S&I FRAMEWORK

Data Provenance

Functional Requirements Document: Developed in Response to the Data Provenance Task Force Recommendations

Version 1.0

May 2015

Version History

Version Number	Revision Date	Author	Description of Change
1.0	*05/29/2015*		*First version published for community review*

1 Contents

2 Introduction

During the harmonization and standards development phase of the Standards and Interoperability (S&I) Data Provenance initiative, two sub-workgroups were created to determine the information interchange (i.e., transport layer) and system (i.e., sending and receiving application) requirements as they relate to exchanging and conveying provenance. The work of the two sub-workgroups was guided by recommendations provided by the Federal Advisory Committee's (FACA) Data Provenance Task Force in its official transmittal letter (February 2015).

This functional requirements document provides an in-depth summary of the work that the Information Interchange and System Requirements sub-workgroups have done and the discussions that they have had from March-May 2015 in response to the Data Provenance Task Force recommendations. Specifically, this document outlines the functional requirements that systems must meet in order to exchange provenance information, provides a short list of the most common health IT standards that support and/or convey the provenance concepts, and addresses the privacy, security, and policy considerations that factor into the methods used to support the provenance requirements.

This document is intended to guide pilot project organizations on the functional requirements of data provenance. The sub-workgroups also developed a list of standards that are needed to adequately support the functional and data requirements of provenance. Ultimately, pilot organizations will make the final decisions on what standards are used in their scenarios and how they are implemented based on their own organizational needs. The Data Provenance Task Force offered its recommendations, but pilot organizations should not feel overly restricted or constrained in their participation by these recommendations (e.g., focusing solely on EHR to EHR exchanges). This functional requirements document, in conjunction with feedback from the pilot organizations, will help to identify gaps where the prevailing standards fail to meet the core requirements and to develop technical implementation guidance for Data Provenance.

2.1 Purpose

The purpose of the S&I Functional Requirements Document is to:

1. Provide granular-level detail of the information interchange and system requirements outlined in the Use Case
2. Document functional requirements for pilots to demonstrate feasibility, identify gaps, and provide insight into additional technical and policy considerations
3. Act as an input to the development of detailed solution artifacts (e.g. implementation guidance, updates to standards, etc.)

2.2 Audience

This document is designed for analysts and developers who require guidance on developing technical solutions consistent with the goals and requirements of the Data Provenance Initiative. Users of this guide must be familiar with the details of the standards and Implementation Guides that are referenced by regulation; this document is not intended to be a tutorial on those subjects.

2.3 Framing Questions

When clinical information is shared between providers, organizations and/or electronic health record systems, the recipient of the information must know some basic information before it can use the data to make clinical decisions.

To that end, the framing questions that guided the work of the Information Interchange sub-workgroup are “can I trust it,” and “has it been changed?”

The sub-workgroup also considered that for clinical care, if trending the data, the recipient may need to know the degree to which the information can be trusted.

3 Analysis of Use Case Functional and Data Requirements

3.1 Information Interchange Requirements

3.1.1 Framing Questions

1. To what extent is data provenance required on transport protocols used in health IT?
2. To what extent is required data provenance supported in current and proposed transportation standards, e.g., Direct, CONNECT?

3.1.2 Assumptions

In developing a set of core requirements and the corresponding payload and target standards, the Information Interchange sub-workgroup operated under a number of assumptions:

3.1.2.1 Transport

- The interchange of information is a black box, EHR to EHR exchange*; the content of the exchange is not important
- Information interchange begins once the exchange artifact is presented to an interface for movement between applications
- The transport is content neutral; the ‘thing’ doesn’t get changed and is transported intact (e.g. the black box)
- The receiver makes the decision to accept the message

**Although the Data Provenance Task Force recommended that we focus on EHR to EHR exchanges, pilot organizations should not feel overly restricted by this recommendation.*

3.1.2.2 Payload

- The information required for end-to-end routing must be present in the un-encrypted metadata

3.1.3 Out of Scope

In accordance with the FACA Data Provenance Task Force, the intermediaries included in workflows #2 and #3 of the Use Case (transmitter, assembler, and composer) have been deemed out of scope.

3.1.4 Core Requirements and Data Elements

The first task of the Information Interchange sub-workgroup was to determine the core information that a recipient of clinical data needs to know in order to make a trust decision upon reception of the data.

