Data Provenance Initiative

October 16, 2014

Table of Contents

List of Figures:

List of Tables:

1.0 Preface and Introduction

To fully realize the benefits of health IT, the Office of the National Coordinator for Health Information Technology (ONC), as part of the Standards and Interoperability (S&I) Framework is developing Use Cases that define the interoperability requirements for high priority health care data exchange; maximize efficiency, encourage rapid learning, and protect patients' privacy in an interoperable environment. These Use Cases address the requirements of a broad range of Communities of Interests including; patients, their significant others and family members, providers, payers, vendors, standards organizations, public health organizations, and Federal agencies.

The Data Provenance initiative was launched to establish a standardized way of capturing provenance, including provenance at time of creation, modification, and time of exchange. The Data Provenance Use Case aims to capture provenance requirements to improve trust in healthcare data and its applications.

These Use Cases describe:

- User stories which link the functional capabilities with the business goals and needs
- The operational context for the data exchange
- The stakeholders with an interest in the Use Case
- The information flows that must be supported by the data exchange
- The types of data and their specifications required in the data exchange

The Use Case is the foundation for identifying and specifying the standards required to support the data exchange and developing reference implementations and tools to ensure consistent and reliable adoption of the data exchange standards.

2.0 Initiative Overview

The term "provenance" in the context of Health IT refers to evidence and attributes describing the origin of health information as it is captured in a health system. The requirements for data provenance information must support the full lifecycle and lifespan of health IT data. As the exchange of health data increases, so does the demand to track the provenance of this data over time and with each exchange instance. Confidence in the authenticity, trust worthiness and reliability of the data being shared is fundamental to robust, privacy, safety, and security enhanced health information exchange. Truth and trust may be improved by means of a standardized way to capture and express the provenance of the data and by the expectation that systems have the ability to recognize and validate the provenance information. This in turn can lead to uses such as "chain of trust" and "chain of custody" and other business requirements/applications (for example records management, evidentiary support and clinical decision support).

A significant amount of literature has discussed the value, importance and legal necessity of provenance.

The Data Provenance Initiative aims to establish a standardized way of capturing provenance (including inbound, system generated, and outbound provenance), retaining and exchanging the provenance of health information. (Provenance Capabilities) Data provenance captures (embodies) data related to the Who, What, When, Where, Why, and Authenticity of a clinical encounter or data capture event for each type of author and supporting system.

2.1 Initiative Challenge Statement

While there are efforts to address data provenance, no existing authoritative specification, standard, or model for provenance has been commonly adopted to-date within the context of HIT. This is exemplified by the variance in how HIEs, EHRs, and PHRs currently capture, retain, and convey provenance. This variability is problematic for the interoperable exchange (system interoperation), integration, safe use and interpretation of health data.

Even current health information standards, such as the CDA, lack guidance for handling data provenance and chain of trust. Additionally, the receipt and integration of provenance information is equally variable and dependent upon system capabilities. Further challenges are presented if one system can share detailed provenance data but those receiving it cannot capture, retain and convey the same level of detail.

3.0 Use Case Scope

The scope of Data Provenance is broad and there are differing perspectives surrounding priorities and expectations for provenance capabilities. We will address the initiative goals by discussing the following questions to help us understand the broader scope and applicability of data provenance:

- When healthcare data is first created, what is the provenance information that should be captured (previously created provenance information or newly created provenance information), retained, and conveyed?
- Can a receiving system understand and trust that provenance information?
- Do we need to know who touched it along the way? If so, how do we know?
- In situations where the receiving system combines this information with data received from third parties, how do we persist the provenance from multiple sources?
- When multi-sourced data is assembled and sent to another system, how do we convey the provenance of the multiple data sources as well as for the system doing the assembly?
  - Is this considered new data?
  - What if the assembling system "cherry picks" from multiple sources, or adds some new health information of its own?
  - How do we know if the received provenance information is complete?

3.1 In Scope

1. To identify and define guidance on use of standards to facilitate provenance capabilities by specifying the following: ***
   a. Standards for the provenance (e.g. origin, source, custodian(s), FHIR resources, CDA, etc.)
   b. Supportive standards (e.g. integrity, non-repudiation, and privacy & security with respect to provenance)
   c. Vocabulary standard metadata tags for data provenance
2. Variance in granularity of data to which provenance is applied, the way provenance is encoded, and how data and its provenance are conveyed to consuming systems and users
3. To define system requirements that describe how applications capture, retain and convey data provenance and maintain association with the data
4. To ensure sufficient granularity to support chain of custody

***Leveraged from Charter

3.2 Out of Scope

1. Patient identity matching***
2. Third party mechanisms for checking patient consent and the relative merits of existing policies or regulations (such as privacy policies or jurisdictional considerations)***
3. Policy-based decisions (such as records management based policies on record retention)
4. Non-clinical data (such as environmental data)
5. Mechanisms to verify the validity of the original source data

***Leveraged from Charter

3.3 Communities of Interest (Stakeholders)

Member of Communities of Interests	Working Definition
Patient	Healthcare consumers who are recipients of health care services and products.
Health Care Provider	A person or organization that's licensed to give health care. A Provider can enter into an agreement with an insurer to provide services (e.g., Medicare, Health Plan)
Health Care Information System Vendors	Vendors which provide specific EHR/EMR solutions to providers such as software applications and software services. These suppliers may include developers, providers, resellers, integrators, operators, and others who may provide these or similar capabilities.
State HIEs	Health Information Exchange (HIE) is defined as the mobilization of healthcare information electronically across organizations within a region, community or hospital system.
Local, State and Federal Government Health Agencies	Organizations within the federal and/or state governments that deliver, regulate, or provide funding for health, healthcare and clinical or biomedical research. This also includes organizations within the federal and/or state governments that disseminate clinical guidance.
Health Organizations	Organizations whose purpose is to conduct research on healthcare and develop and evaluate quality standards.
Standards Organizations	Organizations whose purpose is to define, harmonize and integrate standards that will meet clinical and business needs for sharing information among organizations and for system interoperability.
Healthcare Payers	A third party entity that establishes indications and limitations of coverage for payments or underwrites coverage for healthcare expense.
Ancillary Health Care Services	Other healthcare vendors

Table 1: Communities of Interest

4.0 Value Statement

The Data Provenance initiative will improve the confidence in the integrity of health information from creation to exchange and integration across multipl e health information systems and between parties. Ultimately these capabilities will improve trust in healthcare data and its applications, which may include clinical care, interventions, analysis, decision making, clinical research, patient engagement and other uses.

