

# **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

- [8:30] **RDF as a Universal Healthcare Exchange Language** -- David Booth – Slides: <http://dbooth.org/2014/rdf-as-univ/>
- [8:55] **Why RDF?** -- David Booth – Slides: <http://dbooth.org/2014/why-rdf/>
- [9:10] **The Ideal Medium for Health Data? A Dive into Lab Tests** – Conor Dowling – Slides: <http://schemes.caregraf.info/presentations/semtech2014/>
- [9:30] **Introduction and RDF Representation of Fast Healthcare Interoperability Resources (FHIR) for Clinical Data** – Josh Mandel – Slides: <http://bit.ly/fhir-semtech-2014>
- [9:45] **Transformations for Integrating VA data with FHIR in RDF** – Rafael Richards – Slides: <http://dbooth.org/2014/richards/>
- [10:30] **Towards a Web of Clinical Knowledge** – Claude Nanjo – Slides: <http://dbooth.org/2014/nanjo/>
- [10:50] **Data-Driven Biomedical Research with Semantic Web Technologies** – Michel Dumontier – Slides: <http://dbooth.org/2014/dumontier/>
- [11:15] **The Yosemite Project: A Roadmap for Healthcare Information Interoperability** -- David Booth – Slides: <http://dbooth.org/2014/yosemite/>
- [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 Curé, PhD, UPEM France
8. Emory Fry, MD, Cognitive Medical Systems
9. Karl Seiler, CEO and founder NIMO Health, a Modus Operandi, Inc. Business
10. Erick Von Schweber, Executive Co-chair SURVEYOR health "Endorses RDF as a universal exchange framework"
11. Tom Munnecke, Independent Consultant
12. Thomas J. Kelly, PMT, Cognizant Technology Solutions
13. Dean Allemang, PhD, Working Ontologist LLC
14. Erich A. Gombocz, CSO, IO Informatics, Inc.
15. Blair Myers, Sr. Enterprise Information Architect, STA Group, LLC
16. Hans Constandt, CEO ONTOFORCE, Gent (Belgium)
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. Kerstin Forsberg, Principal Informatics Scientist, AstraZeneca
21. Niklas Lindström, senior developer, National Library of Sweden
22. Mark Montgomery, Founder & CEO, Kyield
23. Karl Reti, CEO, Crosslink Software
24. David L. Woolfenden President, eVectis Technologies LLC
25. Matthew Vagnoni, MS, CTO KnowMED.com
26. Christopher Regan
27. Doug Burke, President, Cognitive Medical Systems
28. Jerry Scott, Emcee Partners LLC
29. Rick Pope, Cognitive Medical Systems
30. Charles B. Owen, MD, CMIO, Afoundria
31. Conor Dowling, CTO, Caregraf
32. James McCusker, Yale University
33. Cartik Kothari, PhD, CEO, Perfect Informatics, IN/DIA
34. Carl Matlocks, Founder, Wellness Intelligence Institute
35. Lee Feigenbaum, VP and Founder, Cambridge Semantics
36. Jamie Ferguson, VP Health IT Policy, Kaiser Permanente
37. Christian Seppa, Senior Developer, Squishymedia Inc.
38. Dr. Matthias Samwald, Medical University of Vienna
39. Michael Ulschold, PhD, Senior Ontology Consultant, Semantic Arts, USA
40. Jon McBride, BACS, MBRA, CIO
41. Kathrin Dentler, PhD student, VU University Amsterdam & University of Amsterdam
42. Claude Nijmjo, MA MPH, Zynex Health Inc
43. Murray Bent, e-researcher
44. Pedro Lopes, PhD, University of Aveiro
45. Sibi Jacob, Senior Information Analyst, Ramsay Healthcare
46. Carlton Northern, Senior Software Engineer, The MPTRE Corporation
47. Michael Denny, PhD, ontology consultant
48. Robert Stanley, CEO, IO Informatics
49. Renato Iannella, PhD, Semantic Identity
50. Junica Kite MBBS, MEd, A.I.M. Consulting Ltd, UK
51. Jeff Nimmo, co-founder, Ugly Research
52. Stephane Jellali, CTO, smartKalen LLC
53. Frank van Harmelen, Prof., VU University Amsterdam

