



# Health IT Standards Committee

A Public Advisory Body on Health Information Technology to the National Coordinator for Health IT

June 22, 2011

Farzad Mostashari, MD, ScM  
National Coordinator for Health Information Technology  
Department of Health and Human Services  
200 Independence Avenue, SW  
Washington, DC 20201

Dear Dr. Mostashari:

The HIT Standards Committee (HITSC) gave the following broad charge to the Metadata Analysis Power Team (Power Team):

### **Broad Charge for Metadata Analysis Power Team**

The Metadata Analysis Power Team (Power Team) was formed as a workgroup to make recommendations to the HIT Standards Committee (HITSC) to identify the metadata elements necessary to support a new exchange architecture with a universal exchange language (UEL) to enable clinicians and patients to assemble a patient's data across organizational boundaries. Specifically, the Power Team was asked to identify and recommend metadata elements and standards in three areas: patient identification, provenance, and privacy.

On June 22, 2011, the Power Team presented its findings to the Committee, which were discussed and subsequently approved. This letter provides recommendations to the Department of Health and Human Services (HHS) on metadata elements and standards in three areas: patient identification, provenance, and privacy.

### **Background and Discussion**

As background, in December 2010, the President's Council of Advisors on Science and Technology (PCAST) issued a report entitled, "*Realizing the Full Potential of Health Information Technology to Improve Healthcare for All Americans: The Path Forward.*" The report stated the best way to manage and store data to support advanced data-analytical techniques is to break data down into the smallest individual pieces that make sense to exchange or aggregate. These individual pieces are called "tagged data elements" (TDE) because each unit of data is accompanied by a mandatory "metadata tag" that describes the attributes, provenance, and required security protections of the data. Both the data and the tag are protected against unauthorized access or data corruption during transport.

The Power Team reached several important decisions regarding which metadata elements and health information technology exchange standards should be contained with patient identity, provenance and privacy. These metadata elements are often referred to as an envelope that will wrap TDEs.

## **Recommendations**

The recommendations fall into one of 3 categories:

1. Patient Identification Metadata Elements and Standard
2. Provenance Metadata Elements and Standard
3. Privacy Metadata Elements and Standard

### **1. Patient Identification Metadata Elements and Standard**

Metadata data pertaining to patient identity should include the following:

- patient's name
- date of birth
- current zip code
- patient identifiers
- address

#### Rationale:

We believe that this set of patient identity metadata represent the minimum elements that are required to uniquely select a patient from a population with a guaranteed degree of accuracy. The identity elements are sufficient for cases where the listed elements are all present. There are situations (not yet named newborns, comatose patients, etc.) where the listed elements may not be available and the ability to match to a unique patient would be in doubt.

A study conducted by RAND

([http://www.rand.org/pubs/monographs/2008/RAND\\_MG753.pdf](http://www.rand.org/pubs/monographs/2008/RAND_MG753.pdf)) identified 4 components that provide an error free composite index: name, date of birth (DOB), zip code and the last four digits of the SSN, in an 80 million record demographic database.

Additional recommendations for patient identification:

- Add a display name element to the HL7 CDA R2 standard to accommodate non-western names. This may require an extension of the HL7 schema.
- Use a URI to act as a namespace for the identifier, as opposed to an OID as used in HL7 CDA R2. This allows for an extensible, flexible mechanism to uniquely identify an individual, without having to specify explicitly what type of identifier is used.

#### Standard:

We recommend that metadata expressing patient identifiers use the HL7 CDA R2 header format. We believe that this XML based format for describing generic clinical documents can best accommodate international representation of names. Additional information supported by the CDA R2 header could be included if desired.

## 2. Provenance Metadata Elements and Standard

Metadata contained within the provenance envelope should include the following:

- tagged data element (TDE) identifier
- a time stamp
- the actor
- the actor's affiliation
- the actor's digital certificate

### Rationale:

The envelope will provide information permitting the recipient to judge whether a trusted source sent the data, when it was packaged, and whether any content tampering took place. The TDE identifier will preserve clinical context by allowing other TDEs to link to this particular instance and allow users to keep a log of the set of TDEs used for a particular task. The time stamp metadata will express when the envelope was signed. The metadata describing the actor (in the form of a digital certificate) should include the name of the actor who signed the envelope and the organizational affiliation of the actor. Note that this scheme allows for exchanges involving either organizational "actors" and individual "actors."

