CDA Source of Information

Guidelines and Strategy

September 28, 2011

Table of Contents

Revision History

Revision Number	Revision Date	Summary of Changes	Author
0.1	July 11, 2011	Created initial draft.	Michael LaRocca, Ready Computing
0.2	July 19, 2011	Incorporated workgroup feedback: Made nullFlavor usage consistent throughout XML examples.	Michael LaRocca, Ready Computing
0.3	August 1, 2011	Incorporated workgroup feedback: Distinguish more clearly between "aggregate" and "non-aggregate" CDA content creators.	Michael LaRocca, Ready Computing
0.4	August 5, 2011	Added California State as an example of a "non-aggregate CDA" content creator. Incorporated workgroup feedback: Clarify exchange of CDA instances in an XDS environment.	Michael LaRocca, Ready Computing
0.5	August 9, 2011	Added Massachusetts as a "non-aggregate CDA" content creator.	Michael LaRocca, Ready Computing
1.0	September 28, 2011	Incorporated ONC feedback and prepared final copy for submission.	Michael LaRocca, Ready Computing

Introduction and Overview

Background and Document Purpose

As part of the Multi-State/Multi-Vendor Electronic Health Record (EHR)/Health Information Exchange (HIE) Interoperability Workgroup centered on CDA and CCD usage, discussions among workgroup members revealed inconsistencies in how "source of information" is represented in CDA. It can be challenging representing this information consistently for a CDA that comes from a single source. In cases where a "data aggregation" service is used to create a single CDA patient summary from several underlying data providers, such as those used in regional and national exchanges, reporting the "source of information" becomes an even greater challenge.

This document proposes a strategy for expressing this information in cross-community projects that exchange CDA-based content[1]. While this topic may be relevant to any healthcare IT professionals interested in learning about CDA-based exchange, the primary audience targeted by this document includes:

- Technical committee members within the Office of the National Coordinator focused on harmonizing CDA utilization across HIE implementations
- Technical management staff representing a Statewide HIE
- Technical management staff representing a community participating within a Statewide HIE
- Technical Management Staff representing a hospital or hospital chain participating in a Statewide HIE
- Developers of clinical systems that create content for use within a HIE
- Developers of integration frameworks that connect systems and create content for use within an HIE

Stakeholders Represented in this Document

Stakeholders within HIE projects around the world come in many shapes and sizes, and each has its own nuances and special circumstances to consider. As this document focuses on the complexities of CDA exchange when many organizations are involved – particularly when documents are exchanged within and across States – three examples were selected: New York State, California State, and the Commonwealth of Massachusetts.

Through these three examples, we see different, yet sound, approaches to CDA interoperability, with a sufficient level of complexity representative of real-life projects around the United States and the world.

New York State

New York's statewide health information exchange is managed by the New York eHealth Collaborative (NYeC) in partnership with the New York State Department of Health (NYS DOH) and other key stakeholders focused on the advancement of health information technology within a coordinated care setting. NYeC's key activities in support of this goal of coordinated care include:

- Collaboration: Through its Statewide Collaborative Process (SCP), stakeholders within the state develop common policies and procedures, standards, technical approaches and services for New York's health information infrastructure.

- Statewide Health Information Network for New York (SHIN-NY): In partnership with the NYS DOH and other key constituents, NYeC is working to develop a statewide network of health information technology to providers, allowing providers to share patient health information in a timely and secure manner.

- Technical and Adoption Services: As a federally-designated Regional Extension Center (REC) for New York (outside of New York City), NYeC REC provides tailored and personal technical and adoption services to providers through the EHR adoption process. Some of the services provided include, but are not limited to, consultative services to choose the right EHR software and hardware, discounted pricing

[1] While the intention of this document is to specifically address how CDA-based content is represented, it should be made clear that in no way is this document meant to preclude non-CDA content from being shared.

and terms for preferred vendors to enable cost effective EHR purchase and implementation, and skilled project management services to oversee the adopting and implementation process.

- Education: NYeC is actively developing its outreach program for educating the health care community on the State's vision and solution for supporting collaborative care through the use of securely-shared information.

New York State was selected for this document for two primary reasons:

1. It has a very diverse collection of participants within its statewide network including community-wide HIEs, large Integrated Delivery Networks (IDNs), hospitals, clinics, group practices, laboratories, pharmacy exchanges, Medicaid, and more.
2. It plans to offer a statewide patient summary, representing the aggregate of its stakeholders' parts in a single CDA instance.

California State

California's statewide health information exchange is led by the California eHealth Initiative, a group of organizations committed to adopting health information technology (HIT) and electronic health in California. The Initiative's member organizations share a mission consistent with ARRA/HITECH, and collaborate to support healthier people and communities. California's key goals include:

- To ensure patients have safe, secure access to their personal health information and the ability to share that information with others involved in their care,
- To engage in an open, inclusive, collaborative, public-private process that supports widespread electronic health records (EHR) adoption and a robust, sustainable statewide health information exchange and technology infrastructure,
- To improve health care outcomes and reduce costs,
- To maximize California stakeholders' access to critical ARRA stimulus funds,
- To integrate and synchronize the planning and implementation of health information exchange (HIE), health information technology (HIT), telehealth, telemedicine, and provider incentive program components of the federal stimulus act,
- To ensure accountability in the expenditure of public funds,
- To improve public health through stronger health program integration, bio-surveillance, and emergency response capabilities

California State was selected for this document for two primary reasons:

1. Like New York, it too has a very diverse collection of participants within its statewide network.
2. Unlike New York, it does not plan to offer a statewide patient summary, and instead plans to promote organization-specific CDA instances.

The Massachusetts Commonwealth

Massachusetts' statewide health information exchange is managed by the Massachusetts eHealth Institute (MeHI), a division of the Massachusetts Technology Collaborative. MeHI is responsible for advancing the dissemination of health information technology across the Commonwealth, including the deployment of electronic health records systems in all health care provider settings that are networked through a statewide health information exchange.

The Massachusetts Commonwealth was selected for this document because it presented an interesting example of a statewide HIE that was capable of generating an aggregate CDA patient summary, but elected not to, effectively aligning the State's strategy with that of California. For this reason, when the document refers to the example of California, the reader should also recognize that the guidelines prescribed within that example also apply to Massachusetts.