5.0 Use Case Assumptions

1. Clinical information that already exists within the EHR, PHR and HIE system (without the use of the CDA artifact) is found at the level appropriate for the implementation
2. The original sources (intent) are valid
3. Representation of the party providing information follows standards practices and is of high quality/integrity
4. All patient-related health data/records are originated, retained, attested, accessed and conveyed for legitimate purposes, including treatment, payment and healthcare operations
5. Data provenance, if properly captured, retained, secured, managed and conveyed from the point of origination forward is presumed a necessary condition to establish a verifiable chain of trust
6. All necessary communication acknowledgements are correct

6.0 Pre-Conditions

1. Where it exists, the assembling software is the component of a system such as an EHR, PHR, and HIE that facilitates the person or device authoring the data (record, document, entry, image, etc.)
    - The type of assembly software functionality (e.g., aggregator, composer, translator, transformer, etc.), and the organization or person responsible for deploying it, are information that is captured, retained, and conveyed by the authoring or sending system to the receiving system for use in calculating reliability, authenticity, and trustworthiness
2. Systems that are involved in originating the data that is being exchanged are responsible for capturing the provenance associated with that data as specified herein
3. The assembler is able to capture and retain provenance information on predecessor artifacts
4. The newly assembled artifact will have new provenance information associated with the assembly
5. All parties involved are authorized users
6. Systems participating must be able to appropriately maintain privacy and security (e.g. confidentiality, integrity, and availability of information)

7.0 Post Conditions

1. Receiving system has incorporated source data and its associated provenance information into its database
2. Sending and receiving systems have recorded the exchange transactions in their security audit records
3. Data provenance information has been securely retained and conveyed, and remains verifiable and immutable.
4. Data provenance information has been successfully conveyed in standards-based exchange artifacts (e.g., HL7 CCDA)

5. Data provenance information remains fit for use and the purpose intended, i.e., fit for primary (clinical care, interventions and decision making) or secondary use

8.0 Actors and Roles

Actor	System	Role
*Start Point (e.g., patient, provider/provider organizations, labs)	Data Capture System (e.g., EHR, PHR)	Create Data Send Data
Transmitter (e.g., State HIEs)	Intermediary System (e.g., HIE, HISP)	Receive Data Send Data
Assembler (e.g., provider/provider organizations, State HIEs)	Data Integration System (e.g., EHR, HIE, PHR)	Receive Data Consolidate Data Send Data
End Point	Data Storage System (e.g., EHR, HIE, PHR)	Receive Data
Composer	Data Integration Composer System	Receive Data Consolidate Selected Data Send Data

*Start point and end point are points on the workflow that can be carried out by a variety of actors carrying out a variety of roles. For example, the start point may be the point of origin of the data, or it may be the start of an information interchange which includes pre-existing data from multiple sources.

Table 2: Actors and Roles

9.0 Use Case Diagram

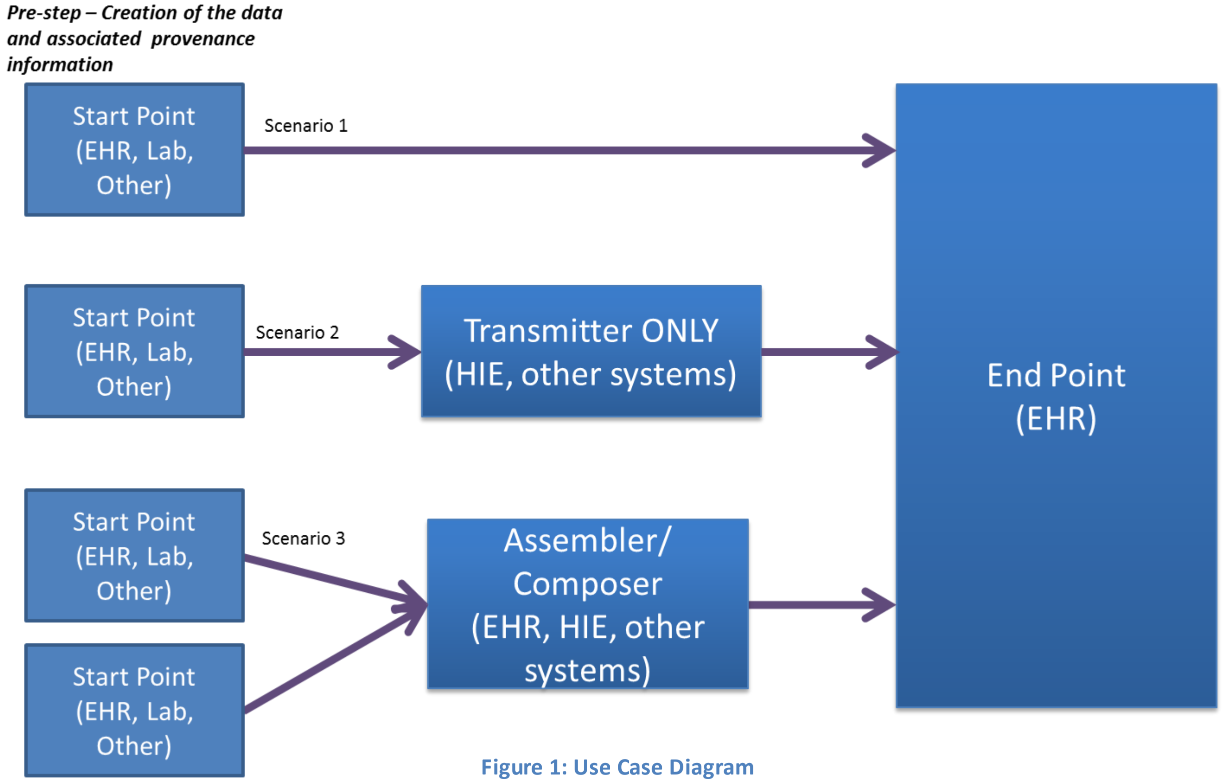
Pre-step – Creation of the data and associated provenance information
Start Point (EHR, Lab, Other)
Scenario 1
Start Point (EHR, Lab, Other)
Scenario 2
Transmitter ONLY (HIE, other systems)
Start Point (EHR, Lab, Other)
Scenario 3
Start Point (EHR, Lab, Other)
Assembler/ Composer (EHR, HIE, other systems)
End Point (EHR)

Figure 1: Use Case Diagram

10A.0 Scenario

Scenario 1: Start Point→ End Point

10A.1 User Story

User Story 1: A patient arrives at the ophthalmologist's office for her annual eye exam. The ophthalmologist conducts an eye exam and captures all of the data from that visit in his EHR. The ophthalmologist electronically sends the information back to the patient's PCP (where all data in the report sent was created by the ophthalmologist).