Additional recommendations for provenance:

- The use of an X.509 certificate to digitally sign the envelope contents. The use of a digital signature fulfills two requirements outlined by the Power Team:
  - Non-repudiation. If a medical decision is based on the contents of a TDE, it is important that the contents cannot be later denied by the origin.
  - Tamper-resistance. It is important to be able to verify that the content and metadata of a TDE have not been tampered with after production otherwise new content could be substituted by a malicious intermediary and TDE could not be trusted.
  - Without such a security mechanism the metadata will be at best advisory and any real use will require the kind of security requirements discussed above.
- While the actor and actor's affiliation are expressed within the X.509 certificate, we feel there should be the additional optional metadata fields for actor/affiliation for the following reasons:
  - In cases where the message signing authority is different from the party that generated the message, a distinction is necessary between the information in the X.509 certificate, and the actual originator/owner of the information.
  - Software which understands HL7 CDA R2 headers, but does not have the ability to process more complex cryptographic signatures can have access to the information.
  - More granular information can be contained in the optional metadata elements than in the certificate, to include additional properties of the actor that may be relevant for provenance but which are not supported or necessary for the digital signature.

### Standard:

We recommend that these metadata components be expressed using the HL7 CDA R2 format. While many standards support the required metadata elements, the Power Team

determined that the use of HL7 CDA R2 for provenance would be complimentary with its use for identity, and should therefore be adopted as the base standard. We also propose that an optional portion of the actor/affiliation metadata should point to the entity record in the Enterprise-Level Provider Directory, which may be a URL. If available, this will decrease complexity in the TDE and enable lookup of the source information in a provider directory.

### 3. Privacy Metadata Elements and Standard

Metadata pertaining to privacy should include the following:

- **Policy Pointer.** The Policy Pointer is a URL that points to the privacy policy in effect at the time the tagged data element (TDE) is released.
- **Content Metadata.** Content metadata are needed to enforce current federal and state policies as well as to anticipate more granular policies that may be defined. Content metadata comprises 2 components:
  - o Datatype: information category from a clinical perspective
  - o Sensitivity: indicates special handling that may be necessary per the referenced policy

#### Rationale:

We initially considered three components necessary to enforce privacy: the policy, metadata about the content, and metadata about the requestor. However, information about the requestor would be used by the sender to mediate the request, but would not need to be tagged onto the data exchanged in an authorized request. It was determined that including the policy with each TDE was not feasible because policy changes over time. Therefore it was agreed that a pointer to an external policy registry would be most appropriate, though we did not address the specifics of how these registries might be implemented. We are therefore restricting our metadata recommendation to the content.

Additional recommendations for privacy:

- In order to provide coded values for Datatype, the LOINC codes specified in the CDA document code element are suggested. LOINC codes are suggested because they provide additional granularity.
- The HL7 vocabulary for sensitivity will need to be expanded. A proposed starter set could include:
  - Substance Abuse
  - Reproductive Health
  - Sexually Transmitted Disease
  - Mental Health
  - Genetic Information
  - Violence
  - Other

#### Standard:

We recommend that these metadata components be expressed using HL7 CDA R2 header

elements since this standard is already being used for defining the metadata elements for Patient Identity and Provenance.

### **Summary**

The HIT Standards Committee recommends that metadata for patient identification, provenance, and privacy be expressed using elements from the HL7 Clinical Document Architecture-Version 2 (HL7 CDA R2) header with the following additions:

- Extend name to include an XML element whose value is a string that captures the patient's name as it should be displayed or written.
- Extend the existing HL7 id element that allows a URI to be used as the value of the root attribute instead of the currently allowed UUID, OID or HL7 identifier.