Other Stakeholders

While the document focuses on large Statewide HIE instances, it must be noted that each of these States is itself comprised of many other stakeholders including community-wide HIEs, IDNs and hospitals, group practices, clinics, laboratories, pharmacy exchanges, and other organizations. These other organizations play a very significant role in the examples, and will be introduced later in the document.

Benefits of the Proposed Approach and Value Proposition

The approach proposed in this document gives us a consistent mechanism for answering certain key questions pertaining to information sources in every CDA instance that's generated and exchanged. The questions vary slightly in aggregate CDAs (as created by New York) and non-aggregate CDAs (as created by California or Massachusetts):

1. In cases where an aggregate CDA is exchanged:
    a. What regional exchange (national, state, community, or other) created the CDA summary?
    b. Which qualified organizations[1] and sub-organizations within that regional exchange contributed content to the CDA summary?
    c. Through which clinical software solution was the CDA content reported?
    d. Under what single global patient identifier (such as a Master Patient Identifier, or an "affinity domain[2]" identifier) was the CDA content reported?
    e. Under what organization-assigned Medical Record Numbers (MRNs) was the CDA content reported?
    f. Did a patient, payor, or provider contribute the content?

2. In cases where a non-aggregate, organization-specific, CDA is exchanged:
    a. Which organization created the CDA summary?
    b. Through which clinical software solution was the CDA content reported?
    c. Under what single organization-assigned patient identifier was the CDA content reported?
    d. Did a patient, payor, or provider contribute the content?

If as a community we were to adopt the practices prescribed in this document, we would harmonize what is today a somewhat disjointed set of approaches, and dramatically increase the quality of data exchanged between systems and organizations.

Relationship to ONC's Advance Notice of Proposed Rulemaking (ANPRM) on Metadata Standards

On August 9th, 2011, ONC released an ANPRM titled *Metadata Standards to Support Nationwide Electronic Health Information Exchange*. In this ANPRM, ONC solicited public comments (up to the closing date of September 23rd, 2011) on the metadata standards recommended to ONC by the HIT Standards Committee.

Within this ANPRM, the term "provenance metadata" is used to refer to additional content placed in the CDA header to help consumers of that document make determinations about whether the information contained within the CDA instance could be trusted. The proposed metadata includes a tagged data element (TDE) identifier, a timestamp, the actor, the actor's affiliation, and the actor's digital certificate, which when combined, provide answers to the "who, what, where, and when" questions a document consumer might ask.

The ANPRM provides an excellent mechanism for ensuring document integrity and non-repudiation, but it doesn't address the problems raised in this document that arise when many data sources coordinate to produce a longitudinal or summary representation of a patient's data in CDA. This document proposes a strategy that compliments the ANPRM to define an even more usable and trusted CDA.

[1] The term "qualified organization" will be defined later.
[2] The term "affinity domain" will be defined later.

Key Definitions

In this document, we'll refer to certain terms and concepts that should be explicitly defined so the reader understands the intention of their use in this proposal.

Author

The CDA-R2 definition for "author" is as follows:

> "Represents the humans and/or machines that authored the document."

This proposal maintains the same definition, and uses the author for two purposes:

1. To indicate the primary document author, conforming to the *HITSP C83 Information Source* entry module
2. To indicate the clinical software – which depending on context may be an HIE framework, an EMR, a Lab System, a Pharmacy Hub, or other solution – through which the content was reported.

Informant

The CDA-R2 definition for "informant" is as follows:

> "An informant (or source of information) is a person that provides relevant information, such as the parent of a comatose patient who describes the patient's behavior prior to the onset of coma."

In this proposal, we extend this definition to include organizations, in addition to people, to accommodate cases where an organization (such as the North Shore Long Island Jewish Health System) was the source of information for content contained within the CDA generated for a region (such as New York State).

Additionally, this proposal makes use of another extension, defined by the Structured Documents Technical Committee that, where relevant, allows an informant to indicate the local patient identifier associated with the supplied information.

Qualified Organization and Sub-Organization

Recent thinking on HIE has coined the term "qualified organization" to refer to any data-providing organization – a small clinical practice, a hospital, a hospital chain or IDN, a community, a region, or even a nation – that's authorized to connect to an HIE and export patient information for use by other trusted sharing partners.

What's interesting is that some of these qualified organizations – such as RHIOs/HIEs or IDNs/Hospital Chains – are themselves comprised of other organizations, or "sub-organizations". In such cases, there's an implied hierarchy of qualified organizations and sub-organizations that would need to be represented in the generated CDA summary as "sources of information".

In the context of New York State, for example, a qualified organization might be the Long Island Patient Information Exchange (LIPIX), which is comprised of sub-organizations such as Catholic Health Services, Nassau Health Care Corporation, and South Nassau Communities Hospital. Or, also in New York, the State Medicaid system itself might be a qualified organization. In these examples, the qualified organizations connect directly to the State's network while the sub-organizations connect to the State network only by way of the qualified organization.

CDA Content Creator and Content Consumer

The terms "content creator" and "content consumer" refer to systems that produce content and systems that interpret that content, respectively. These terms are used often in IHE actor definitions, such as those seen in the XPHR, MS, and XD-LAB integration profiles.

Medical Record Number and Global Patient Identifier

Identifying patients is one of the single biggest challenges in any HIE project as it's typical for a patient to have multiple identifiers within any single institution they visit, or across institutions within a community. The term "medical record number" refers to the local identifier assigned to the patient when he or she registers during a visit with his or her healthcare provider. The term "global patient identifier" refers to the master identifier under which the patient's various local identifiers are linked.

Affinity Domain

Related to the topic of patient identity is the concept of an "affinity domain". An affinity domain refers to an organization or group of cooperative organizations that – among many things – make use of a single shared Master Patient Index (MPI) solution, called a PIX Manager in IHE-speak. The implication within an affinity domain with regard to patient identity is that all data source systems within the domain agree to link their local patient identifiers under a common global patient identifier assigned by a common PIX Manager. This concept is key for understanding the relative scope of a CDA, and how a content creator may or may not create an aggregated CDA.