User Story 2: A patient has a PHR that allows them to record their daily dietary intake. The patient accesses the PHR and requests that their dietary intake for the past month be transmitted to their PCP prior to their visit next week. The patients uses a PHR to transmit the dietary record to the PCP. The PCP understands from the document's provenance that the data was generated by the patient and that it is authentic, reliable, and trustworthy.

10A.2 Activity Diagram

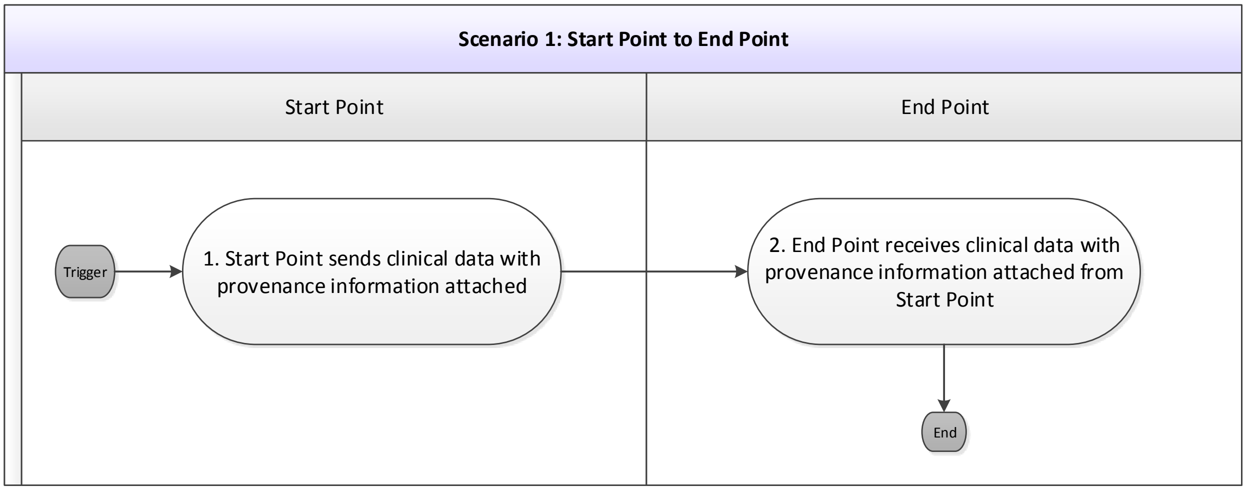
Scenario 1: Start Point to End Point
Start Point
End Point
Trigger
1. Start Point sends clinical data with provenance information attached
2. End Point receives clinical data with provenance information attached from Start Point
End

Figure 2: Activity Diagram for Scenario #1

10A.2.1 Base Flow

Step #	Actor	Role	Event/Description	Inputs	Outputs
1	Start Point	Send clinical data	Start Point sends clinical data with provenance information attached	Start Point selected clinical data with provenance information attached	Clinical data with provenance information in standard format and content specification where possible
2	End Point	Receive clinical data	End Point receives data from Start Point	Clinical data with provenance information in standard format and content specification where possible	End

Table 3: Base Flow for Scenario #1

10A.3 Functional Requirements

10A.3.1 Information Interchange Requirements

Initiating System	(describes action)	Information Interchange Requirement Name	(describes action)	Receiving System
Start Point	Send	Send Clinical Data with Provenance Information	Receive	End Point

Table 4: Information Interchange Requirements for Scenario #1

10A.3.2 System Requirements

System	*Minimum System Requirement

System	*Minimum System Requirement
Start Point	1. Create clinical data and provenance data 2. Maintain Clinical data and provenance data 3. Create exchange artifact 4. Attest clinical data and provenance data (where feasible)
End Point	1. Retain/Consume clinical data and provenance data 2. Access clinical data and provenance data

Note: System Requirements presented in this table are the minimum necessary to support the Use Case and are not all inclusive

Table 5: System Requirements for Scenario #1

10A.4 Sequence Diagram

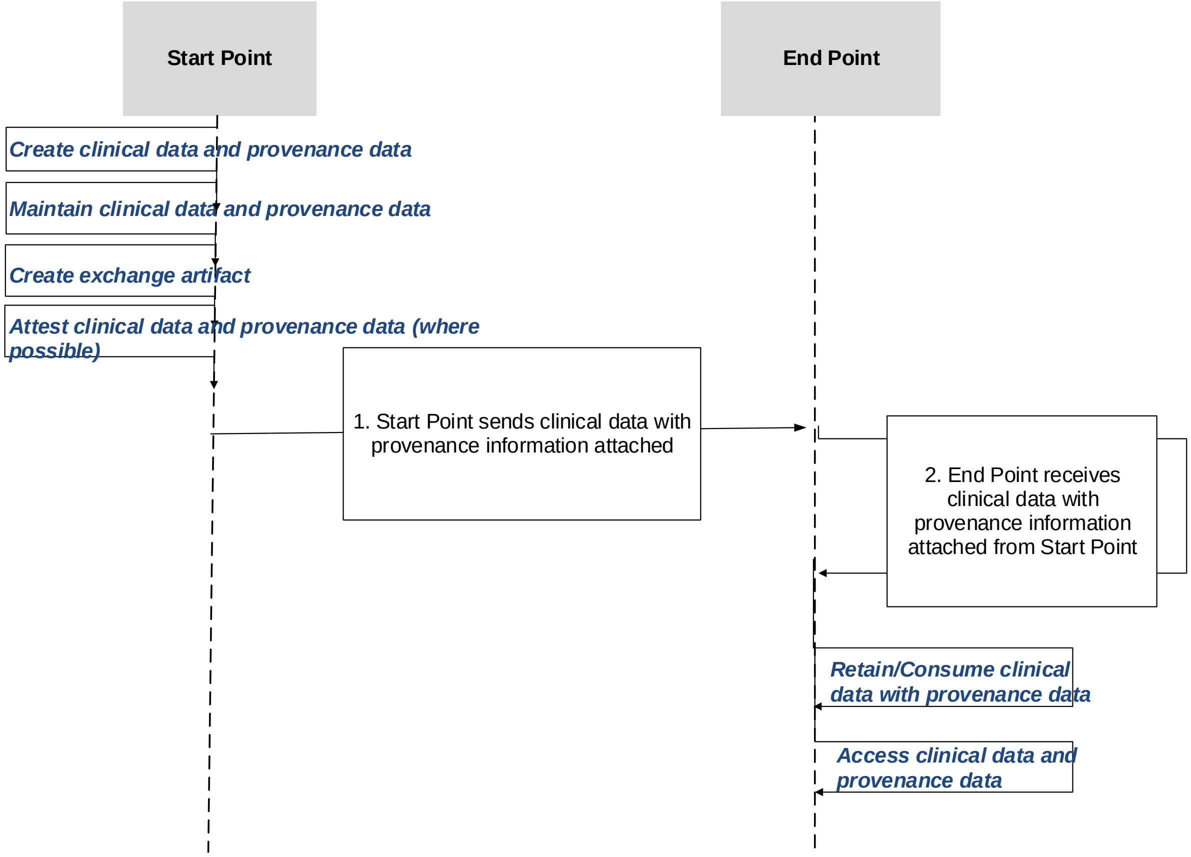
Start Point
End Point
Create clinical data and provenance data
Maintain clinical data and provenance data
Create exchange artifact
Attest clinical data and provenance data (where possible)
1. Start Point sends clinical data with provenance information attached
2. End Point receives clinical data with provenance information attached from Start Point
Retain/Consume clinical data with provenance data
Access clinical data and provenance data