54. Tim Finin, Professor, University of Maryland, Baltimore County
55. François Schaffré, Maître de conférences, Université Montpellier 2
56. Varish Mulkwad, PhD candidate, Computer Science, UMRBC
57. Deborah M Cooper, Principal, Deborah M Cooper Consulting LLC
58. Joana S. Luciano, BSc MSc PhD, Research, Associate Professor, Rensselaer Polytechnic Institute, President, Predictive Medicine, Inc.
59. M. Scott Marshall, PhD, MARSFRO Clinic, Maastricht, The Netherlands
60. Xalina Bontcheva, PhD, University of Sheffield
61. Alan Rasmberg, Director of Data Warehouse at Institute for Health Informatics, University at Buffalo
62. Dan Brickley, Google
63. Kristina Kumar Karkal, MS, KnowMED Incorporated.
64. Sergey Krikov, MS, University of Utah
65. Shelly Kulesza, Project Manager, KnowMED
66. Sofia T. Amiri, MEd, MS, KnowMED Inc.
67. Xuyi Mogul, healthcare software
68. Mary Dee Harris, PhD, independent consultant
69. David Cesar, PhD, University of Aberdeen, UK
70. Christophe Lambert, PhD, Golden Medix Inc.
71. Javier Fernandez Iglesias, Independent Consultant, Spain
72. Paolo Ciccarelli, MS PhD, Harvard Medical School
73. Michael Egan, MD, MD, Radiation Cancer Center
74. Mikael Erik, Ltd
75. Foster Carr, MD, Sciencelink.com
76. Sreyan Tildan, MPhil, MS
77. Silvia Waga, MD, IT Project Manager, Scientific Society of General Medicine - Belgium
78. Markus Schneider, Healthcare Data Analyst
79. Ted Slater, CTO, OpenHLS Consortium
80. Lucian Ionescu, CTO, HCL Group
81. Andrea Ispoglami, director, intelligiSoft
82. RJ Harris, The US, The Consortium
83. Erwin Schiffer, Switzerland, independent consultant
84. Erik Brown, MSc., Long Beach University
85. Alan T. Kuhl, MD, PhD, FRCR, FRCR (1982-2009)
86. Ralf Kapte, McMaster University
87. David Murali, Murali Computing
88. Sachin Bhatia, PhD, Stanford University
89. Marc Truongmanica, MD, PhD, Agile Healthcare N/A
90. Stuart Turner, PhD, MS, Logical Informatics
91. Hong Sun, PhD, Agile Healthcare
92. Avdi Zappa, PhD, FRCGP at NCI gateway - The Centre for Data Analytics
93. Yoshinori Kamezaki, MD, PhD
94. Barry Kelson, Original Architect of QJGEL
95. Graham Hughes, MD, PhD, Australia
96. Suresh Arundani, MD, MS, ONCOPHD
97. Sarah Bates, MS, Midcom
98. Salvatore Magagnoli, Bioinformatics, Study University
99. Benedikt Wittenberg, Research Associate, Karlsruhe Institute of Technology, Germany
100. Brian Nove, CTO, Ringier Technology LLC
101. Stefan Döcker, PhD, MARSFRO Clinic
102. Sébastien Lefebvre, PhD, Health Entrepreneur & Developer
103. Marcelo Bay, PhD, Federal University of Minas Gerais - Brazil
104. Natalia Dias Rodrigues, M. Sc., Philips Research
105. N. D. Patel, MD, ONCOE 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/>