Fundamentals of CDA Content Creation

Before proposing how to deal with the "source of information" in CDA, it's important to briefly address a few issues that will become the foundation of what's later discussed.

As alluded to earlier, there are several kinds of CDA content creators that will be sharing clinical documents with authorized content consumer partners. At a high level, CDA content creators differ in three ways:

1. What is the size and breadth of the content creator (i.e., a State, Community, IDN, Hospital, or other)?
2. What type of content will the content creator share (i.e., a Patient Summary, Discharge Summary, Referral Summary, Lab Report, or other)?
3. Can the content creator identify the patient through a single master patient identifier?

Size and Breadth of Content Creators

A CDA document may vary quite a bit in scope from one source to the next; for example, a CDA may contain content from a single system, a single organization, a community of institutions, a State with a variety of "qualified organizations", or a Nation of various States.

When discussing the idea of a single "aggregate CDA" instance, we'll use New York State as an example. When discussing the idea of "non-aggregate CDA" instances, which are essentially organization-specific documents, we'll use California State as an example.

In both cases, we'll show how the use of CDA's <author> and <informant> tags, as well as the <representedOrganization> that appears within these tags, allows us to represent the source of information for virtually any content creator, large or small, single-organization or multi-organization.

Flavors of CDA

CDA comes in many flavors such as CCD, C32, C37, C48, XPHR, MS, XD-LAB, and others. Regardless of CDA flavor, the challenge of representing the information source effectively still remains. For example, in a C32 document generated by the Rochester RHIO in New York, how can we effectively represent that 2 of the 5 medications originated from Medina Memorial Hospital while the 3 other medications on the same list originated from FF Thompson Hospital and Affiliates?

The proposal made in this document addresses this issue, and is fully compatible with any instance of CDA. CDA generate according to these guidelines have also been proven to pass all publicly-recognized instance validation tools, such as those offered by the National Institute of Standards and Technology (NIST).

The Effect of Affinity Domains on the CDA's Scope

Patient identity in a dynamically-generated CDA can be a real challenge depending on whether or not the content creator can identify the patient through a single master patient identifier or not. IHE gives us an important concept – the *affinity domain* – that has great relevance to this topic.

In general, some content creators generate a single aggregate CDA based on many data sources, others generate a collection of CDAs, each of which represents content from a single data source, and others utilize a combination of both techniques.

Consider the following examples:

1. New York State, where a centralized MPI is planned, may generate a Statewide aggregate CDA, as well as an organization-specific CDA for each underlying data source. When a user outside of New York asks the basic question "what data is available for this patient", he or she will see a list of documents that includes a New York State aggregate CDA, as well as the collection of organization-specific CDAs that would comprise the aggregate CDA.
2. Massachusetts, where a centralized MPI is planned but elects not to generate a Statewide aggregate CDA, will generate a collection of organization-specific CDAs, one per underlying data source. When a

user outside of Massachusetts asks the basic question "what data is available for this patient", he or she will only see the list of organization-specific CDAs.

3. California, where a centralized MPI is not planned, may generate a collection of organization-specific CDAs, one per underlying data source. When a user outside of California asks the basic question "what data is available for this patient", he or she will only see the list of organization-specific CDAs.

In these examples, we see that both New York and Massachusetts are capable of generating an aggregate CDA – although Massachusetts chooses not to – because they not only represent a collection of underlying affinity domains, but also themselves manage a single, higher-level affinity domain through a centralized MPI. On the other hand, California at present is not capable of generating an aggregate CDA because it only represents the underlying collection of affinity domains.

Unification of Patient Identifiers in CDA and XDS

Cross-Enterprise Document Sharing (XDS) is an IHE-defined integration profile that defines how systems should share content with one another. In principle, a document such as a CDA instance is "wrapped" with metadata defined by XDS which describes certain characteristics about the document, such as the patient it pertains to, the document's creator, the type of document, and other information.

One of the fundamental rules of XDS is that documents are "registered" under the patient's global identifier; that is, the patient's master identifier as issued by an assigning authority within the affinity domain. In cases where XDS is used to supply a CDA document, this proposal maintains that both the XDS and CDA headers ***must*** use the same global patient identifier.

Representing "Source of Information" in an Aggregate CDA

In this section, we'll focus on the "source of information" as it pertains to an aggregate CDA.

Organizations and Sub-Organizations as Information Sources: The New York State Example

New York serves as a good example of a State that would likely offer an aggregate CDA, in addition to non-aggregate CDAs and other documents. New York is capable of offering the aggregate CDA because its plans include the use of a centralized master patient index.

If we were to imagine a scenario where New York State exported a single CDA patient summary in response to a request from a partnering State, say Colorado, we would see many examples of the information source hierarchy we described earlier. The following table illustrates how this hierarchy could be structured[1]:

Primary Document Author	Qualified Organizations	Qualified Sub-Organizations
New York State	Albany Medical Center	
	Brooklyn Health Information Exchange (BHIX)	Lutheran Medical Center
		Maimonides Medical Center
		Sephardic Nursing Center
	Health Information Exchange of New York (HIXNY)	Columbia Memorial Hospital
		St. Peter's HealthCare Services
		Seton Health
	Greater Rochester RHIO	Alexander Medical Group
		FF Thompson Hospital and Affiliates
		Medina Memorial Hospital
	Interboro RHIO	Bellevue Hospital Center
		Coney Island Hospital
		Woodhull Medical Center
	Long Island Patient Information Exchange (LIPIX)	Catholic Health Services
		Nassau Health Care Corporation
		South Nassau Communities Hospital
	New York State Medicaid	
	North Shore Long Island Jewish Health System	
	Surescripts	
	Visiting Nurse Service of New York	

In "CDA speak", using this example New York State would act as the "primary document author" and each qualified organization and sub-organization would appear as "document informants" within the appropriate regions of the CDA.

"Levels" of Organizations

It's probable that data-contributing sources could span more levels than the two presented here – qualified organizations and sub-organizations. While situations such as this do exist, we have to be careful to work within CDA's basic constraints of supporting two levels through the <representedOrganization> and <wholeOrganization> structures.