Figure 3: Sequence Diagram for Scenario #1

10B.0 Scenario

Scenario 2: Start Point → Transmitter → End Point

10B.1 User Story

User Story 1 (no alteration in exchange): While training for a marathon, a patient fractures his foot. The patient's PCP conducts a foot exam and captures all of the data from that visit in his EHR. The PCP also calls in a referral for the patient to an orthopedic specialist for further treatment. After the PCP calls in the referral, the summary of care information is made available to the specialist, by passing through a transmitter, before being received by the orthopedic specialist's system. The orthopedic specialist receives the summary of care with provenance information and an indication that the data passed through a transmitter.

10B.2 Activity Diagram

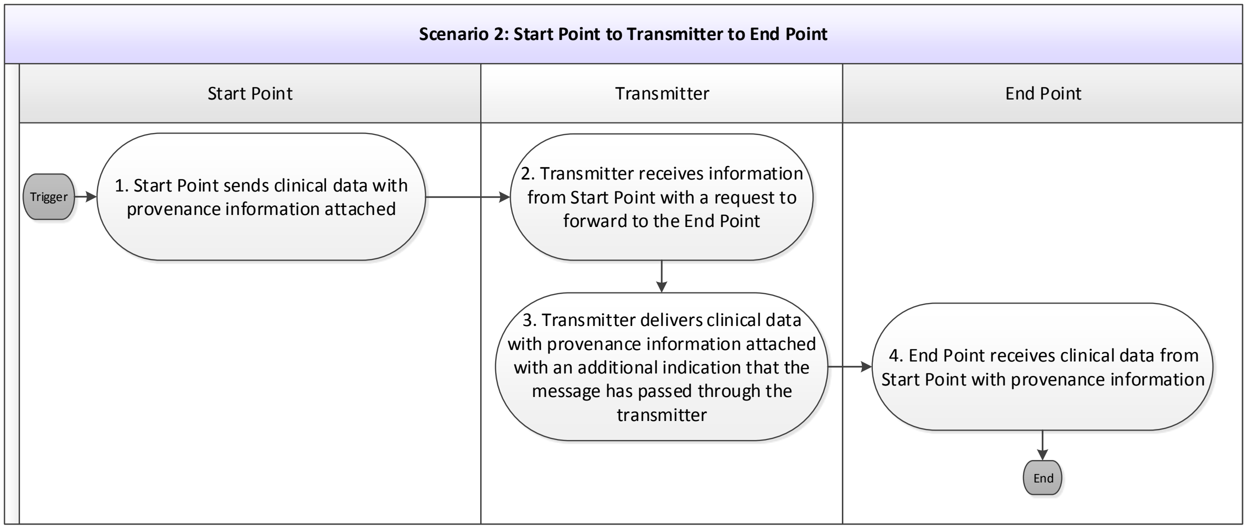
Scenario 2: Start Point to Transmitter to End Point
Start Point
Transmitter
End Point
Trigger
1. Start Point sends clinical data with provenance information attached
2. Transmitter receives information from Start Point with a request to forward to the End Point
3. Transmitter delivers clinical data with provenance information attached with an additional indication that the message has passed through the transmitter
4. End Point receives clinical data from Start Point with provenance information
End

Figure 4: Activity Diagram for Scenario #2

10B.2.1 Base Flow

Step #	Actor	Role	Event/Description	Inputs	Outputs
1	Start Point	Send clinical data	Start Point sends clinical data with provenance information attached	Start Point selected clinical data	Clinical data with provenance information in standard format and content specification where possible
2	Transmitter	Receive clinical data	Transmitter receives data from Start Point with a request to forward to the end point	Clinical data with provenance information in standard format and content specification where possible	Clinical data with provenance information in standard format and content specification where possible

Step #	Actor	Role	Event/Description	Inputs	Outputs
3	Transmitter	Send Clinical Data	Transmitter delivers clinical data with provenance information attached with an additional indication that the message has passed through the transmitter	Clinical data with provenance information in standard format and content specification where possible	Clinical data with provenance information (with indication of passing through transmitter) in standard format and content specification where possible
4	End Point	Receive clinical data	End Point receives data from Start Point	Clinical data with provenance information (with indication of passing through transmitter) in standard format and content specification where possible	END

Table 6: Base Flow for Scenario #2

10B.3 Functional Requirements

10B.3.1 Information Interchange Requirements

Initiating System	(describes action)	Information Interchange Requirement Name	(describes action)	Receiving System
Start Point	Send	Send Clinical Data	Receive	Transmitter
Transmitter	Send	Transmit Clinical Data	Receive	End Point

Table 7: Information Interchange Requirements for Scenario #2

10B.3.2 System Requirements

System	*Minimum System Requirement
Start Point	1. Create clinical data and provenance data 2. Maintain clinical data and provenance data 3. Create exchange artifact 4. Attest clinical data and provenance data (where feasible)
Transmitter	1. Able to add indication that clinical data and provenance data has passed through transmitter
End Point	1. Retain/Consume clinical data and provenance data 2. Access clinical data and provenance data