Below are the key questions that need to be addressed in both the transport and the payload sub-topics. Note that addressing the question does not equate to forcing the conveyance - the outcome could be "doesn't apply" at the transport level but "mandatory/required" at the payload level. In addressing the question, it may also be noted that some of this is inherent in the business agreement between two parties and doesn't have to be explicitly conveyed. (e.g., if we determine that "who is this from" is a critical transport question, that may be known through the fact that only Org A can submit on this port.)

What do I need to know?	Corresponding Data Elements	Required for Transport?	Required for Payload?	Implied or Explicit System Requirements (In the Absence of Information Interchange)
WHO is the sender? ***WHERE*** is the sender? It may be the original organization, original individual or a combination of both	Required: - Organization Name - Organization ID - Sender Location Optional: - Individual Name - Individual ID	*Required*	*Conditionally Required*	*Required/ OPTIONAL - at the discretion of the implementers* - On Behalf of (e.g., type) - Device (might come up as System Requirement activities) - Author (too complex for initial goal)
WHAT is being sent? Do we need to identify anything about the content? Message that can be wrapped and sent over content neutral transport	- Transaction, Transaction Type (CCDA, v2.x message) - Provider Directory Content Profile - FHIR Resources	*Optional*	*Required*	*Required*
HOW will system identify which request it is responding to? Request Response ID	- Query ID to respond to request (echo back original data)	*Conditional (Required in either transport or payload, not in both)*	*Conditional(Required in either transport or payload, not in both)*	*Not Required*
WHEN was the message sent?	- Timestamp	*Required*	*N/A*	*Not Required*

WHO is the intended Recipient?	• Receiver	*Required*	*Conditionally Required*	*Not Required*

3.1.5 Information Interchange Standards for Consideration

3.1.5.1 Transport

The Information Interchange sub-workgroup created a short list of target standards that will support the recommended data provenance payloads, as well as the previously defined core requirements of clinical information exchanges that include provenance information.

- Direct
- HealtheWay/CONNECT
- X12 EDI
    - X12 275 as a metadata wrapper can transport payload and can wrap content
- HL7 v2 MDM, Lab, and other HL7 V2.x messages in common use or prescribed by MU
- REST

One primary objective of the data provenance initiative is to be as agnostic as possible, and therefore, pilot organizations should try to pick a standard that can support transport between various organizations and systems, such as from an EHR to a Practice Management System or from an EHR to a payer. Furthermore, in determining which standard(s) would be best for information interchange, the sub-workgroup recommends that pilot organizations consider the implications of security aspects related to information interchange (traceability, audit, etc.) and what the impact may be on the trust decision. These considerations will be discussed in further detail in the privacy and security section. **Error! Reference source not found.**

3.1.5.2 Payload

***In general, provenance requirements are agnostic of the standards used to define the payload*.** We suggest that pilot organizations start with the recommended standards below - standards which systems are required to support under Meaningful Use regulations- and iterate on this initial exchange with other payload types.

Given the core set of requirements, Meaningful Use 3 requirements, and the recommendations from the FACA Data Provenance Task Force, the Information Interchange sub-workgroup recommends that pilot organizations employ C-CDA R2 templates as the payloads that will convey clinical information between EHR systems.

As FHIR becomes more mature, pilot organizations can begin to work with FHIR-based content and resources.

3.2 System Requirements

3.2.1 Assumptions

In developing the set of core requirements and the corresponding definition of change and data elements, the System Requirements sub-workgroup operated under the following assumptions:

- The exchange of information is point to point from EHR System to EHR System*
- Once the clinical data with provenance information attached is imported by the receiving system, the process starts over again
- **Although the Data Provenance Task Force recommended that we focus on EHR to EHR exchanges, pilot organizations should not feel overly restricted by this recommendation.*

3.2.2 Core Requirements and Data Elements

The following provenance event system requirements and data elements were taken from the system events requirements matrix, which is available here.