This proposal takes the position that two most-significant organizational levels, which must be represented in the CDA, are (1) the sub-organization that actually produced the content under a local medical record number and (2) the qualified organization that is formally registered with the State as a data-providing source.

[1] Note, ***the organizations listed here are for example only***, and do not necessarily represent the actual directory of New York State participants.

Managing Sub-Organizations that are Members in Multiple Qualified Organizations

In some cases, a sub-organization may be a member of multiple qualified organizations; for example, if Bellevue Hospital Center appeared as a sub-organization to Interboro RHIO, and at the same time appeared as a sub-organization to NYC Health and Hospitals Corporation (if it were to later connect to the State network directly as a qualified organization), what should happen?

While the suggested framework of qualified organizations and sub-organizations would correctly represent this case, New York's aggregate CDA would contain duplicate information unless one of the following occurred:

1. The sub-organization, in this case Bellevue Hospital Center, would agree to contribute content through one and only one whole organization.
2. The State's data aggregation capabilities would include the ability to detect duplication and suppress extraneous copies in accordance with guidelines established with the sub-organization.

Primary Document Author

The primary document author represents the "highest level" entity that supplied the CDA summary, and conforms to the *HITSP C83 Information Source* entry module.

Representing the document's primary author is accomplished through use of the <author> element in the CDA header[1], as follows[2]:

```
<ClinicalDocument  xmlns="urn:hl7-org:v3">
  …
  <author typeCode="AUT">
    <assignedAuthor classCode="ASSIGNED">
      <id nullFlavor="NI"/>
      <addr nullFlavor="NI"/>
      <telecom nullFlavor="NI"/>
      <assignedPerson><name nullFlavor="NI"/></assignedPerson>
      <representedOrganization>
        <id root="1.2.3.4.5" extension="NYeC"/>
        <name>New York eHealth Collaborative</name>
        <telecom use="WP" value="tel:(212) 562-0100"/>
        <addr use="WP">
          <streetAddressLine>220 Church Street, 5th Floor</streetAddressLine>
          <city>New York</city>
          <state>NY</state>
          <postalCode>10013-2988</postalCode>
          <country>USA</country>
        </addr>
      </representedOrganization>
    </assignedAuthor>
  </author>
  …
</ClinicalDocument>
```

As shown here, the primary document author's information is contained in the <representedOrganization> element.

Document Informants

Document informants refer to the qualified organizations and sub-organizations that contribute a particular CDA coded entry to a patient's aggregate CDA summary. For example, when New York State reports a

---

[1] Note, while the <author> can appear in both the CDA header and body to reflect the default and overridden document authors, respectively, this proposal utilizes <author> strictly in the CDA header and does not discuss areas where overridden authors may apply.

[2] Note, this example, as with all CDA examples in this document, is a CDA fragment and targets the primary concept discussed here. The example doesn't necessarily represent all fields required by the CDA-R2 schema and related schema extensions and schematrons. Additionally, the primary document author has no human attached to it in this context (something that's been raised before in the HL7 structured documents technical workgroup), and must be supplied as null values or instead pulled from a static configuration file.

patient's allergy list, Albany Medical Center might have contributed one CDA <entry> while another <entry> in the same allergy section might have been contributed by the Long Island Patient Information Exchange (LIPIX).

This proposal takes the position that the primary document author will also serve as the primary document informant, and that each coded CDA entry may optionally override the default information with the qualified organization (and if relevant, sub-organization) that contributed that particular entry. For example:

```
<ClinicalDocument xmlns="urn:hl7-org:v3">
  …
  <informant>
    <assignedEntity>
      <id nullFlavor="NI"/>
      <addr nullFlavor="NI"/>
      <telecom nullFlavor="NI"/>
      <assignedPerson><name nullFlavor="NI"/></assignedPerson>
      <representedOrganization>
        <id root="1.2.3.4.5" extension="NYeC"/>
        <name>New York eHealth Collaborative</name>
        <telecom use="WP" value="tel:(212) 562-0100"/>
        <addr use="WP">
          <streetAddressLine>220 Church Street, 5th Floor</streetAddressLine>
          <city>New York</city>
          <state>NY</state>
          <postalCode>10013-2988</postalCode>
          <country>USA</country>
        </addr>
      </representedOrganization>
    </assignedEntity>
  </informant>
  …
  <component>
    <structuredBody>
      <component>
        <section>
          <templateId root="2.16.840.1.113883.3.88.11.83.102"/>
          <templateId root="2.16.840.1.113883.10.20.1.2"/>
          <templateId root="1.3.6.1.4.1.19376.1.5.3.1.3.13"/>
          <code code="48765-2" displayName="Allergies, Adverse Reactions, Alerts" codeSystem="2.16.840.1.113883.6.1"
codeSystemName="LOINC"/>
          <title>Allergies, Adverse Reactions, Alerts</title>
          <text><!-- Allergy Narrative Here --></text>
          <entry typeCode="DRIV">
            <act classCode="ACT" moodCode="EVN">
              …
              <informant>
                <assignedEntity>
                  <id nullFlavor="NI"/>
                  <addr nullFlavor="NI"/>
                  <telecom nullFlavor="NI"/>
                  <assignedPerson><name nullFlavor="NI"/></assignedPerson>
                  <representedOrganization>
                    <id root="6.7.8.9.10" extension="AMC"/>
                    <name>Albany Medical Center</name>
                    <telecom nullFlavor="NI"/>
                    <addr use="WP">
                      <streetAddressLine>43 New Scotland Avenue</streetAddressLine>
                      <city>Albany</city>
                      <state>NY</state>
                      <postalCode>12208</postalCode>
                      <country>USA</country>
                    </addr>
                  </representedOrganization>
                </assignedEntity>
              </informant>
              …
            </act>
          </entry>
```

```
        <entry typeCode="DRIV">
          <act classCode="ACT" moodCode="EVN">
            ...
            <informant>
              <assignedEntity>
                <id nullFlavor="NI"/>
                <addr nullFlavor="NI"/>
                <telecom nullFlavor="NI"/>
                <assignedPerson><name nullFlavor="NI"/></assignedPerson>
                <representedOrganization>
                  <id root="3.9.2.1.9" extension="CHS"/>
                  <name>Catholic Health Services</name>
                  <telecom use="WP" value="tel: (516) 705-3700"/>
                  <addr use="WP">
                    <streetAddressLine>992 N. Village Avenue</streetAddressLine>
                    <city>Rockville Centre</city>
                    <state>NY</state>
                    <postalCode>11570</postalCode>
                    <country>USA</country>
                  </addr>
                  <asOrganizationPartOf>
                    <effectiveTime nullFlavor="NI"/>
                    <wholeOrganization>
                      <id root="2.16.840.1.113883.3.176" extension="LIPIX"/>
                      <name>Long Island Patient Information Exchange</name>
                      <telecom use="WP" value="tel:(212) 562-0100"/>
                      <addr use="WP">
                        <streetAddressLine>347 W 36th Street, Suite 201</streetAddressLine>
                        <city>New York</city>
                        <state>NY</state>
                        <postalCode>10018</postalCode>
                        <country>USA</country>
                      </addr>
                    </wholeOrganization>
                  </asOrganizationPartOf>
                </representedOrganization>
              </assignedEntity>
            </informant>
            ...
          </act>
        </entry>
      </section>
    </component>
  </structuredBody>
 </component>
</ClinicalDocument>
```