<u>Note</u>: System Requirements presented in this table are the minimum necessary to support the Use Case and are not all inclusive

Table 8: System Requirements for Scenario #2

10B.4 Sequence Diagram

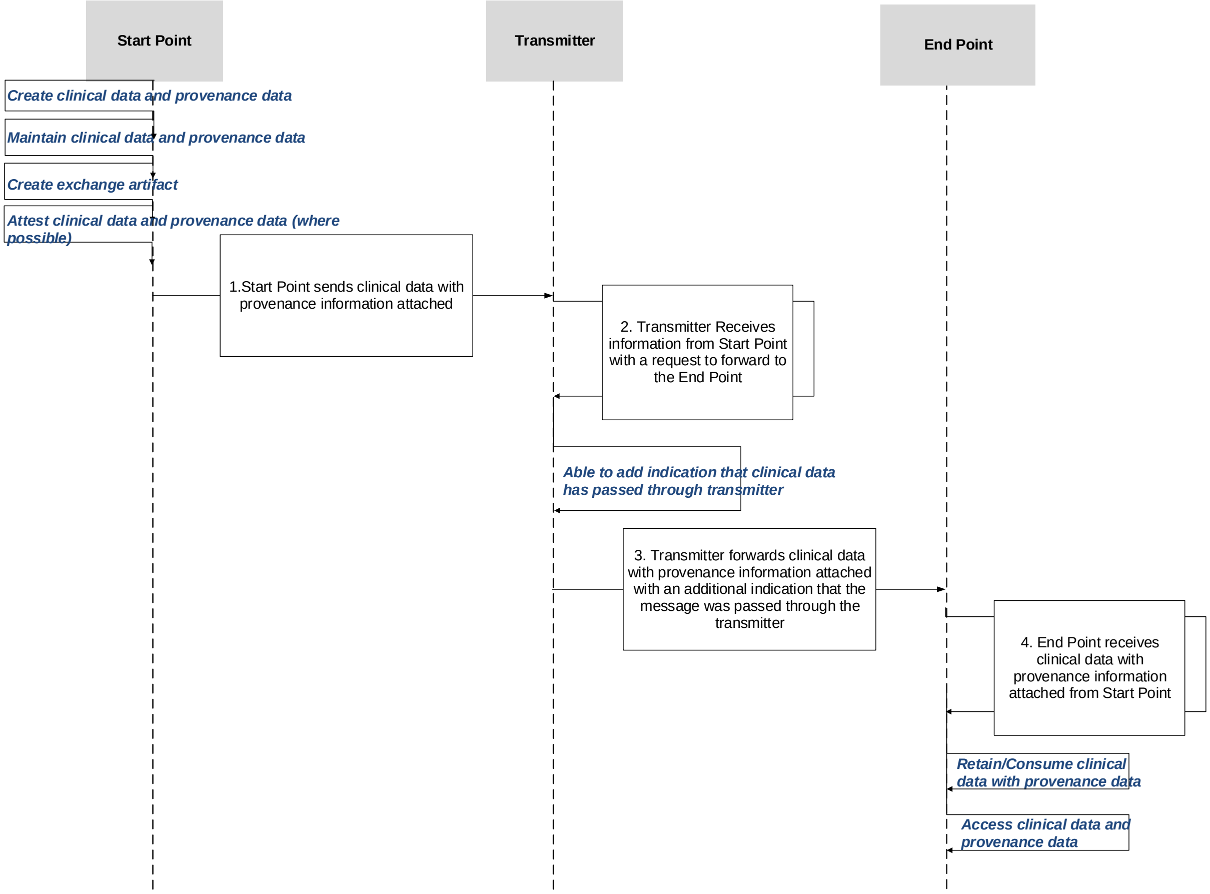
Start Point
Transmitter
End Point
Create clinical data and provenance data
Maintain clinical data and provenance data
Create exchange artifact
Attest clinical data and provenance data (where possible)
1.Start Point sends clinical data with provenance information attached
2. Transmitter Receives information from Start Point with a request to forward to the End Point
Able to add indication that clinical data has passed through transmitter
3. Transmitter forwards clinical data with provenance information attached with an additional indication that the message was passed through the transmitter
4. End Point receives clinical data with provenance information attached from Start Point
Retain/Consume clinical data with provenance data
Access clinical data and provenance data

Figure 5: Sequence Diagram for Scenario #2

10C.0 Scenario

Scenario 3: Start Point → Assembler /Composer→ End Point

10C.1 User Story

<u>Note:</u> *A community of providers have established a data use agreement that allows them to upload data to an HIE repository. When data is sent to the repository, the provenance information is also included.*

User Story 1: A patient is rushed to the Emergency Department due to a car accident. The physician wants to obtain the patient's summary record as part of the delivery of care. The physician queries the HIE repository which delivers a summary record or records from the past six months. The data received includes the provenance information from the originating sources and also information that identifies the assembler and the provenance related actions each has taken.

User Story 2: A patient with diabetes goes to Lab A to have his blood drawn. Lab A sends the lab results in a standard lab format to the PCP's EHR with provenance information attached. Upon reviewing the lab results, the PCP decides to refer the diabetic patient to a specialist for consultation. The PCP electronically sends a referral to the specialist. The referral document includes relevant data originating in the PCP's EHR along with provenance information from Lab A that is transformed into a representation that is compatible with the referral document.

