attributes, provenance, and required privacy protections of the data."

Metadata can be developed and managed through a metadata repository that indexes and consolidates metadata from different documents and information systems. From there the repository can integrate with an electronic health record system. Queries against a patient name will retrieve all relevant data for that patient.

In this manner, metadata can be used in a decision support tool to identify contradicting medications, allergies, and other factors that would affect the patient's care. Take for example a physician using a medication system that is not tied to other aspects of the patient's record and does not reflect the patient's other disease states. If the physician prescribes the patient aspirin for a chronic headache, metadata could be used to retrieve other patient information, alerting the physician that the patient currently takes a blood thinner.

On the other hand, metadata could be viewed as a security vulnerability. If not secured, they could expose information and software assets to theft or infiltration via hijacking, compromised certification, or viruses. In healthcare, breach of information can lead to damaging disclosure of confidential patient information.

Metadata can provoke controversy related to confidentiality, compliance, and litigation because they can be used to confirm who has seen or edited a record and in some cases establish a series of actions. This can benefit physicians by giving them a more accurate picture of the sequence of events in a patient file. It could also offer patients and their attorneys evidence during litigation that malpractice took place-or, conversely, offer providers evidence that it did not. Metadata trails can alert organizations that staff have been accessing patient records they are not authorized to view, which could in turn be an effective deterrent against snooping.

ONC's advance notice proposes three categories of metadata, which include recommendations from the Health IT Standards Committee's Metadata Power Team: patient identity, provenance, and privacy.

Patient identity metadata are used to select a particular patient from a population. The proposed standard would require:

- Name
- Date of birth
- Address
- Zip code
- Patient identifier(s)-unique identifying information such as the last four digits of Social Security numbers, driver's license number, the provider's patient identification number, or some combination thereof

ONC sought comments on whether additional elements are needed for patient identity categories or if any of the listed elements should be removed. ONC also sought input on what to do when address information is not available, proposing that the healthcare institution's address be used. Commenters were almost evenly divided on this approach.

It appears that, at the very least, multiple identifiers must be used to avoid patient identification issues. For example, there may be multiple John Smiths born on the same day.

What Do Metadata Look Like?

Computers are alerted to metadata by common tags that identify them. ONC includes examples of how the metadata items proposed in the advance notice could be expressed using the Clinical Document Architecture, Release Two standard. Several of the examples are included below.

CDA R2 is used in health information exchange to describe the encoding, structure, and semantics of clinical documents. It was developed by HL7 International and became an ANSI-approved standard in 2005. CDA R2 uses XML, a common computer markup language whose code is readable by humans and computers both.

Metadata Element	Expressed According to HL7 CDA R2 Requirements
Patient ID-Date of birth	<birthTime value="19600427"/>
Provenance-Timestamp	<effectiveTime value="20101217093047"/>
Privacy-Content data type	<code code="11488-4" displayName="Consultation note" codeSystemName="LOINC"/>

Provenance metadata describe the data set's history, origin, and any modifications made to it since its creation. This would include:

- Tagged data element identifier (i.e., linking other tagged data elements to each other, such as linking a diagnostic study to the patient encounter that led to the test)
- Time stamp
- Actor and actor affiliation

Provenance metadata thus would inform patients and healthcare providers who created the record and when it was created. As noted, such data also could be highly significant in legal proceedings. A plaintiffs' attorneys, for example, could have a field day seeking information about alterations or late amendments to a record. Poorly managed provenance metadata could prove a liability, misrepresenting or invalidating a file's history, such as its author or modifications. The proper use of metadata offers substantial benefits in responding to discovery requests by facilitating more effective and efficient searching and retrieval of information.

Understanding when metadata need to be preserved represents one of the biggest challenges in electronic discovery. The extent to which metadata are discoverable will depend on the needs of the individual case. In assessing their preservation strategies, organizations should note that the failure to produce metadata may deprive them of the opportunity to contest the authenticity of the document later if the metadata are material to that determination.

Privacy metadata would convey a patient's consent related to the sharing of his or her information in health information exchanges. The advance notice proposed these metadata include a policy pointer (e.g., a URL pointing to the current privacy policy) and metadata about the content itself. This standard would require that metadata be secure (i.e., encrypted) and would only address patient preferences related to sharing information within an exchange.