The above example illustrates how both the default and overridden <informant> elements are used. In the process, it demonstrates two important concepts:

1. In cases where a qualified organization contributes content to the aggregate CDA, the qualified organization will be expressed in the informant's <representedOrganization> element.
2. In cases where a qualified sub-organization contributes content to the aggregate CDA, the sub-organization will be expressed in the informant's <representedOrganization> element, and the associated organization will be expressed in asOrganizationPartOf's <wholeOrganization> element.

Clinical Software Solutions

Representing which clinical software solution was used to produce a particular document, or region of a document, is also a desired goal. This proposal comes from the thinking that, within the context of an aggregate CDA, there are two levels at which such clinical software solutions are used:

1. At the regional level, typically data aggregation software used to generate the CDA
2. At the organizational level, in which case two uses are envisioned:
    a. Data aggregation software used by a RHIO acting as a qualified organization
    b. EMR, PHR, or other software used by interactive users at a qualified organization

Regional Data Aggregation Software

Expressing which software was used to generate the CDA at the regional level is done through used of a header-level <author> structure[1], as follows:

```
<ClinicalDocument  xmlns="urn:hl7-org:v3">
  …
  <author typeCode="AUT">
    <assignedAuthor classCode="ASSIGNED">
      <id nullFlavor="NI"/>
      <addr nullFlavor="NI"/>
      <telecom nullFlavor="NI"/>
      <assignedAuthoringDevice>
        <softwareName>Aggregation Software Name (such as InterSystems HealthShare or Axolotl Elysium or other)</softwareName>
      </assignedAuthoringDevice>
      <representedOrganization>
        <id root="1.2.3.4.5" extension="NYeC"/>
        <name>New York eHealth Collaborative</name>
        <telecom use="WP" value="tel:(212) 562-0100"/>
        <addr use="WP">
          <streetAddressLine>220 Church Street, 5th Floor</streetAddressLine>
          <city>New York</city>
          <state>NY</state>
          <postalCode>10013-2988</postalCode>
          <country>USA</country>
        </addr>
      </representedOrganization>
    </assignedAuthor>
  </author>
  …
</ClinicalDocument>
```

Organizational Data Aggregation or Interactive Software

Expressing which software was used by a qualified organization or sub-organization is done by an entry-level <author> structure, as follows:

```
<ClinicalDocument  xmlns="urn:hl7-org:v3" >
  …
  <component>
    <structuredBody>
      <component>
        <section>
          …
          <entry typeCode="DRIV">
            …
            <author typeCode="AUT">
              <assignedAuthor classCode="ASSIGNED">
                <id nullFlavor="NI"/>
                <addr nullFlavor="NI"/>
                <telecom nullFlavor="NI"/>
                <assignedAuthoringDevice>
```

[1] Note, as mandated by the CDA-R2 schema, the software device used to author content must appear in a separate <author> structure than the one used to represent the human being that authored content. Combining the use of software devices and human authors under the same <author> structure is under consideration for CDA-R3, which is under development by the HL7 committee.

```
                    <softwareName>Aggregation Software Name (such as HealthShare or Elysium or other)</softwareName>
                  </assignedAuthoringDevice>
                  <representedOrganization>
                    <id root="6.7.8.9.10" extension="AMC"/>
                    <name>Albany Medical Center</name>
                    <telecom nullFlavor="NI"/>
                    <addr use="WP">
                      <streetAddressLine>43 New Scotland Avenue</streetAddressLine>
                      <city>Albany</city>
                      <state>NY</state>
                      <postalCode>12208</postalCode>
                      <country>USA</country>
                    </addr>
                  </representedOrganization>
                </assignedAuthor>
              </author>
              ...
            </entry>
          </section>
        </component>
      </structuredBody>
    </component>
    ...
</ClinicalDocument>
```

Patient Identity

As described earlier, this proposal takes the position that a CDA's scope is limited to an affinity domain, and that only regions with a shared Master Patient Index should generate aggregate CDA instances.

