RDF as a Universal Healthcare Exchange Language

David Booth, Hawaii Resource Group
Conor Dowling, Caregraf
Michel Dumontier, Stanford University
Josh Mandel, Harvard University
Claude Nanjo, Cognitive Medical Systems
Rafael Richards, Veterans Affairs

Semantic Technology and Business Conference
21-Aug-2014

Download the latest version of these slides from http://dbooth.org/2014/rdf-as-univ/
Workshop Agenda
21-Aug-2014


• [10:50] **Data-Driven Biomedical Research with Semantic Web Technologies** – Michel Dumontier – Slides: http://dbooth.org/2014/dumontier/


• [11:35] **Panel Discussion** – All
Imagine a world
Imagine a world in which all healthcare systems speak the same language with the same meanings covering all healthcare.
Healthcare today

Tower of Babel, Abel Grimmer (1570-1619)
REPORT TO THE PRESIDENT
REALIZING THE FULL POTENTIAL OF
HEALTH INFORMATION TECHNOLOGY
TO IMPROVE HEALTHCARE
FOR AMERICANS:
THE PATH FORWARD

Executive Office of the President
President’s Council of Advisors
on Science and Technology

December 2010
"PCAST has also concluded that to achieve these objectives it is crucial that the Federal Government facilitate the nationwide adoption of a universal exchange language for healthcare information"
2013 Workshop on "RDF as a Universal Healthcare Exchange Language"

• 32 participants

• Ended up creating the Yosemite Manifesto . . .
Yosemite Manifesto Endorsements

1. David Booth, Ph.D., KnowMED, Inc.
2. Charlie Mead, M.D., MSc., Octo Consulting Group
3. Tracy Allison Altman, Ph.D., PepperSlice
4. Michel Dumontier, Associate Professor of Bioinformatics, Carleton University
5. Rafael Richards MD MS, Johns Hopkins School of Medicine
6. Stanley M. Huff, MD, CMIO Intermountain Healthcare
7. Olivier Cari, PhD, UPEM France
8. Emory Fry, MD, Cognitive Medical Systems
9. Karl Seiler, CEO and founder NEIMO Health, a Modus Operandi, Inc. Business
10. Erick Von Schweber, Executive Co-chair SURVEYOR
11. Tom Munnecke, Independent Consultant
12. Thomas J. Kelly, PMB, Cognizant Technology Solutions
13. Dean Allemang, PhD, Working Ontologist LLC
15. Dean Allemang, PhD, Working Ontologist LLC
16. David Booth, Ph.D., KnowMED, Inc.
17. Dave McComb,, Semantic Arts
18. Manuel Wahle, Dipl.­Inform, MS, The University of Texas Health Science Center at Houston
19. Michael Erdmann, PhD, DIQA Gmbh (Germany)
20. Regin Forsberg, Principal Informatics Scientist, AstraZeneca
22. Mark Montgomery, Founder & CEO, Xf-yield
23. Karl Riefi, CEO, Crossflow Software
24. David L. Wosnenden President, eXevic Technologies LLC
25. Matthew Vagnoni, MD, CTO KnowMED.com
26. Christopher Ryan
27. Doug Burke, President, Cognitive Medical Systems
28. Jerry Scott, Emvee Partners LLC
29. Rick Pope, Cognitive Medical Systems
30. Charles B. Brown, MD, CMIO, Mckesson
31. Conor Dougan, CDI, Coringle
32. James McCusker, CIO University
33. Cartik Kothari, PhD, CEO, Perfect Informatics, IN/DIA
34. Carl Mattise, Founder, Wellness Intelligence Institute
35. Laz Yagourdis, VP and Founder, Cambridge Semantics
36. Yves Févrais, VP Health IT Policy, Kaiser Permanente
37. Christian Joppa, Senior Developer, Apachejena Inc.
38. Dr. Matthias Samwald, Medical University of Vienna
39. Michael Goble, PhD, Senior Ontology Consultant, Semantic Arts, USA
40. Zef McBride, MSc, NBB, CIO
41. Katrina Dreher, PhD student, UvA University Amsterdam & University of Amsterdam
42. Claude Thoya, MSc MBB, Zynx Health Inc
43. Murray West, a researcher
44. Pedro Lopez, PhD, University of Sharjah
45. Sih Jacob, Senior Information Architect, Ramag Healthcare
46. Carsten Schrader, Senior Software Engineer, The MAPS Corporation
47. Michael Davis, PES, ontology engineer
48. Robert Stanley, CEO, IO Informatics
49. Ryoatsu, Senior PD, PhD, Semantic Identity
50. Ryoatsu, Senior PD, PhD, Semantic Identity
51. Ryoatsu, Senior PD, PhD, Semantic Identity
52. Stephane Fellah, CTO, smartRealm LLC
53. Frank van Harmelen, Prof., VU University Amsterdam
54. Tracey Allison Altman, Ph.D., PepperSlice
55. Tim Finin, Professor, University of Maryland, Baltimore County
56. François Aurenche, M.D. de médecine, Université Montpellier 2
57. Charles E. Cooper, Professor, National Cancer Institute (NIH)
58. James J. Lavan, MD, PhD, Rened Networks Professor, Syrmaic Biological Institute, President, Proteinic Medicine, Inc.
59. M. J. Neklani, MD, PhD,メーカー, Chief, Materials, The Netherlands
60. Mikael Diabetes, PhD, University of Munich
61. Alan Kruppa, Director, Director of Data Warehouse at Institute for Health Informatics, University of Buffalo
62. Ben Birgmeier (2014
63. Frank van Harmelen, Pro, VU University Amsterdam
64. Tracy Allison Altman, Ph.D., PepperSlice
65. Rafael Richards MD MS, Johns Hopkins School of Medicine
66. Stanley M. Huff, MD, CMIO Intermountain Healthcare
67. Olivier Cari, PhD, UPEM France
68. Emory Fry, MD, Cognitive Medical Systems
69. Karl Seiler, CEO and founder NEIMO Health, a Modus Operandi, Inc. Business
70. Erick Von Schweber, Executive Co-chair SURVEYOR
71. Tom Munnecke, Independent Consultant
72. Thomas J. Kelly, PMB, Cognizant Technology Solutions
73. Dean Allemang, PhD, Working Ontologist LLC
74. Erich A, Gombocz, CSO, IO Informatics, Inc.
75. Dean Allemang, PhD, Working Ontologist LLC
76. David Booth, Ph.D., KnowMED, Inc.
77. Michael Erdmann, PhD, DIQA Gmbh (Germany)
78. Jamie Ferguson, VP Health IT Policy, Kaiser Permanente.
79. Lee Feigenbaum, VP and Founder, Cambridge Semantics
80. Michael Uschold, PhD, Senior Ontology Consultant, Semantic Arts, USA
81. Dr. Matthias Samwald, Medical University of Vienna
82. Michael Goble, PhD, Senior Ontology Consultant, Semantic Arts, USA
83. James McCusker, CIO University
84. Zef McBride, MSc, NBB, CIO
85. Katrina Dreher, PhD student, UvA University Amsterdam & University of Amsterdam
86. Peter Grant, PhD, University of Buffalo
87. Charles B. Brown, MD, CMIO, Mckesson
88. Rick Pope, Cognitive Medical Systems
89. Jerry Scott, Emvee Partners LLC
90. Jon McBride, BACS, MBA, CIO
91. Charles B. Owen, MD, CMIO, Afoundria
92. Rick Pope, Cognitive Medical Systems
93. Jerry Scott, Emvee Partners LLC
94. Jon McBride, BACS, MBA, CIO
95. Charles B. Owen, MD, CMIO, Afoundria
96. Rick Pope, Cognitive Medical Systems
97. Jerry Scott, Emvee Partners LLC
98. Jon McBride, BACS, MBA, CIO
99. Charles B. Owen, MD, CMIO, Afoundria
100. Rick Pope, Cognitive Medical Systems