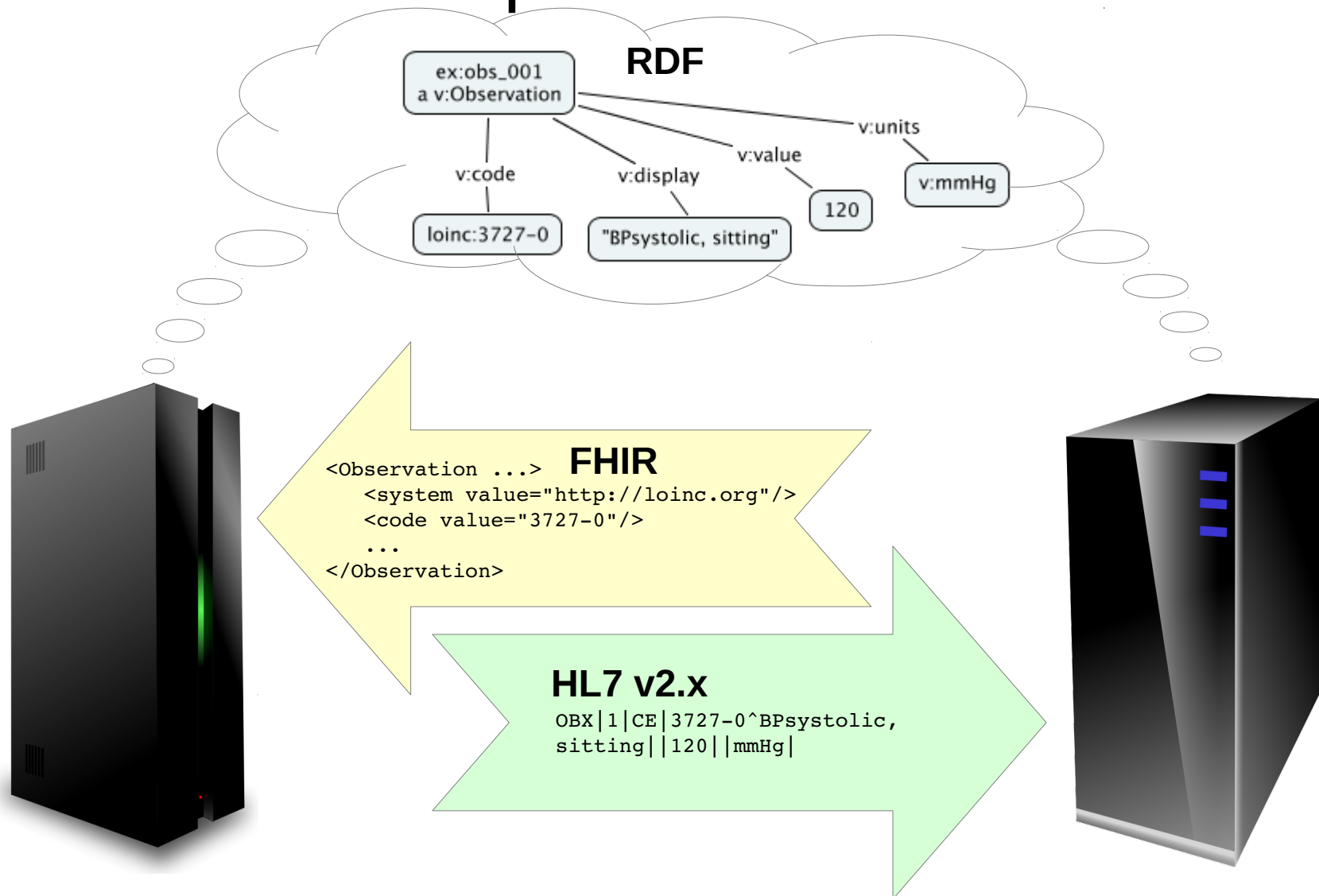
*"2. Electronic healthcare information should be exchanged in a format that either: (a) is an RDF format directly; or (b) has a standard mapping to RDF."*