ONC was open to comments beyond this narrow scope of patient summary care records, as it sees metadata supporting the growth of nationwide health information exchange. ONC believes that this granular level of data exchange and accuracy will increase healthcare providers' confidence in data exchange networks and promote their widespread use.

In addition, metadata promote the healthcare reform goal that providers use real-time and accurate data for quality improvement activities. In the future, information contained in metadata could be used for research and public health purposes.

Metadata and Meaningful Use

If ONC requires metadata use in stage 2 of meaningful use, electronic health record (EHR) technology will need to meet specified metadata standards in order to become certified for use in the program.

In its advance notice of proposed rulemaking ONC sought answers to 20 questions about elements of metadata in electronic health information. It wrote that the "immediate scope" of its notice was the association of metadata with "summary care records," or how patients get a summary of their records from a provider's EHR, either on paper or through an electronic

transmission to a personal health record. However, metadata tagging does have the potential to increase the reliability, dependability, and trustworthiness of health information exchanges by better describing the data being exchanged.

Industry response to the advance notice was generally supportive. ONC received 51 comments on its advance notice, and 48 commenters supported the use of metadata (two did not; one made no recommendation).

There is consensus on the need to exchange health information and conduct clinical decision support to increase quality, improve patient care, and reduce costs. With the shared savings program, a provision outlined in the Affordable Care Act, accountable care organizations will need to work in concert to reduce costs associated with Medicare beneficiaries. This serves as an excellent backdrop from which to build community health information exchange.

Comments to ONC acknowledged the value of leveraging metadata standards to improve data usefulness and liquidity in support of health data exchange. ONC was applauded for embracing the recommendations and vision outlined in the PCAST report: to enable the ability to share and understand data from disparate systems; develop uniform standards; and promote interoperability among health IT systems.

However, there is a common concern in the industry over rushing the required use of metadata. Metadata-and the standards governing metadata use-have not been widely implemented within the healthcare community to date. Some commenters expressed misgivings about implementing metadata requirements through the federal rulemaking process.

ONC's summary of the comments noted that many wanted industry standards development organizations to set metadata standards. Nine of the 51 commenters stated specifically that metadata standards were not ready for inclusion in stage 2 of meaningful use.

Apart from the need to further develop metadata standards, testing and implementing metadata standards in time for stage 2 would be a tall order. The healthcare industry is not ready or capable of supporting regulatory requirements to integrate metadata standards, as its bandwidth is stressed with transitioning to ICD-10, creating accountable care organizations, and preparing for other requirements in stage 2.

Barriers and Considerations

Although ONC received general support for moving the PCAST recommendations forward, many of the comments it received on the advance notice emphasized barriers and challenges to speedy implementation.

In its report, PCAST acknowledged that the industry lacked standards on metadata use. It noted the need to develop initial minimal standards for the metadata associated with tagged data elements and a road map for a more comprehensive set of standards over time. ONC, it recommended, should "facilitate the rapid mapping of existing semantic taxonomies into tagged data elements, while continuing to encourage the longer-term harmonization of these taxonomies by vendors and other stakeholders."

The providers, payers, vendors, and policy organizations submitting comments on the advance notice raised additional challenges associated with implementing metadata standards in the current health IT environment.

Data sets. Before the industry can begin development of metadata standards, ONC must identify a uniform and unambiguous data set. In the advance notice ONC described a proposal that metadata be expressed according to the requirements in the Clinical Document Architecture, Release Two header, a recommendation put forward by its HIT Standards Committee. CDA R2, a document format standard developed by HL7 International, provides wide coverage across metadata elements, and working from a single standard would make implementation easier, the committee reasoned.

However, the proposal has perceived drawbacks. CDA R2 inextricably merges data and presentation in the same format. Secondly, as ONC acknowledges, the standard does not support some of the metadata elements outlined in the proposed rule, such as the patient's "display name," which ONC proposes be included in the header.

An argument for not managing metadata as proposed is that success will be better achieved by making the standardized data set widely adopted through multiple, flexible implementation methods.

The American Academy of Family Physicians commented that by restricting the metadata to a single implementation ONC would reduce the ability of the market to innovate. AAFP firmly recommended that ONC separate the data definition of the metadata needed to support health information exchange from its representation in a particular standard. The important work is to define an unambiguous data set. How to represent it, such as in XML, is unimportant or, more accurately, fertile ground for innovation.

Other commenters also recommended that ONC specify only the metadata elements to be used, not the representation structure for those elements. If ONC focuses on the data required to support health information exchange, the presentation of the data will follow and should not be prescriptive.

