Health IT Standards Committee

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

Data Provenance Task Force

Final Recommendations

Lisa Gallagher, HIMSS, Chair

January 27, 2015

Outline

- Background
  - Task Force Members
  - Data Provenance Task Force Charge and Supporting Questions
- Recommendations for Supporting Questions
  - Supporting Question #1
  - Supporting Question #2
  - Supporting Question #3
- Committee Discussion

BACKGROUND

Task Force Members

Member Name	Organization
Lisa Gallagher, Chair	HIMSS
Rebecca D. Kush	CDISC
John Moehrke	GE
Floyd Eisenberg, MD	iParsimony, LLC
Aaron Seib	National Association for Trusted Exchange (NATE)
Mike Davis (Workgroup Federal Ex Officio)	US Department of Veterans Affairs

Task Force Charge

Specific Question from ONC:

Given the community-developed S&I Data Provenance Use Case, what first step in the area of data provenance standardization would be the most broadly applicable and immediately useful to the industry?

Supporting Questions

1) **Do the 3 scenarios in the Use Case, and the Use Case's identified scope, address key data provenance areas, or is something missing?**
   a) Yes, the scenarios address key provenance areas
   b) No, some key data provenance areas are missing

2) **The Use Case is broad and spans a lot of challenges. Where in the Use Case should the Initiative start in terms of evaluating standards to meet Use Case requirements?**
   a) At the point of data creation in a Patient Controlled Device (PCD) or PHR?
   b) At the point of origin/data creation in an EHR or HIE?
   c) With the transfer of data from a PCD/PHR to an EHR system?
   d) With exchange of data between EHRs?

3) **Are there any architecture or technology specific issues for the community to consider?**
   a) Content: Refining provenance capabilities for CDA/C-CDA while supporting FHIR?
   b) Exchange: Push (e.g. DIRECT), Pull (SOAP and REST-based query responses)?
   c) Others?

RECOMMENDATIONS

Question #1
High level Recommendation

Do the 3 scenarios in the Use Case, and the Use Case's identified scope, address key data provenance areas, or is something missing?

a) Yes, the scenarios address key provenance areas

b) No, some key data provenance areas are missing

RESPONSE:

The Use Case may be over-specified. The Task Force recommends that the Data Provenance Initiative should focus on the following:

A. Where did the data come from? ("source provenance")

B. Has it been changed?

C. Can I trust it (the data)?

Question #1
Detailed Recommendations

1. Begin focus from the perspective of an EHR - Provenance of the intermediaries is only important if the source data is changed. Therefore, begin focus from the perspective of an EHR, including provenance for information created in the EHR ("source provenance") and when it is exchanged between two parties.

   The notion of "who viewed/used/conveyed without modification along the way" is not important for provenance, as long as the information was not changed.

Question #1
Detailed Recommendations (Cont.)

2. Clearly differentiate between Communication/Information Interchange requirements and System Requirements

   Both are important. For the purposes of this use case-- Start with the assumption that at the point for information interchange, the "source provenance" is good, complete, trusted.

   a. Address Communication/Information Interchange requirements
      - Note: As a basic requirement, converting between different transport protocols should be lossless, i.e., retain integrity, in terms of provenance of the payload/content.

   b. Address System Requirements for provenance (including "source provenance") by looking at provenance data at time of import, creation, maintenance, and export.
      - Note: Agnostic of transport technologies
      - Consider FDA Project, Guidance and Regulations - There are 12 requirements and use cases for the use of EHRs and eSource applications (e.g. patient reported information/eDiaries) requiring provenance described in an *eSource Data Interchange Document, FDA Guidance*, which includes a definition for **"the source"** and regulation for Electronic Records.

Question #1
Detailed Recommendations (Cont.)

3. Consider the definition of "change" to data (for example, amend, update, append, etc.) and the implications for provenance. If the content changes, the change should be considered a "provenance event."
4. Consider the implications of security aspects – Traceability, audit, etc. – what is the impact on the trust decision?
5. If applicable, capture policy considerations and request further guidance from the HITPC. For example,

Can I trust it and has it been changed? Consider that, for clinical care, if trending the data, one may need to know the degree to which the information can be trusted.

Defining levels of trust would be a ***policy issue.***

Question #2
High-Level Recommendations

The Use Case is broad and spans a lot of challenges. Where in the Use Case should the Initiative start in terms of evaluating standards to meet Use Case requirements?

RESPONSE:

- Given the recommendations above, the TF recommends addressing the Use Case in the following priority order:
  a) With exchange of data between EHRs
  b) At the point of origin/data creation in an EHR or HIE
  c) With the transfer of data from a PCD/PHR to an EHR system
  d) At the point of data creation in a Patient Controlled Device (PCD) or PHR
- The Initiative should clearly differentiate a set of basic/core requirements for provenance.