- 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



*"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



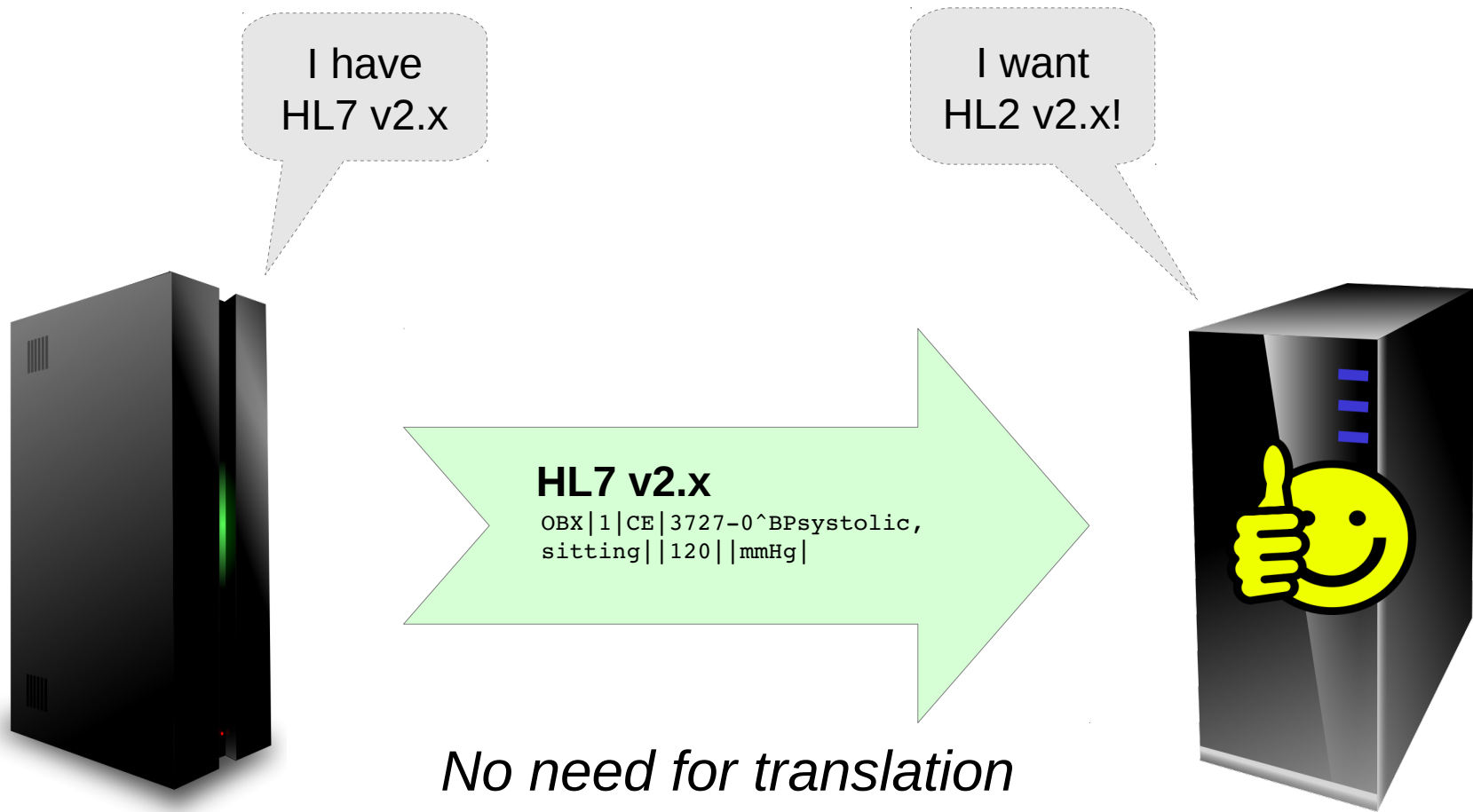
*"5. Exchanged healthcare information should be self-describing, using Linked Data principles, so that each concept URI is de-referenceable to its free and open definition."*

- Clickable URIs for concepts:
  - Easy to find the definition
- "Free and open definition":
  - Encourages interoperability