10C.2 Activity Diagram

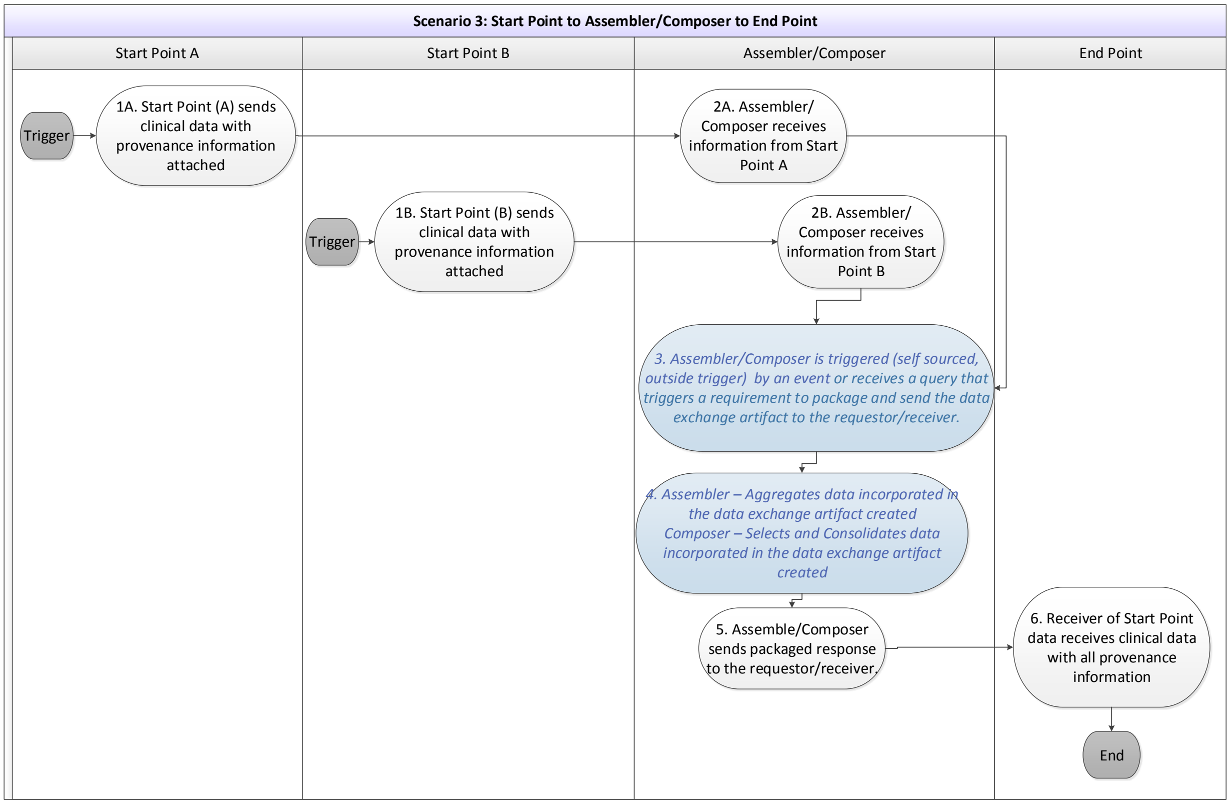
Scenario 3: Start Point to Assembler/Composer to End Point
Start Point A
Start Point B
Assembler/Composer
End Point
Trigger
1A. Start Point (A) sends clinical data with provenance information attached
2A. Assembler/ Composer receives information from Start Point A
Trigger
1B. Start Point (B) sends clinical data with provenance information attached
2B. Assembler/ Composer receives information from Start Point B
3. Assembler/Composer is triggered (self sourced, outside trigger) by an event or receives a query that triggers a requirement to package and send the data exchange artifact to the requestor/receiver.
4. Assembler – Aggregates data incorporated in the data exchange artifact created Composer – Selects and Consolidates data incorporated in the data exchange artifact created
5. Assemble/Composer sends packaged response to the requestor/receiver.
6. Receiver of Start Point data receives clinical data with all provenance information
End

Figure 6: Activity Diagram for Scenario #3

10C.2.1 Base Flow

Step #	Actor	Role	Event/Description	Inputs	Outputs

Step #	Actor	Role	Event/Description	Inputs	Outputs
1A	Start Point A	Send clinical data	Start Point sends clinical data with provenance information attached	Start Point selected clinical data	Clinical data with provenance information in standard format and content specification where possible
2A	Assembler	Receive clinical data	Assembler receives information from Start Point A	Clinical data with provenance information in standard format and content specification where possible	Clinical data with provenance information in standard format and content specification where possible
1B	Start Point B	Send clinical data	Start Point sends clinical data with provenance information attached	Start Point selected clinical data	Clinical data with provenance information in standard format and content specification where possible
2B	Assembler/ Composer	Receive clinical data	Assembler/Composer receives information from Start Point B	Clinical data with provenance information in standard format and content specification where possible	Clinical data with provenance information in standard format and content specification where possible
3	*Assembler/Composer*	*Package clinical data*	*Assembler/ Composer is triggered (self-sourced, outside trigger) by an event or receives a query that triggers a requirement to package and send the data exchange artifact to the requestor/receiver*	*Clinical data from sources A and B with provenance information in standard format and content specification where possible*	*Clinical Data with provenance information from sources A and B*

Step #	Actor	Role	Event/Description	Inputs	Outputs
4	*Assembler/Composer*	*Package clinical data*	*Assembler – Aggregate Data to be incorporated in the data exchange artifact created* *Composer – Select and consolidate data to be incorporated in the data exchange artifact created*	*Clinical Data with provenance information from sources A and B*	*Packaged clinical data with provenance information in standard format and content specification where possible*
5	Assembler/ Composer	Sends clinical data	Assembler sends packaged response	Packaged clinical data with provenance information in standard format and content specification where possible	Clinical data with provenance information (with indication of it being assembled) in standard format and content specification where possible
6	End Point	Receive clinical data	End Point receives data from Start Point	Clinical data with provenance information (with indication of it being assembled) in standard format and content specification where possible	END

Table 9: Base Flow for Scenario #3

10C.3 Functional Requirements

10C.3.1 Information Interchange Requirements

Initiating System	(describes action)	Information Interchange Requirement Name	(describes action)	Receiving System
Start Point	Send	Sent Clinical Data	Receive	Assembler/Composer
Assembler/Composer	Send	Compiled Clinical Data	Receive	End Point

Table 10: Information Interchange Requirements for Scenario #3

10C.3.2 System Requirements

System	*Minimum System Requirement
Start Point	1. Create clinical data and provenance data 2. Maintain clinical data and provenance data 3. Create exchange artifact 4. Attest clinical data and provenance data (where feasible)
Assembler	1. Access clinical data 2. Extract and aggregate clinical data 3. Maintain clinical data and provenance data 4. Create exchange artifact
Composer	1. Access clinical data 2. Extract, **select** and consolidate clinical data 3. Maintain clinical data and provenance data 4. Create exchange artifact
End Point	1. Retain/Consume clinical data and provenance data 2. Access clinical data and provenance data

Note: System Requirements presented in this table are the minimum necessary to support the Use Case and are not all inclusive.