Primary Patient Identifier

In CDA, the <patientRole><id> element is a required, single-value element that contains the unique identifier for the patient represented in the clinical summary. In cases of an aggregate CDA, this patient identifier should be the patient's global patient identifier as issued by the region's assigning authority[1][2]. For example:

```
<ClinicalDocument xmlns="urn:hl7-org:v3">
  …
  <recordTarget>
    <patientRole>
      <id root="1.2.3.4.5" extension="1122334455" assigningAuthorityName="New York eHealth Collaborative"/>
    </patientRole>
  </recordTarget>
  …
</ClinicalDocument>
```

Secondary Patient Identifiers

The use of local Medical Record Numbers is important for understanding how a patient is identified by a qualified organization or sub-organization represented in the aggregate CDA document. The use of the content within CDA requires use of extensions defined by the HITSP Structured Documents Technical Committee (SDTC), and is represented within <informant> using the <sdtc:patient> structure. For example:

```
<ClinicalDocument xmlns="urn:hl7-org:v3" xmlns:sdtc="urn:hl7-org:sdtc">
  …
  <component>
    <structuredBody>
      <component>
        <section>
          …
          <entry typeCode="DRIV">
            …
            <informant>
              <assignedEntity>
                <id nullFlavor="NI"/>
                <addr nullFlavor="NI"/>
                <telecom nullFlavor="NI"/>
                <assignedPerson><name nullFlavor="NI"/></assignedPerson>
                <representedOrganization>
                  <id root="6.7.8.9.10" extension="AMC"/>
                  <name>Albany Medical Center</name>
                  <telecom nullFlavor="NI"/>
                  <addr use="WP">
                    <streetAddressLine>43 New Scotland Avenue</streetAddressLine>
                    <city>Albany</city>
                    <state>NY</state>
                    <postalCode>12208</postalCode>
                    <country>USA</country>
                  </addr>
                </representedOrganization>
                <sdtc:patient>
                  <sdtc:id root="6.7.8.9.10" extension="MRN-12345" assigningAuthorityName="Albany Medical Center"/>
                </sdtc:patient>
              </assignedEntity>
            </informant>
```

[1] In the case of State-generated CDA summaries, the implication here is that the State itself has a Master Patient Index. In the case of national CDA summaries, a similar implication exists, requiring that the nation issue a national health identifier – something we're not yet equipped to do in the United States.

[2] As described earlier, when exchanging a CDA document through XDS protocols, the primary patient identifier used in the CDA header must correspond to the global patient identifier used in the XDS header.

```
            ...
          </entry>
        </section>
      </component>
    </structuredBody>
  </component>
  ...
</ClinicalDocument>
```

In the above example, we see how <sdtc:patient> is used within the <informant>[1]. While technically <sdtc:patient> could be used in the primary document informant found in the CDA header, in practice for this purpose of identifying the local MRN, we're more likely to use this extension in the entry-specific <informant> instances.

Patient, Provider, or Payor Data Source

Clinicians examining content assembled from many data sources may trust certain data sources more than others. In particular, we need a mechanism to distinguish patient-, payor-, and provider-supplied information. Distinguishing effectively between these data source types allows a consumer of the CDA to visually represent the data differently to the user; for example, perhaps provider-supplied data is marked green, payor-supplied data blue, and patient-supplied data orange.

In CDA, the <author> structure would be most appropriate for containing the data source type, and within <author>, the <functionCode> or <assignedAuthor><code> elements potentially could be used to represent the actual type designation. This proposal suggests the use of <functionCode>, utilizing codes from the HL7 Role Class code system (OID = 2.16.840.1.113883.5.110) to represent the patient, provider, and payor concepts.

Examples of such usage of <functionCode> are as follows:

Patient Data Sources

Representing that a patient authored a particular document or coded entry would be represented as follows:

```
<author typeCode="AUT">
  <functionCode code="PAT" codeSystem="2.16.840.1.113883.5.110" displayName="Patient"/>
  ...
</author>
```

Payor Data Sources

Representing that a payor authored a particular document or coded entry would be represented as follows:

```
<author typeCode="AUT">
  <functionCode code="PAYOR" codeSystem="2.16.840.1.113883.5.110" displayName="Payor"/>
  ...
</author>
```

Provider Data Sources

Representing that a healthcare provider authored a particular document or coded entry would be represented as follows[2]:

```
<author typeCode="AUT">
  <functionCode code="PROV" codeSystem="2.16.840.1.113883.5.110" displayName="Healthcare Provider"/>
  ...
</author>
```

---

[1] Note, use of <sdtc:patient> requires that the sdtc namespace (xmlns:sdtc="urn:hl7-org:sdtc") be defined in the CDA document.

[2] Note, "PROV" should be the default code in all cases where the <functionCode> is omitted from the <author> tag; that is, the CDA consumer should assume a healthcare provider authored the CDA content if no <functionCode> explicitly designates the data source type.

Representing "Source of Information" in a Non-Aggregate CDA

In this section, we'll focus on the "source of information" as it pertains to a non-aggregate CDA.

Organizations and Sub-Organizations as Information Sources: The California State Example

In the case of California, we'll also have many qualified organizations and sub-organizations as we saw in the New York example. What's different in California is that no aggregate CDA will be offered; instead, a collection of CDAs – one per qualified organization – will be offered.

The following table lists examples of qualified organizations and sub-organizations in California[1]:

Qualified Organizations	Qualified Sub-Organizations
Alameda County Medical Center	
Cedars-Sinai Medical Center	
John Muir Health Information Exchange	Brentwood Health Center
	Diablo Cardiology Medical Group
	John Muir Medical Center
	John Muir Physicians Network
Hill Physicians Medical Group	
Sutter Health System	Alta Bates Summit Medical Center
	California Pacific Medical Center
	Eden Medical Center
Surescripts	
Visiting Nurse Association and Hospice of Southern California	

Each generated CDA's scope is limited to a qualified organization, which in some cases is a single-entity organization and in other cases is a multi-entity organization comprised of one or more sub-organizations.

Multi-entity qualified organizations would themselves produce aggregate CDAs, and therefore follow the guidelines listed in the New York State example when generating their CDA documents. The following section specifically focuses on single-entity qualified organizations – in particular, Alameda County Medical Center – and how those organizations would represent their information sources.

Primary Document Author

As described earlier, representing the document's primary author is accomplished through use of the <author> element in the CDA header, as follows:

```
<ClinicalDocument xmlns="urn:hl7-org:v3">
  …
  <author typeCode="AUT">
    <assignedAuthor classCode="ASSIGNED">
      <id nullFlavor="NI"/>
      <addr nullFlavor="NI"/>
      <telecom nullFlavor="NI"/>
      <assignedPerson><name nullFlavor="NI"/></assignedPerson>
      <representedOrganization>
        <id root="2.1.4.4.5.6.7.19.1.222" extension="ACMC"/>
        <name>Alameda County Medical Center</name>
        <telecom use="WP" value="tel:(510) 437-4800"/>
        <addr use="WP">
          <streetAddressLine>1411 East 31st Street</streetAddressLine>
          <city>Oakland</city>
          <state>CA</state>
          <postalCode>94602</postalCode>
```

[1] Note, ***the organizations listed here are for example only***, and do not necessarily represent the actual directory of California State participants.

```
            <country>USA</country>
          </addr>
        </representedOrganization>
      </assignedAuthor>
    </author>
    …
</ClinicalDocument>
```

As shown here, the primary document author's information is contained in the <representedOrganization> element.