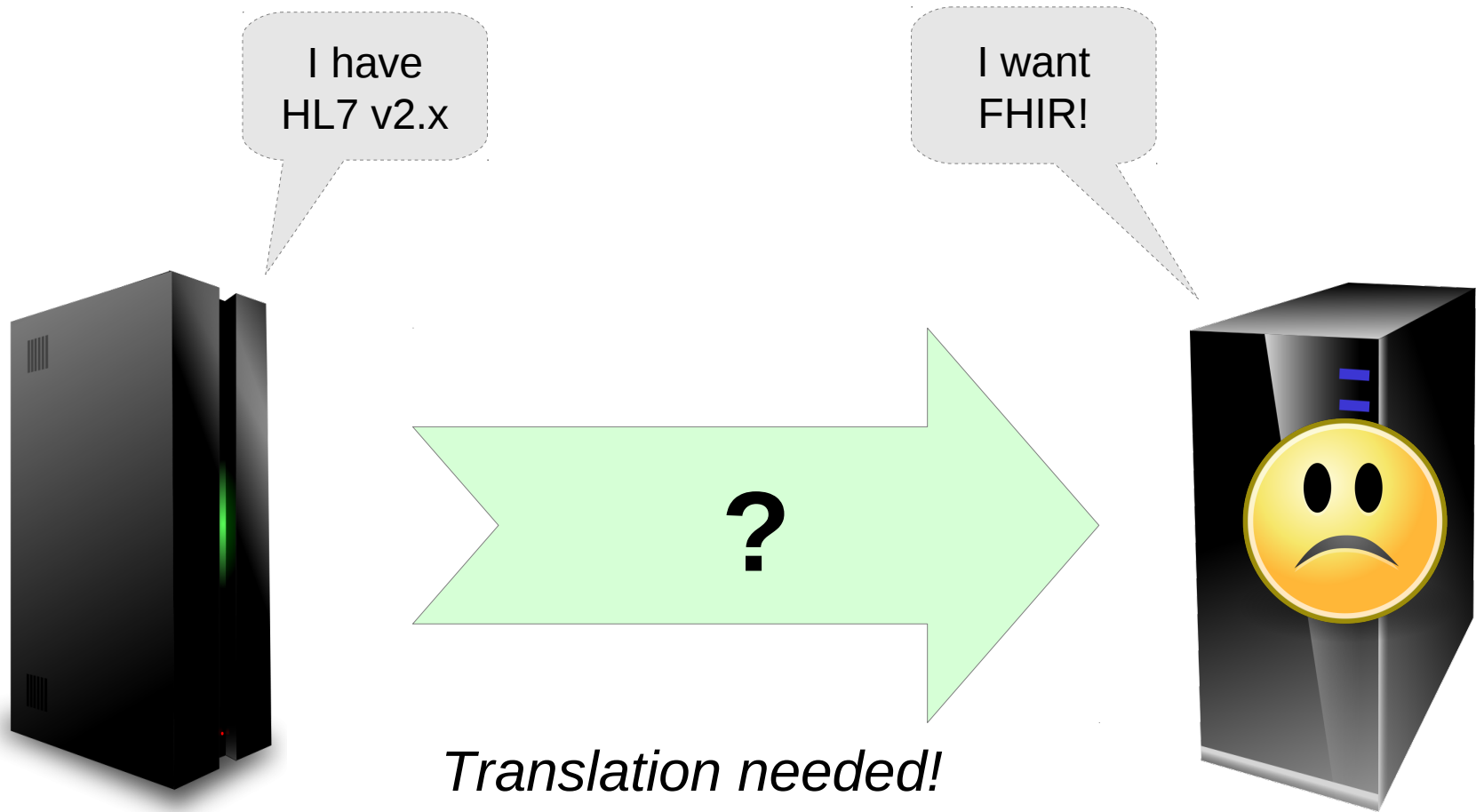


How would RDF work as a  
universal healthcare exchange language?

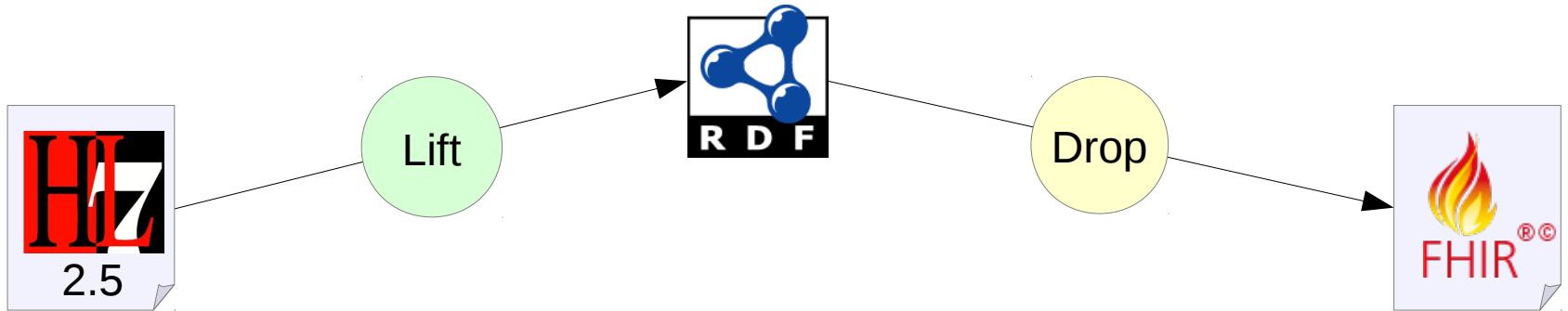
# If sender and receiver speak the same format and semantics . . .



