Linked Data: Fast, low cost semantic interoperability for health care?
About the presentation

• **Part I: Motivation**
  • Why we need semantic operability in health care
  • Why enhancing existing systems to increase semantic interoperability may be necessary

• **Part II: A Linked Data-based solution**
  • Linked Data background
  • Applying the Linked Data approach in health care
  • Work at AEHRC
Pt 1: Semantic interoperability in health care

- **Definitions**
  - ... means that the information exchanged by different computer systems can be interpreted by both computer applications and human users (NeHTA)
  - ... means ensuring that the precise meaning of exchanged information is understandable by any other system or application not initially developed for this purpose (EC Recommendation COM(2008) 3282 final)
  - ... is the seamless inter- and intra-organisational co-operation between health information systems of fully machine interpretable, standardised and coded data (Stroetmann, 2009; Walker et al., 2005)

- **Maturity models (e.g. Walker’s Levels 1-4)**

- **High semantic interoperability maturity is assumed**
  - Canada Health InfoWay
  - NHS Connecting for Health
  - Australian Department of Health & Aging HealthConnect
  - Interoperability is a significant part of NeHTA’s work program
Options for increasing maturity

- **The reality of the health care system is that budgets are less important than treating patients**
  - IT projects rarely get funded to simply replace existing functionality
  - Increasing the scope of a system implementation increases costs

- **Enhancing existing systems to support higher levels of semantic interoperability may be a preferred (or only) option**

- **Using open standards is a good idea**
  - Avoid silos (and vendor lock-in) down the track
  - Information models: openEHR
  - Clinical terminologies: SNOMED CT
So where are we?

- Increasing semantic interoperability maturity is a Good Thing.

- It’s not always possible or appropriate to increase semantic interoperability maturity with new system development.

- We’d like a way to quickly increase semantic interoperability maturity for existing systems at a low cost.
Pt 2: Linked data for health system interoperability?

• **Linked data background**
  - Conceptual difference in integration styles
  - What is it
  - How does it work

• **Applying the linked data approach to health care**
  - Vocabulary
  - What tools are and are not available?
  - Applications
  - Issues
Approaches to Integration: Typical

- Flat file/CSV
- HL7 Message
- API/XML

Inspired by Heath 2009
Approaches to Integration: Linked Data

Clinical models (Archetypes) refer to:

- Clinical terminology (SNOMED CT)

Typed Links:

- EDIS → CIS
- CIS → Path
- EDIS → Path

Typed Links:

- CIS → EDIS
- Path → CIS
- Path → EDIS
What is Linked Data?

• Linked data is about using Web technologies to create typed links between data from different sources.

• A recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Web using URIs and RDF (Wikipedia, 16 August 2009)

• A step on the road to the Semantic Web that provides a method for publishing data that:
  • encourages reuse
  • reduces redundancy
  • maximises its (real and potential) inter-connectedness
  • enables network effects to add value to data (Health 2009)
Linked Data Technology Stack

• **URIs**
  - Uniform Resource Identifier
  - A simple and extensible means for identifying a “resource”

• **HTTP**
  - Hypertext Transfer Protocol
  - A widely accepted data access mechanism

• **RDF**
  - Resource Description Framework
  - Data format for describing things and their interrelations

• **(RDFS/OWL)**
  - RDF Vocabulary Definition Language (RDFS)
  - Web Ontology Language (OWL)
  - More expressive ways for describing “things” in the world and how they are related using classes and properties
  - Vocabularies: collections of RDFS/OWL classes and properties
Linking Open Data Cloud (Bizer 2009)
Linked Open Drug Data (LODD) Cloud
Clinical Linked Data Vocabulary

- Many useful vocabularies already exist (re-use)
  - LODD vocabularies
  - FOAF – describes people (identities, affiliations, social networks)
  - Geo-names - describes places
- Clinical Model
  - Map archetypes to OWL (Kilic 2005)
- Terminology
  - OWL representation of SNOMED-CT
Clinical Linked Data Tooling

• **Clinical Linked Data publishing tool**
  • Generating compliant Linked Data for clinical data based on openEHR mappings

• **Semantic link discovery**
  • Use vocabulary property-based matching algorithms to find relationships between data
    • Identity resolution -> patient master list
    • Clinical case matching

• **Patient Journey Browser**
  • Faceted browsing Linked Data for a particular patient

• **Archetype Query Language**
  • Domain specific SPARQL
  • Queries involving other vocabularies (e.g. Gene Ontology)
Summary

• Linked data is an increasingly popular approach to semantic integration on the Web
• The approach can be applied to systems integration both within and between health care organisations
  • Existing tools and vocabularies provide assistance, although some tools are missing
  • Enables the development of applications that leverage other Linked Data sets
• Our work is in building tools for and applying the approach in health care
Thank you

Simon McBride
Stream Leader (Health Informatics)
Project Leader (Health Information Systems)
Phone: +61 7 3253 3631
Email: simon.mcbride@csiro.au
Web: http://aehrc.com