Pilots and demonstrations. Many of the industry responses called for the need to conduct further evaluation, analysis, and testing of the metadata approach before incorporating it into the rulemaking process for stage 2. Current metadata standards are not mature and robust enough to support HIE at this time, and it would be premature for ONC to incorporate them into stage 2 without having the opportunity to fully test, demonstrate, and evaluate this technological approach.

Infrastructure. For metadata standards to be formally adopted as national standards in the meaningful use program or any other program, ONC must establish a policy framework and infrastructure that will support and govern the integration and use of these standards.

Alignment with other policy objectives. ONC received strong encouragement and recommendations to align and harmonize metadata performance standards with other policy objectives under development or currently in place to prevent an increase in burden and conflicting processes.

Privacy policy pointers. The proposal to use privacy policy pointers solicited strong feedback from the healthcare community. Believing that it would not be feasible to include a privacy policy with each tagged data element because policy can change over time, ONC suggested that a pointer link to an external registry. Commenters expressed concern regarding the ability of this function to be scalable as objects age and policies are changed.

The policy pointer standards would have to incorporate all federal and state regulations, patient preferences, and other privacy requirements for implementation. This does not address the need for variation in the ability to parse the data for limited viewing from specific providers or other recipients of health information. The ability to support a variety of circumstances in defining the privacy standards is complex and should be considered further.

Digital signatures. Several organizations recommended that ONC remove the requirement to use digital signatures for nonrepudiation. There was support in using this functionality; however, there were suggestions that digital signatures coupled with metadata should not be applied. The use of digital signature technology can be applied separately at the metadata level rather than the content level to ensure authenticity and integrity of the data, allowing for a layered approach.

According to comments submitted by GE Healthcare, "More basically, not all uses of data require the very high level of assurance of non-repudiation that a digital signature provides. Forcing digital signatures as metadata will make the model very expensive. The rationale for this approach is the same as that for the proposed, and in our view correct, justification for a separate confidentiality layer."

Two initiatives currently directed by ONC under the Standards and Interoperability Framework can inform development of metadata initiatives: the Data Segmentation and Query Health pilots.

In particular, the Data Segmentation pilot program will "enable the implementation and management of varying disclosure policies in an electronic health information exchange environment in an interoperable manner with the goal to produce a pilot project allowing providers to share portions of an electronic medical record while not sharing others, such as information related to substance abuse treatment, which is given heightened protection under the law," according to a description on the framework Web site.

Based upon patient consent decision, applicable law, and policy, use cases will include metadata tagging to highlight privacy attributes in the patient's records. Relevant standards, policies, and technologies will be evaluated for use in developing this model.

More such work needs to be done to develop and verify metadata standards before they can be broadly implemented across the healthcare community. Using metadata to make better clinical decisions may ultimately benefit patients and healthcare providers in the long run, but it will take investment and development to realize the full benefits.

References

Office of the National Coordinator for Health Information Technology, Department of Health and Human Services. "Metadata Standards to Support Nationwide Electronic Health Information Exchange." *Federal Register* 76, no. 153 (Aug. 9, 2011): 48769–76. www.gpo.gov/fdsys/pkg/FR-2011-08-09/pdf/2011-20219.pdf.

Office of the National Coordinator for Health Information Technology, Department of Health and Human Services. "S&I Framework Data Segmentation." <http://wiki.siframework.org/Data+Segmentation>.

President's Council of Advisors on Science and Technology. "Realizing the Full Potential of Health Information Technology to Improve Healthcare for Americans: The Path Forward." December 8, 2010. www.whitehouse.gov/administration/eop/ostp/pcast.

The Sedona Conference. "The Sedona Principles: Best Practices Recommendations and Principles for Addressing Electronic Document Production, 2nd edition. June 2007. www.thesedonaconference.org/content/miscFiles/TSC_PRINCP_2nd_ed_607.pdf.

Allison Viola is director of federal relations at AHIMA. Shefali Mookencherry (smookencherry@hayesmanagement.com) is a principal healthcare consultant for Hayes Management Consulting.

Article citation:

Viola, Allison F.; Mookencherry, Shefali. "Metadata and Meaningful Use" *Journal of AHIMA* 83, no.2 (February 2012): 32-38.

Driving the Power of Knowledge

Copyright 2022 by The American Health Information Management Association. All Rights Reserved.