# If sender and receiver speak different format or semantics . . .

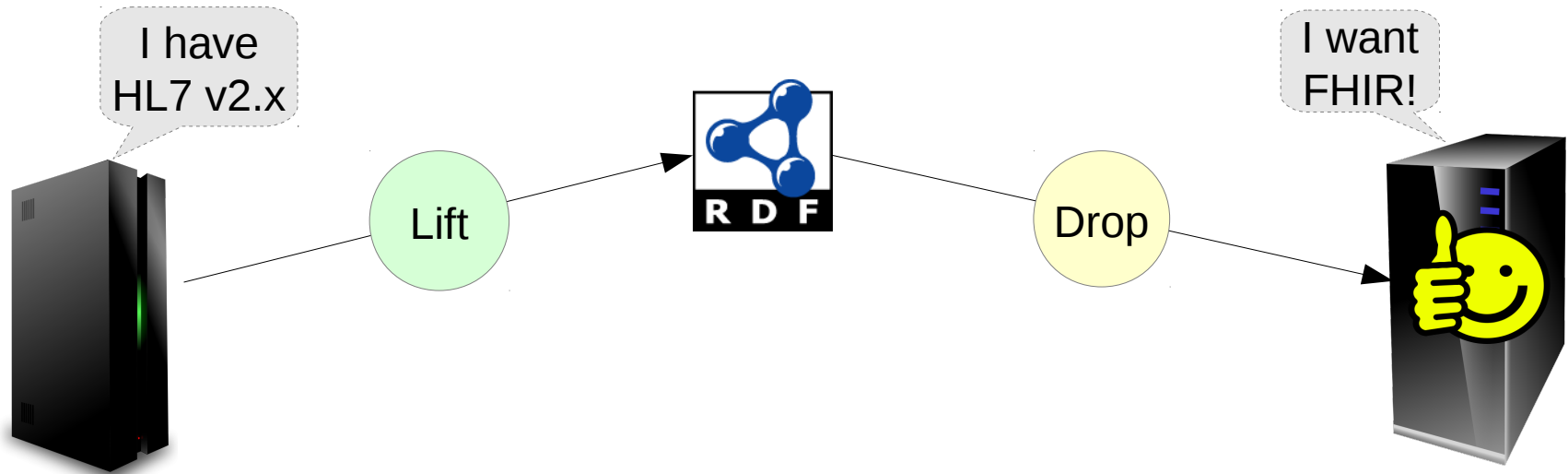


# Lift and Drop



- Lift: Maps to RDF
- Drop: Maps from RDF
- Simple syntactic translation
- Retains data models and vocabularies

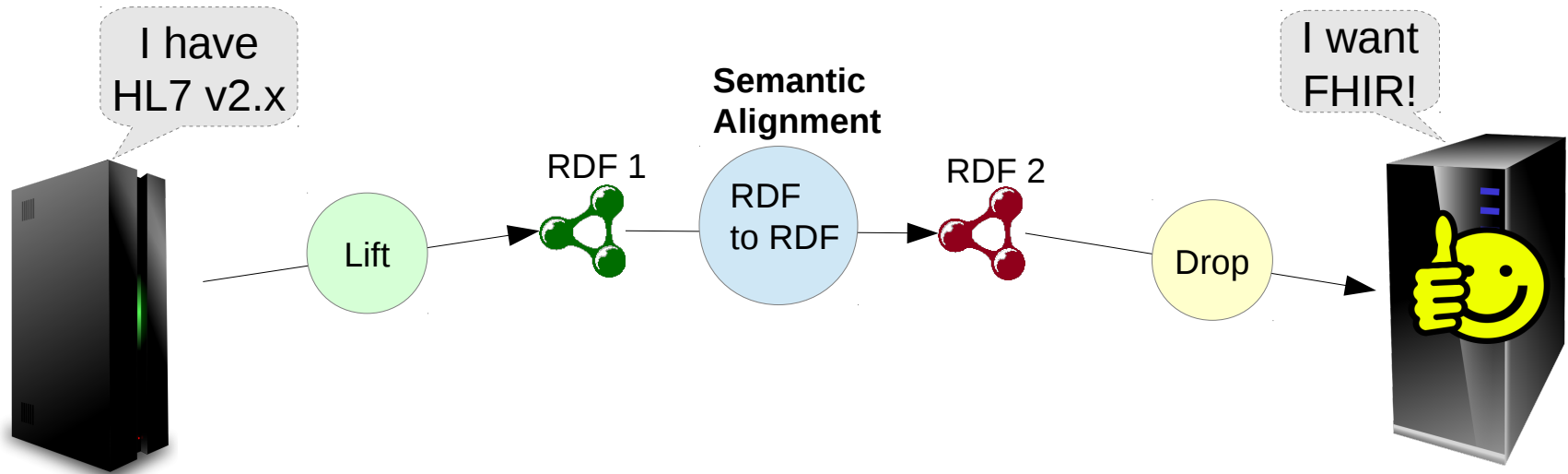
# Translation (Naive view)