Table 11: System Requirements for Scenario #3

10C.4 Sequence Diagram

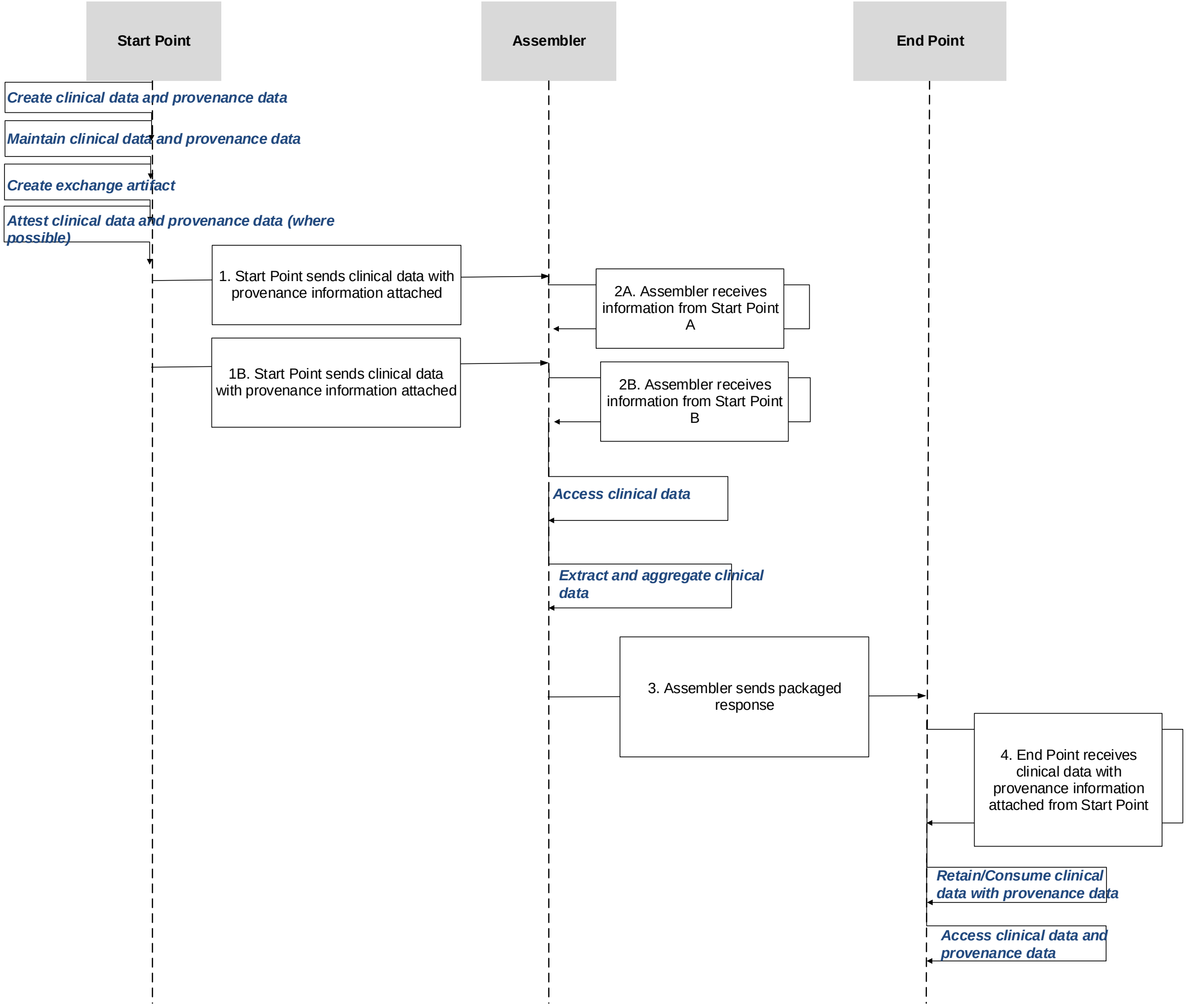
Start Point
Assembler
End Point
Create clinical data and provenance data
Maintain clinical data and provenance data
Create exchange artifact
Attest clinical data and provenance data (where possible)
1. Start Point sends clinical data with provenance information attached
2A. Assembler receives information from Start Point A
1B. Start Point sends clinical data with provenance information attached
2B. Assembler receives information from Start Point B
Access clinical data
Extract and aggregate clinical data
3. Assembler sends packaged response
4. End Point receives clinical data with provenance information attached from Start Point
Retain/Consume clinical data with provenance data
Access clinical data and provenance data

Figure 7: Sequence Diagram for Scenario #3 (Assembler)

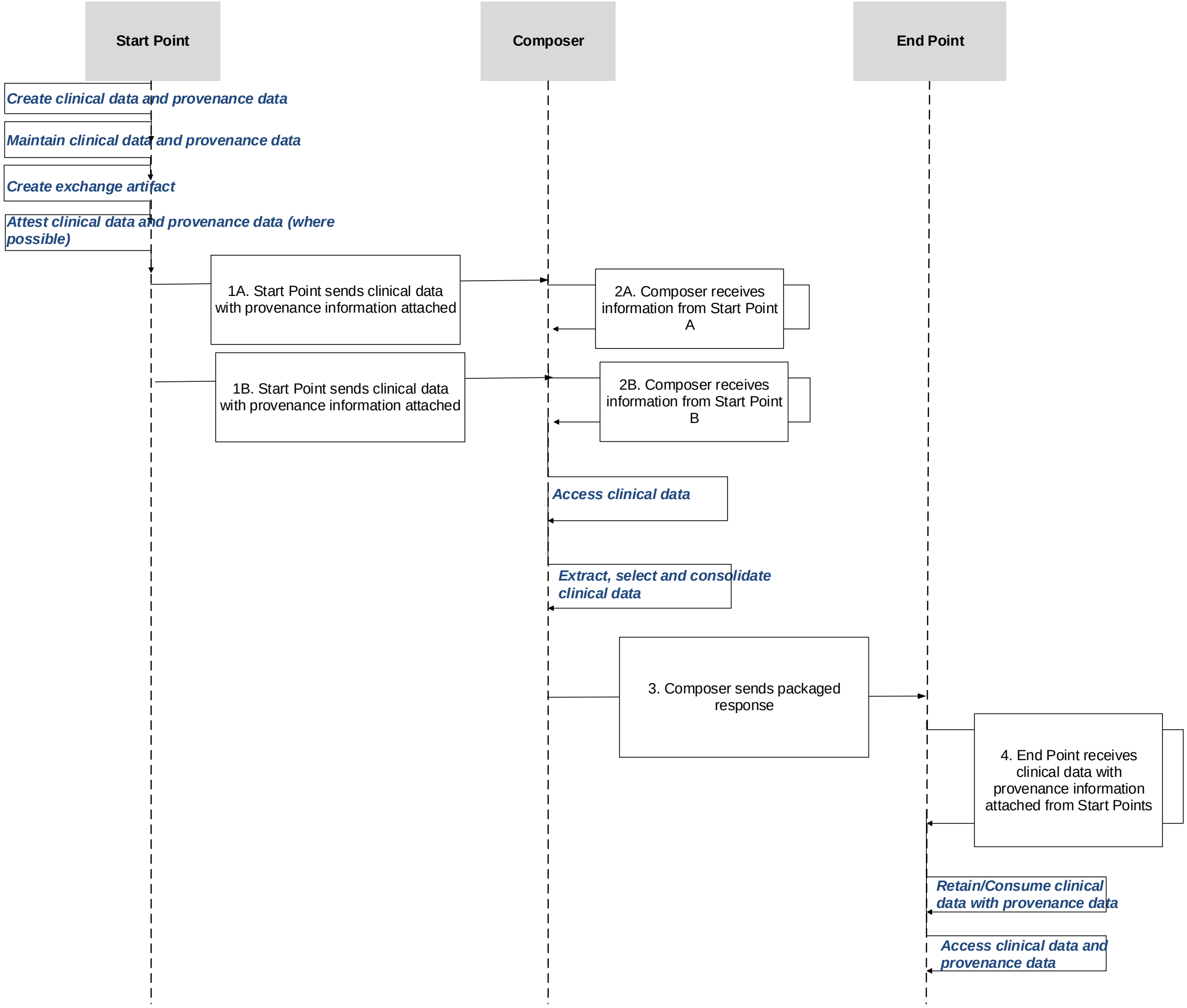
Start Point
Composer
End Point
Create clinical data and provenance data
Maintain clinical data and provenance data
Create exchange artifact
Attest clinical data and provenance data (where possible)
1A. Start Point sends clinical data with provenance information attached
2A. Composer receives information from Start Point A
1B. Start Point sends clinical data with provenance information attached
2B. Composer receives information from Start Point B
Access clinical data
Extract, select and consolidate clinical data
3. Composer sends packaged response
4. End Point receives clinical data with provenance information attached from Start Points
Retain/Consume clinical data with provenance data
Access clinical data and provenance data