Use of an established standard allows the relevant standards organization to continue to maintain the standard and provide ongoing support. As HL7 CDA R2 header is a current health standard, many EHRs can already understand and disseminate data in this format. Future support and maintenance of the standard is viewed to be better with HL7 as this is an active standard. HL7 CDA also provides wide coverage across the metadata elements, and a single standard would lead to easier implementation than multiple standards. The Committee does suggest, however, that HL7 eliminate licensing fees for use of the CDA R2 header to make use of the standard more widely available.

An example of the metadata elements using HL7 CDA R2 is provided in Attachments A and B.

The aforementioned recommendations were targeted to address a set of questions raised by the Office of the National Coordinator for Health IT (ONC). We propose these standards with the understanding that ONC will conduct further testing and evaluation prior to proposing these through rulemaking.

Sincerely,

/s/

Jonathan Perlin  
Chair  
Health IT Standards Committee

John Halamka  
Vice Chair, Health IT Standards Committee

/s/

cc: Doug Fridsma

### Attachments

Attachment A: Metadata Elements Using HL7 CDA R2 - TABLE

Attachment B: Metadata Elements Using HL7 CDA R2 - XML

## Attachment A: Metadata Elements Using HL7 CDA R2 - TABLE

Metadata Element	CDA R2 Example	Extensions
Envelope	<pre>&lt;?xml version="1.0" encoding="UTF-8"?&gt; &lt;ClinicalDocument xmlns="urn:hl7-org:v3"&gt;</pre>	
Provenance - TDE ID	<pre>&lt;id extension="http://stelsewhere.com/id/12345" assigningAuthority="St. Elsewhere Hospital"/&gt;</pre>	
Privacy - Content Data Type	<pre>&lt;code code="34788-0" displayName="Psychiatric Consult note" codeSystemName="LOINC"/&gt;</pre>	
Provenance - Timestamp	<pre>&lt;effectiveTime value="20011217093047"/&gt;</pre>	
Privacy - Content Sensitivity	<pre>&lt;confidentialityCode code="PSY"/&gt;</pre>	The Power Team discussed extending the current set of allowable codes
Boilerplate	<pre>&lt;recordTarget&gt; &lt;patientRole&gt;</pre>	
Patient ID - ID	<pre>&lt;id extension="1234567" root="http://www.nh.gov/safety/divisions/dmv"/&gt;</pre>	Note that in a CDA R2 Header, the root attribute would typically be an OID
Patient ID - Address	<pre>&lt;addr use="HP"&gt; &lt;streetAddressLine&gt;1234 Main St. Apt 3&lt;/streetAddressLine&gt; &lt;city&gt;Bedford&lt;/city&gt; &lt;state&gt;MA&lt;/state&gt; &lt;postalCode&gt;01730&lt;/postalCode&gt; &lt;/addr&gt;</pre>	
Boilerplate	<pre>&lt;patient&gt;</pre>	

Metadata Element	CDA R2 Example	Extensions
Patient ID - Name	<pre> &lt;name&gt;   &lt;prefix qualifier="AC"&gt;Dr.&lt;/prefix&gt;   &lt;given&gt; John&lt;/given&gt;   &lt;given&gt;William&lt;/given&gt;   &lt;family&gt;Smith&lt;/family&gt;   &lt;displayName&gt;Dr. John William   Smith&lt;/displayName&gt; &lt;/name&gt; </pre>	<p>Note that displayName is not part of the CDA R2 Header and is being suggested by the Power Team</p>
Patient ID - DOB	<pre> &lt;birthTime value="19600427"/&gt; </pre>	
Boilerplate	<pre> &lt;/patient&gt; &lt;/patientRole&gt; &lt;/recordTarget&gt; </pre>	
Boilerplate	<pre> &lt;author&gt;   &lt;assignedAuthor&gt;     &lt;assignedPerson&gt;       &lt;providerDirectoryEntry         href="http://providerdirectory.org/1234"/&gt;         &lt;name&gt;           &lt;family&gt;Smith&lt;/family&gt;           &lt;given&gt;John&lt;/given&gt;           &lt;prefix&gt;Dr.&lt;/prefix&gt;         &lt;/name&gt;       &lt;/assignedPerson&gt;     &lt;/author&gt;   &lt;/assignedAuthor&gt; </pre>	<p>Note that providerDirectoryEntry is not part of the CDA R2 Header and is being suggested by the Power Team</p>
Provenance - Actor	<pre> &lt;representedOrganization&gt;   &lt;id extension="http://stelsewhere.com/"     assigningAuthority="St. Elsewhere Hospital"/&gt;   &lt;name&gt;St. Elsewhere Hospital&lt;/name&gt;   &lt;telecom use="1-800-555-1234"/&gt; &lt;/representedOrganization&gt; </pre>	
Boilerplate	<pre> &lt;/assignedAuthor&gt; </pre>	