Question #2
Detailed Recommendations

1. Determine if "Origination of the Patient Care Event Record Entry" is in scope
   a. Address "source provenance" data within an EHR
   b. Consider those provenance events which an EHR would need for:
      a. import, create, maintain, export
   c. Define "source" (consider FDA definition below)
      - **Source Data:** All information in original records and certified copies of original records of clinical findings, observations, or other activities (in a clinical investigation) used for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies).

Question #2
Detailed Recommendations (Cont.)

2. Add CDISC ODM to the candidate standards list.

3. Consider if there are related requirements that may have implications (i.e., regulatory, program specific), for example:
   - Medical Record retention
   - Data receipts
   - esMD (digital signature)

Question #3 Recommendations

Are there any architecture or technology specific issues for the community to consider?

a) Content: Refining provenance capabilities for CDA/C-CDA while supporting FHIR?

RESPONSE: Consider related work in HL7 projects, such as:

- CDA/C-CDA provenance
- FHIR Provence Project
- Privacy on FHIR Projects

b) Exchange: Push (e.g. DIRECT), Pull (SOAP and REST-based query responses)?

RESPONSE: In Information Interchange – The provenance of content should be lossless (retain integrity).

COMMITTEE DISCUSSION

BACKUP SLIDES

Supporting Information

- **21 Code of Federal Regulations (CFR) Part 11 -Electronic Records; Electronic Signatures — Scope and Application, U.S. Department of Health and Human Services Food and Drug Association**
  http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm125125.pdf

- **eSource Data Interchange Document** - Output of a team that convened at the request of FDA to help move to new technology and electronic source data (vs. paper) while adhering to global and FDA regulations and guidance. It contains use cases and **12 basic requirements around provenance related to exchanging data between patients, clinical sites and research sponsors.**
  http://www.cdisc.org/system/files/all/reference_material_category/application/pdf/esdi.pdf

- **FDA eSource Guidance** based upon 21CFR11 and the above work, r.e. a continued need to move to electronic data collection and new technologies (eDiaries, EHRs).
  http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM328691.pdf

NOTE: While focused on research, the principles around provenance are very similar to those that the FDA was addressing with their requirement for traceability of the data that they receive (and thus the Federal Regulations for 'audit trails' for electronic data exchange).

Definitions from FDA Guidance

- **Electronic Health Record (EHR):** An electronic record for healthcare providers to create, import, store, and use clinical information for patient care, according to nationally recognized interoperability standards. NOTE: The EHR has the following distinguishing features: able to be obtained from multiple sources, shareable, interoperable, accessible to authorized parties.
- **Electronic Record:** Any combination of text, graphics, data, audio, pictorial, or other information representation in digital form that is created, modified, maintained, archived, retrieved, or distributed by a computer system (21 CFR 11.3(b)(6)).
- **Electronic Signature:** A computer data compilation of any symbol or series of symbols executed, adopted, or authorized by an individual to be the legally binding equivalent of the individual's handwritten signature (21 CFR 11.3(b)(7)).
- **Electronic Source Data:** Electronic source data are data initially recorded in electronic format.
- **Source Data:** All information in original records and certified copies of original records of clinical findings, observations, or other activities (in a clinical investigation) used for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies).

CDISC Operational Data Model (ODM)

Operational Data Model (ODM)

- Supports forms-based data transport (case records) and use cases of exchanging data between and among eDiaries, ePRO devices, EHRs, Electronic Data Capture tools…
- Mature global standard in use since 1999; developed as a global consensus-based standard and supported through the XML Technologies team led by CDISC
- ONC Structured Data Capture Initiative, IHE SDC profile and community-driven Interoperability Specification #158 (2010)
- In use in Europe, U.S. and Japan to export data from EHRs for research, public health and safety reporting
- ODM Certification Program established in 2007
- Adheres to 21CFR11 requirements for audit trail and supports electronic signatures

Data Provenance Initiative
Use Case Summary

Scenario 1: Start Point -> End Point.

Describes simple provenance requirements when transferring healthcare data from a Start Point (sending system) to an End Point (Receiving System).

Scenario 2: Start Point -> Transmitter -> End Point.

Includes use of a third party as a conduit/transmitter to transfer information from Start Point to End Point. There may be use cases where it is important to know how the information was routed, as well as who originated it and who sent it.

Scenario 3: Start Point ->Assembler / Composer -> End Point.

Uses a third party system to aggregate or combine information from multiple sources, either in whole or in part, to produce new healthcare artifacts. The new artifacts may contain information previously obtained from multiple sources, as well as new information created locally.

Data Provenance Initiative Use Case Summary