Figure 8: Sequence Diagram for Scenario #3 (Composer)

11.0 Dataset Considerations

Note: This is a starting set of Dataset Requirements and will be further refined during Harmonization, which may include adding or removing elements.

ROLE	DATA CATEGORY	DATA ELEMENT
START POINT	*Who*	Sending System
		Sending System Organization
		Author
		Custodian
		Role
	When	Send Date
		Send Time

	Where	Address
		State
		Zip
	Type (What)	Software
		Device
	Why	Clinical Context
		Purpose
	Integrity/Authenticity	Digital Signature
	Additional	Patient Record Target Assigned Author Informant Service Event Performer Authenticator Legal Authenticator
TRANSMITTER	*Who*	Transmitter Organization
		Transmitter System
	When	Transmission Time Sent
		Transmission Date Sent
	Where	Transmitter Location
		Transmitter System Location
	Type (What)	Transmission Device
		Transmission Software
		Transmission Hardware
		Transmission Method
	Why	Purpose of Transmission
	Routing	Transmitter Sender Address
		Receiver Address
	Integrity/Authenticity	Digital Signature
	Who	Transmitter Organization
		Transmitter System
	Additional	Patient Record Target
ORIGINATOR	*Who*	Originator Organization
		Originator Author
		Originator Enterer
		Originator Attester
		Originator Verifier
		Originator System
	When	Originator Time Created

	Where	Originator Locations
		Originator System Location
	Type (What)	Originator Event
	Additional	Patient Record Target Author Assigned Author Authoring System Authoring Organization Informant Service Event Performer Participant Custodian Authenticator Legal Authenticator
ASSEMBLER	*Who*	Assembler System
		Assembler Organization
		Intended Recipient
	When	Assembly Date
		Assembly Time
	Where	Address
		State
		Zip
	Type (What)	Software
		Device
	Why	Assembly Purpose
	Integrity/Authenticity	Assembly Participants
		Attestation/Nonrepudiation of data
	Additional	Patient Record Target Author Assigned Author Authoring System Authoring Organization Informant Service Event Performer Participant Custodian Authenticator Legal Authenticator
COMPOSER	*Who*	Composer System

		Composer Organization
	When	Composition Date
		Composition Time
	Where	Address
		State
		Zip
	Type (What)	Software
		Device
	Why	Composing Purpose
	Integrity/Authenticity	Composing Participants
		Selector
	Additional	Patient
		Record Target
		Author
		Assigned Author
		Authoring System
		Authoring Organization
		Informant
		Service Event
		Performer
		Participant
		Custodian
		Authenticator
		Legal Authenticator

Table 12: Dataset Requirements

12.0 Risks, Issues and Obstacles

1. the management of data provenance will become harder due to record/data extraction and re-assembly, and its movement through the life cycle and multiple locations
   - This is not an auditing consideration; this refers to the accumulation of provenance and the movement of data
   - When a provenance event occurs, the provenance information is changed accordingly; the accumulation of provenance changes poses a challenge
   - Derivations of information with initial provenance create the need for new provenance of the derived data; this can create multiple threads extending from the original information
2. There is a lack of experience working with provenance data
3. Granularity at which provenance data is recorded and exchanged may not be sufficient to address user requirements

4. With technology constantly changing and new data origination sources constantly emerging, there is room for future documenting challenges
5. The format in which data is recorded and exchanged may impact the level and form in which provenance data is provided, adding complexity to storage and maintenance
6. Variability in the data assembly processes may impact the ability to maintain and complete provenance data
7. If source data has different levels of assurance of assertions, (e.g. digital signatures) this creates issues with trust and certification of authenticity of data
    - Assertions have different levels of assurance associated with them (different levels with digital signatures for example)
8. Conflation of processes that rely on provenance with the capture of provenance itself

Appendices

Appendix A: References

- Data Provenance Homepage: http://wiki.siframework.org/Data+Provenance+Initiative
- Data Provenance Use Case: http://wiki.siframework.org/Data+Provenance+Use+Cases
- Data Provenance Glossary: http://wiki.siframework.org/Data+Provenance+Glossary
- Data Provenance References: http://wiki.siframework.org/Data+Provenance+References
- Other References:
    - "Report to the President: Realizing the Full Potential of Health Information Technology to Improve Healthcare for Americans: The Path Forward. "*President's Council of Advisors on Science and Technology*.
        - http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-health-it-report.pdf