<b>Metadata Element</b>	<b>CDA R2 Example</b>	<b>Extensions</b>
	</author>	
<b>Boilerplate</b>	<pre>&lt;custodian&gt;   &lt;assignedCustodian&gt;     &lt;representedCustodianOrganization&gt;</pre>	
<b>Privacy - Policy Pointer</b>	<pre>&lt;id extension="http://policy.example.org/9876543" root="policy_pointer_oid"/&gt;</pre>	<p>Note that this particular representation of a policy pointer was not directly discussed by the Power Team</p>
<b>Boilerplate</b>	<pre>&lt;/representedCustodianOrganization&gt; &lt;/assignedCustodian&gt; &lt;/custodian&gt;</pre>	
<b>Envelope</b>	</ClinicalDocument>	



## Attachment B: Metadata Elements Using HL7 CDA R2 - XML

```
<?xml version="1.0" encoding="UTF-8" ?>
<ClinicalDocument xmlns="urn:hl7-org:v3">
  <!-- Provenance - TDE Id -->
  <id extension="http://stelsewhere.com/id/12345" assigningAuthority="St. Elsewhere Hospital" />
  <!-- Privacy - Content Data Type -->
  <code code="34788-0" displayName="Psychiatric Consult note" codesystemName="LOINC" />
  <!-- Provenance - Timestamp -->
  <effectiveTime value="20011217093047" />
  <!-- Privacy - Content Sensitivity -->
  <confidentialityCode code="PSY" />
  <recordTarget>
    <patientRole>
      <!-- Patient Id - Note that in a CDA R2 Header, the root attribute would
      typically be an OID -->
      <id extension="1234567" root="http://www.nh.gov/safety/divisions/dmv/" />
      <!-- Patient Address -->
      <addr use="HP">
        <streetAddressLine>1234 Main St. Apt 3</streetAddressLine>
        <city>Bedford</city>
        <state>MA</state>
        <postalCode>01730</postalCode>
      </addr>
    </patient>
    <!-- Patient Name -->
    <name>
      <prefix qualifier="AC">Dr.</prefix>
```

```

<given>John</given>
<given>William</given>
<family>Smith</family>
<!-- Note that displayName is not part of the CDA R2 Header
and is being suggested by the Power Team -->
<displayName>Dr. John William Smith</displayName>
</name>
<!-- Patient Date of Birth -->
<birthTime value="19600427" />
</patient>
</patientRole>
</recordTarget>
<author>
<assignedAuthor>
<!-- Provenance - Actor -->
<assignedPerson>
<!-- Note that providerDirectoryEntry is not part of the CDA R2 Header
and is being suggested by the Power Team -->
<providerDirectoryEntry href="http://providerdirectory.org/1234" />
<name>
<family>Smith</family>
<given>John</given>
<prefix>Dr.</prefix>
</name>
</assignedPerson>
<!-- Provenance - Affiliation -->
<representedOrganization>
<id extension="http://stelsewhere.com/" assigningAuthority="St. Elsewhere Hospital" />
<name>St. Elsewhere Hospital</name>
<telecom use="1-800-555-1234" />
</representedOrganization>

```

```
</assignedAuthor>
</author>
<!-- Privacy - Policy Pointer -->
<custodian>
  <assignedCustodian>
    <representedCustodianOrganization>
      <id extension="http://policy.example.org/9876543" root="policy_pointer_oid" />
    </representedCustodianOrganization>
  </assignedCustodian>
</custodian>
</ClinicalDocument>
```