- 100+ signatures at http://YosemiteManifesto.org/
- Led to Yosemite Project in 2014
Yosemite Manifesto
on RDF as a Universal Healthcare Exchange Language

1. RDF is the best available candidate for a universal healthcare exchange language.
2. Electronic healthcare information should be exchanged in a format that either: (a) is an RDF format directly; or (b) has a standard mapping to RDF.
3. Existing standard healthcare vocabularies, data models and exchange languages should be leveraged by defining standard mappings to RDF, and any new standards should have RDF representations.
4. Government agencies should mandate or incentivize the use of RDF as a universal healthcare exchange language.
5. Exchanged healthcare information should be self-describing, using Linked Data principles, so that each concept URI is de-referenceable to its free and open definition.
"1. RDF is the best available candidate for a universal healthcare exchange language."

• Several reasons:
  – Self describing
  – Easy to map from other data representations
  – Captures information content instead of syntax
  – Multi-schema friendly
  – Enables inference

• See: Why RDF as a Universal Healthcare Exchange Language?
  http://dbooth.org/2014/why-rdf/
"2. Electronic healthcare information should be exchanged in a format that either: (a) is an RDF format directly; or (b) has a standard mapping to RDF."

• Q: Convert all data to RDF format?
• A: No! Convert only:
  – If recipient does not understand the sender's data format or semantics; or
  – To determine the data's normative meaning
"3. Existing standard healthcare vocabularies, data models and exchange languages should be leveraged by defining standard mappings to RDF, and any new standards should have RDF representations."

- Allows RDF to act as a universal information representation across all healthcare information standards
RDF as a common semantic representation

```
<Observation ...>
  <system value="http://loinc.org"/>
  <code value="3727-0"/>
  ...
</Observation>
```

```
HL7 v2.x
OBX|1|CE|3727-0^BP|systolic, sitting|120|mmHg|
```
"4. Government agencies should mandate or incentivize the use of RDF as a universal healthcare exchange language."