Document Informants

As described in the New York example, this proposal takes the position that the primary document author will also serve as the primary document informant. For example:

```
<ClinicalDocument xmlns="urn:hl7-org:v3">
    …
    <informant>
      <assignedEntity>
        <id nullFlavor="NI"/>
        <addr nullFlavor="NI"/>
        <telecom nullFlavor="NI"/>
        <assignedPerson><name nullFlavor="NI"/></assignedPerson>
        <representedOrganization>
          <id root="2.1.4.4.5.6.7.19.1.222" extension="ACMC"/>
          <name>Alameda County Medical Center</name>
          <telecom use="WP" value="tel:(510) 437-4800"/>
          <addr use="WP">
            <streetAddressLine>1411 East 31st Street</streetAddressLine>
            <city>Oakland</city>
            <state>CA</state>
            <postalCode>94602</postalCode>
            <country>USA</country>
          </addr>
        </representedOrganization>
      </assignedEntity>
    </informant>
    …
</ClinicalDocument>
```

In the case of a CDA representing a single-entity organization, we typically wouldn't supply an informant within the coded entry, except perhaps to override the informant element's <assignedPerson>[1].

Clinical Software Solutions

Within the context of a single-entity, non-aggregate CDA, in many cases a single clinical software solution will have created the CDA document; this said, however, it's possible that additional clinical software solutions contributed specific regions of the document. Both cases will be described below.

Primary Clinical Software Solutions

Representing the primary clinical software solution that created the non-aggregate CDA is done using the header's <author> tag, as follows:

```
<ClinicalDocument  xmlns="urn:hl7-org:v3">
    …
    <author typeCode="AUT">
      <assignedAuthor classCode="ASSIGNED">
        <id nullFlavor="NI"/>
        <addr nullFlavor="NI"/>
        <telecom nullFlavor="NI"/>
        <assignedAuthoringDevice>
          <softwareName>Clinical Software Name (such as Allscripts Enterprise or other)</softwareName>
```

[1] Overriding the <assignedPerson> isn't specifically described in this proposal.

```
        </assignedAuthoringDevice>
        <representedOrganization>
          <id root="2.1.4.4.5.6.7.19.1.222" extension="ACMC"/>
          <name>Alameda County Medical Center</name>
          <telecom use="WP" value="tel:(510) 437-4800"/>
          <addr use="WP">
            <streetAddressLine>1411 East 31st Street</streetAddressLine>
            <city>Oakland</city>
            <state>CA</state>
            <postalCode>94602</postalCode>
            <country>USA</country>
          </addr>
        </representedOrganization>
      </assignedAuthor>
    </author>
    …
</ClinicalDocument>
```

Secondary Clinical Software Solutions

In some cases, the primary clinical solution may be including content from a secondary solution. For example, if Allscripts Enterprise imported external content created by the DTI Health Patient Portal and then created a CDA, the resulting CDA content may represent Allscripts Enterprise as the primary authoring software solution and, where relevant, the DTI Health Patient Portal as the secondary authoring software solution. In this case, the secondary authoring solution would appear in an overridden author instance within a coded entry, such as:

```
<ClinicalDocument  xmlns="urn:hl7-org:v3" >
    …
    <component>
      <structuredBody>
        <component>
          <section>
            …
            <entry typeCode="DRIV">
              …
              <author typeCode="AUT">
                <assignedAuthor classCode="ASSIGNED">
                  <id nullFlavor="NI"/>
                  <addr nullFlavor="NI"/>
                  <telecom nullFlavor="NI"/>
                  <assignedAuthoringDevice>
                    <softwareName>Clinical Software Name (DTI Health Patient Portal)</softwareName>
                  </assignedAuthoringDevice>
                  <representedOrganization>
                    <id root="2.1.4.4.5.6.7.19.1.222" extension="ACMC"/>
                    <name>Alameda County Medical Center</name>
                    <telecom use="WP" value="tel:(510) 437-4800"/>
                    <addr use="WP">
                      <streetAddressLine>1411 East 31st Street</streetAddressLine>
                      <city>Oakland</city>
                      <state>CA</state>
                      <postalCode>94602</postalCode>
                      <country>USA</country>
                    </addr>
                  </representedOrganization>
                </assignedAuthor>
              </author>
              …
            </entry>
          </section>
        </component>
      </structuredBody>
    </component>
    …
</ClinicalDocument>
```

Patient Identity

Identifying a patient within a CDA generated by single-entity qualified organization is generally done in one of two ways:

1. Through use of a Master Patient Identifier assigned to the patient by the authoring organization
2. Through use of a Medical Record Number assigned to the patient by the authoring software solution

Both of these cases will be described below.

Primary Patient Identifier

Since many organizations use a Master Patient Index to resolve multiple medical record numbers assigned to the same patient, the MPI identifier is often used as the primary patient identifier in the CDA document. On the other hand, there are contexts in which a specific system within an organization will generate a CDA, in which case that system's medical record number would instead be used as the CDA document's primary patient identifier.