	Source EHR System Events				Exchange (could be a push or a pull)	Receiving EHR System Events	
	Create (Originate)	**Maintain (Retain)**	**Change (Update)**	**Exchange (Transmit)**	↔	**Import (Receive)**	**Maintain (Retain)**
Provenance Event?	YES	No	YES	No	↔	No	No

Data Element Set/ Section	Data Element
Who – Entity	Person
Who – Entity	Organization
Who – Entity	System, Device or Software
Who – Entity	Subject/Target
Who – Entity	Author
Who – Entity Roles	Enterer
Who – Entity Roles	Verifier
Who – Entity Roles	Attester
Who – Entity Roles	Performer
Who – Entity	Informant

Roles	
Who – Entity Roles	Participant
Who – Entity Roles	Viewer/Accessor /User
What	Action Taken
What	Chain of Trust Event
What	Provenance Event
When	Action Date/Time
When	Action Duration
When	Data Event
Where	Action Physical Location
Why	Action Reason
Why	Data Reason
Additional Provenance Metadata	Author Signature
Additional Provenance Metadata	System, Device or Software Signature

3.2.3 System Requirements Standards for Consideration

The System Requirements sub-workgroup has provided a short list of recommended standards that support and convey core system requirements and fulfill the needs of the proposed data elements:

- C-CDA R2 (Authorship and Individual Entries – section vs. entry level templates)
- CDISC ODM Production Version 1.3.2
- HL7 EHR Record Lifecycle Model (2008 - currently DSTU ballot)
- HL7 Implementation Guide for CDA®, Release 2: Consent Directives, Release 1
- HL7 Implementation Guide for CDA® Release 2: Data Provenance, Release 1 – US Realm (Data Provenance CDA IG - currently DSTU ballot)
- HL7 Implementation Guide: Data Segmentation for Privacy (DS4P), Release 1
- HL7 Healthcare Classification System Vocabularies
- *HL7 Implementation Guide for CDA® Release 2: Digital Signatures and Delegation of Rights, Release 1
- HL7 Record Lifecycle Event Metadata using FHIR (project underway 2014)
- HL7 Privacy and Security Architecture Framework (under development) as part of FHIM

- HL7 Security Labeling Services
- ISO/HL7 10781 EHR System Functional Model Release 2
- ISO 21089 Health Informatics: Trusted End-to-End Information Flows

** The presence of digital signatures is optional in Meaningful Use regulations, but the System Requirements SWG strongly recommends that pilot organizations/EHR systems include this information.*

If a pilot organization chooses not to utilize digital signatures, then they should have a point to point exchange or other trusted trading environment.

3.2.4 Definition of Change

Any changes to the content that affect the semantic meaning of the content (e.g. changes to values; additions or subtractions to the set of information; changes in intent that change the business use of the information; changes to the status of the artifact in the workflow) are provenance events. (It is permissible that the custodian of the record makes decisions based on those things that constitute a provenance event.)

A digitally signed object should be considered an attestation of the signer's intent. When the information in the signed object is extracted, the digital signature no longer applies to the information that was extracted.

4 Policy Implications

The Information Interchange and System Requirements sub-workgroups were tasked with considering the policy implications of provenance information, in addition to considering how the implications of privacy and security impacted the chain of trust and trust decisions that recipients of clinical information must make..

The following questions arose during the sub-workgroup conversations and should be considered by pilots as they develop their implementation guidance.

4.1 Privacy and Security Implications

4.1.1 Information Interchange Considerations

4.1.1.1 Transport

- The expectation is that the receiving system can and will comply with the privacy stipulations. Consequently, the ability to evaluate provenance metadata must be considered at the point of receipt, not at the point of consumption.
  - How should a receiving system behave if it is unable to consume the provenance as expected by the sending system?
  - If the sender holds the responsibility for determining whether the recipient is able to comply with the provenance and privacy obligations, is that inherent in the trust framework?

- Is there a need to associate provenance to the artifact itself for end to end transport?

4.1.1.2 Payload

- Security may have an impact if the payloads that are being exchanged are encrypted or externally sent.
- Is there a need to accommodate provenance associated with the payload as something necessary for end to end transport of the data (the receiver may want to know something about the sender prior to opening the exchange artifact).

4.1.2 System Requirements Considerations

- Does the downstream/receiving system have to comply with the sender's wishes (e.g. security labels)? If the recipient refuses the sender's wishes and/or rejects the sender's message, what kind of notification would the sender receive, if any?
- How do we regulate/monitor provenance as the clinical information is consumed and redistributed?

5 Appendices

5.1 Data Provenance Use Case

The final Data Provenance Use Case was published in October 2014. A PDF version of the Use Case is available here.

5.2 Data Provenance Task Force Recommendations

The Federal Advisory Committees Data Provenance Task Force shared their final recommendations in January 2015. Their formal recommendation is available for review here.

5.3 Data Provenance System Event Requirements Matrix

The System Requirements sub-workgroup employed a system event requirements matrix created by Gary Dickinson. That spreadsheet is available for download and review here.