- Healthcare providers and vendors have no natural business incentive to make their data interoperable to others
"5. Exchanged healthcare information should be self-describing, using Linked Data principles, so that each concept URI is de-referenceable to its free and open definition."

• Clickable URIs for concepts:
  – Easy to find the definition
• "Free and open definition":
  – Encourages interoperability
How would RDF work as a universal healthcare exchange language?
If sender and receiver speak the same format and semantics . . .

I have HL7 v2.x

I want HL2 v2.x!

HL7 v2.x

OBX|1|CE|3727-0^BpSystolic, sitting||120||mmHg|

No need for translation
If sender and receiver speak different format or semantics . . .

I have HL7 v2.x

I want FHIR!

Translation needed!
Lift and Drop

- Lift: Maps to RDF
- Drop: Maps from RDF
- Simple syntactic translation
- Retains data models and vocabularies
If Sender and Receiver use the same data model and vocabularies:

- Translate HL7 v2.x to RDF
- Translate RDF to FHIR
Translation with semantic alignment

- Usually semantic alignment is required
  - RDF-to-RDF translation
  - Done with SPARQL rules or other methods
- RDF acts as a universal information representation
  - Enables sharable translation rules
Same information, but different data models and vocabularies

- Both are RDF
- RDF supports inference
  - Good for model and vocabulary translation
Workshop Agenda
21-Aug-2014


• [11:35] Panel Discussion – All
BACKUP SLIDES
What if the data does not map?

- Requires intervention
- Can display RDF-enabled default view
Standard Vocabularies in UMLS

AIR ALT AOD AOT BI CCC CCPSS CCS CDT CHV COSTAR CPM CPT CPTSP CSP CST DDB DMDICD10 DMDUMD DSM3R DSM4 DXP FMA HCDT HCPCS HCPT HL7V2.5 HL7V3.0 HLREL ICD10 ICD10AE ICD10AM ICD10AMAE ICD10CM ICD10DUT ICD10PCS ICD9CM ICF ICF-CY ICPC ICPC2EDUT ICPC2EENG ICPC2ICD10DUT ICPC2ICD10ENG ICPC2P ICPCBAQ ICPCDAN ICPCDUT ICPCFIN ICPCFRE ICPCGER ICPCHEB ICPCCHUN ICPCITA ICPCNOR ICPCPOR ICPCSPA ICPCSWE JABL KCD5 LCH LNC_AD8 LNC_MDS30 MCM MEDLINEPLUS MSHCZE MSHDUT MSHFIN MSHFRE MSHGER MSHITA MSHJPN MSHLAV MSHNOR MSHPOL MSHPOR MSHRUS MSHSCR MSHSPA MSHSWE MTH MTHCH MTHHH MTHICD9 MTHICPC2EAE MTHICPC2ICD10AE MTHMST MTHMSTFRE MTHMSTITA NAN NCISEER NIC NOC OMS PCDS PDQ PDNS PPAC PSY QMR RAM RCD RCDAE RCDSA RCDSY SNM SNMI SOP SPN SRC TKMT ULT UMD USPMG UWDA WHO WHOFRE WHOGER WHOPOR WHOSPA
Standard Vocabularies in UMLS

Over 100!
How Standards Proliferate:
(See: A/C chargers, character encodings, instant messaging, etc.)

SITUATION: There are 14 competing standards.

14?! Ridiculous! We need to develop one universal standard that covers everyone's use cases. Yeah!

SITUATION: There are 15 competing standards.
Why does this happen?

And what can we do about it?
Why do standards proliferate?

1. Problem complexity

• Healthcare domain is huge
  – Many medical specialties
  – Administrative and business aspects
  – Research, biology, chemistry, etc.
  – Connects with everything else!

• Need the ability to represent:
  – Any data model
  – Any vocabulary
  – Any granularity
Why do standards proliferate?

1. Problem complexity

- Infeasible to standardize everything at once

- Need to divide and conquer
  - Standardize first, and interconnect later
Standards trilemma: Pick any two

- Comprehensive: Handles all use cases
- Good: High quality
- Timely: Completed quickly
Why do standards proliferate?

2. Diverse Requirements

• Different uses need:
  – Different data
  – Different granularity of data

• No such thing as a perfect standard
WARNING: Data users are myopic

• Each user thinks his/her use case is the most important
  – Ignores other use cases
The granularity dilemma

Different uses want different granularities!

Simplicity!
The granularity dilemma

Different uses want different granularities!
Need the ability to:

- Use all available information
- Ignore unwanted information
Why do standards proliferate?

3. Changing requirements

• Healthcare changes
• Technology changes
• New standards address deficiencies of old ones

Need the ability to:
• Continuously accommodate new standards/versions
• Relate old and new
• Translate between them
Why do standards proliferate?

4. Misaligned Incentives

• Proprietary interests
• Proprietary "standards"

Need:
• Free and open data models & vocabularies
• International, vendor-neutral