In either case, the primary patient identifier is represented in <patientRole><id>[1], as follows:

```
<ClinicalDocument xmlns="urn:hl7-org:v3">
  …
  <recordTarget>
    <patientRole>
      <id root="123.1.2.355.39.1" extension="AC12245" assigningAuthorityName="Alameda County Medical Center"/>
    </patientRole>
  </recordTarget>
  …
</ClinicalDocument>
```

Secondary Patient Identifiers

In cases where an MPI identifier is used as the primary patient identifier, an organization may wish to represent which local medical record numbers within that MPI identifier's linked group were associated with which coded CDA entries. As discussed earlier, overriding the document's primary patient identifier within CDA requires use of extensions defined by the HITSP Structured Documents Technical Committee (SDTC), and is represented within <informant> using the <sdtc:patient> structure. For example:

```
<ClinicalDocument xmlns="urn:hl7-org:v3" xmlns:sdtc="urn:hl7-org:sdtc">
  …
  <component>
    <structuredBody>
      <component>
        <section>
          …
          <entry typeCode="DRIV">
            …
            <informant>
              <assignedEntity>
                <id nullFlavor="NI"/>
                <addr nullFlavor="NI"/>
                <telecom nullFlavor="NI"/>
                <assignedPerson><name nullFlavor="NI"/></assignedPerson>
                <representedOrganization>
                  <id root="6.7.8.9.10" extension="AMC"/>
                  <name>Alameda County Medical Center</name>
                  <telecom nullFlavor="NI"/>
                  <addr use="WP">
                    <streetAddressLine>1411 East 31st Street</streetAddressLine>
                    <city>Oakland</city>
                    <state>CA</state>
                    <postalCode>94602</postalCode>
                    <country>USA</country>
```

[1] As described earlier, when exchanging a CDA document through XDS protocols, the primary patient identifier used in the CDA header must correspond to the global patient identifier used in the XDS header.

```
                    </addr>
                  </representedOrganization>
                  <sdtc:patient>
                    <sdtc:id root="6.7.8.9.10" extension="MRN-12345" assigningAuthorityName="Alameda County Medical Center"/>
                  </sdtc:patient>
                </assignedEntity>
              </informant>
              …
            </entry>
          </section>
        </component>
      </structuredBody>
    </component>
    …
</ClinicalDocument>
```

Patient, Provider, or Payor Data Source

Identifying whether a patient, provider, or payor created the document is also relevant within a non-aggregate CDA. From a technical perspective, accomplishing this goal would follow the same guidelines as described in the New York example.

Proposal Summary and Conclusion

In this proposal, we've summarized an approach for representing sources of information in CDA, that if implemented, increases the value, consistency, and quality of data expressed as CDA in health information exchange projects. In concluding this document, we'd like to highlight specific areas of that proposal and our rationale for suggesting them.

Organizations and Sub-Organizations as Information Sources

Accurately representing organizations and sub-organizations as information sources within a CDA document is key to ensuring that users of that document make the best use of the content found within. In the absence of such representation, a content consumer would be unable to:

- Determine if all desired or expected content providers were included in the CDA
- Determine if any undesired or unexpected content providers were included in the CDA
- Apply special filtering of content provided by one or more particular data sources included in the CDA
- Audit the receipt of content provided by one or more particular data sources included in the CDA
- Apply any specialized terminology translations to one or more particular data sources included in the CDA
- Perform analytical activities based on one or more particular data sources included in the CDA
- Diagnose a problem specific to one or more particular data sources included in the CDA

Clinical Software Solutions

Understanding which clinical software solution or solutions contributed data to a CDA document is another important objective. Just as described in the previous section, without this information a content consumer would be unable to:

- Determine if all desired or expected source clinical systems were included in the CDA
- Determine if any undesired or unexpected source clinical systems were included in the CDA
- Apply special filtering of content provided by one or more particular source clinical systems included in the CDA
- Audit the receipt of content provided by one or more particular source clinical systems included in the CDA
- Apply any specialized terminology translations to one or more particular source clinical systems included in the CDA
- Perform analytical activities based on one or more particular source clinical systems included in the CDA
- Diagnose a problem specific to one or more particular source clinical systems included in the CDA

Patient Identity

As a single patient typically has multiple identifiers, it's often important to know which patient identifiers are associated with different regions of the CDA document. With this data present, a content consumer can:

- Derive the relationship between a global patient identifier and its linked group of local medical record numbers
- Audit the receipt of content associated with one or more patient identifiers included in the CDA
- Perform analytical activities based on one or more patient identifiers included in the CDA
- Diagnose a problem specific to one or more patient identifiers included in the CDA

Patient, Provider, or Payor Data Source

Often, users of data want to know the classification of user that contributed content to the CDA document. This allows those users to make a determination of which data element values might be more trustworthy than others. For example:

- A medication list that came from a provider source is likely to be more accurate than a medication list that came from a patient source
- Family and social histories reported by the patient are more likely to be complete than those lists that came from a provider source
- A diagnosis from a provider source is likely to be more clinically accurate than a diagnosis reported by a payor

Distinguishing between these different user classifications gives content consumers tremendous flexibility in how they "rank" different data types from different sources, and assign trust to each source's data accordingly.

Appendix A: Referenced Acronyms

The following list includes acronyms referenced by this proposal, and their definitions

Acronym	Definition
ANPRM	Advance Notice of Proposed Rulemaking
ARRA	American Recovery and Reinvestment Act
CCD	Continuity of Care Document
CDA	Clinical Document Architecture
DTI	Digital Technology International
EHR	Electronic Health Record
EMR	Electronic Medical Record
HIE	Health Information Exchange
HIT	Health Information Technology
HITECH	Health Information Technology for Economic and Clinical Health
HITSP	Health Information Technology Standards Panel
HL7	Health Level 7
IDN	Integrated Delivery Network
IHE	Integrating the Healthcare Enterprise
MeHI	Massachusetts eHealth Institute
MPI	Master Patient Identifier
MRN	Medical Record Number
MS	Medical Summary
NIST	National Institute of Standards and Technology
NYS DOH	New York State Department of Health
NYeC	New York eHealth Collaborative
OID	Object Identifier
ONC	Office of the National Coordinator
PHR	Personal Health Record
PIX	Patient Identifier Cross-Referencing
REC	Regional Extension Center
RHIO	Regional Health Information Organization
SCP	Statewide Collaborative Process
SDTC	Structured Documents Technical Committee
SHIN-NY	Statewide Health Information Network for New York
TDE	Tagged Data Element
XD-LAB	Cross-Enterprise Document Sharing of Lab Reports
XDS	Cross-Enterprise Document Sharing
XPHR	Exchange of Personal Health Record Content