If Sender and Receiver use the same data model and vocabularies:

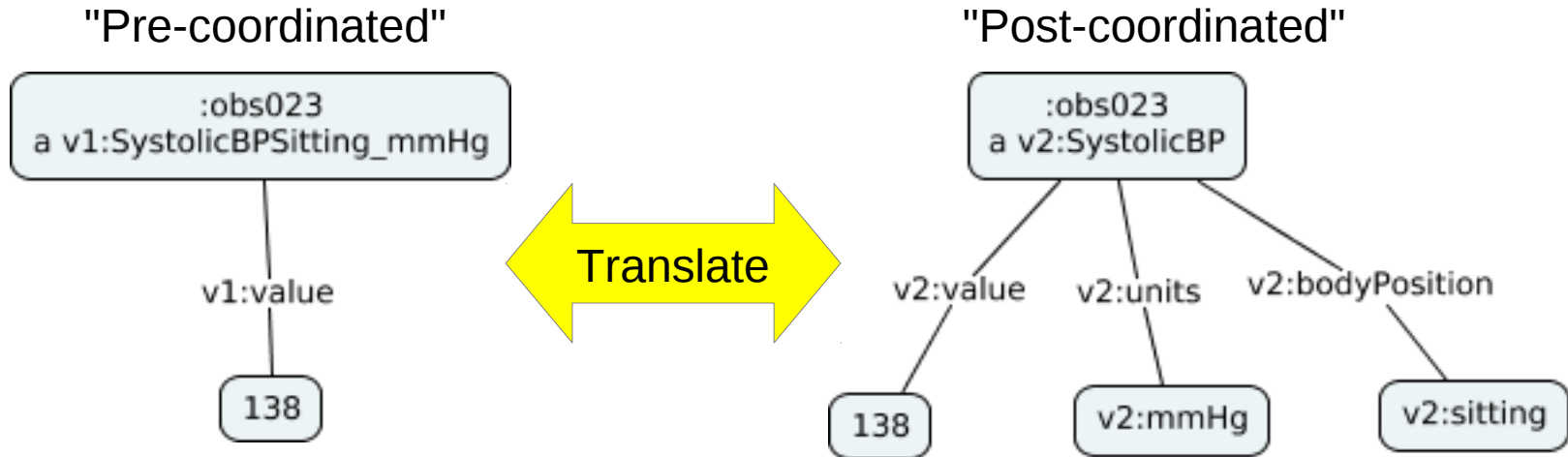
- Translate HL7 v2.x to RDF
- Translate RDF to FHIR

# 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

- [8:30] **RDF as a Universal Healthcare Exchange Language** -- David Booth – Slides: <http://dbooth.org/2014/rdf-as-univ/>
- [8:55] **Why RDF?** -- David Booth – Slides: <http://dbooth.org/2014/why-rdf/>
- [9:10] **The Ideal Medium for Health Data? A Dive into Lab Tests** – Conor Dowling – Slides: <http://schemes.caregraf.info/presentations/semtech2014/>
- [9:30] **Introduction and RDF Representation of Fast Healthcare Interoperability Resources (FHIR) for Clinical Data** – Josh Mandel – Slides: <http://bit.ly/fhir-semtech-2014>
- [9:45] **Transformations for Integrating VA data with FHIR in RDF** – Rafael Richards – Slides: <http://dbooth.org/2014/richards/>
- [10:30] **Towards a Web of Clinical Knowledge** – Claude Nanjo – Slides: <http://dbooth.org/2014/nanjo/>
- [10:50] **Data-Driven Biomedical Research with Semantic Web Technologies** – Michel Dumontier – Slides: <http://dbooth.org/2014/dumontier/>
- [11:15] **The Yosemite Project: A Roadmap for Healthcare Information Interoperability** -- David Booth – Slides: <http://dbooth.org/2014/yosemite/>
- [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 ICPCHUN 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 PNDS 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

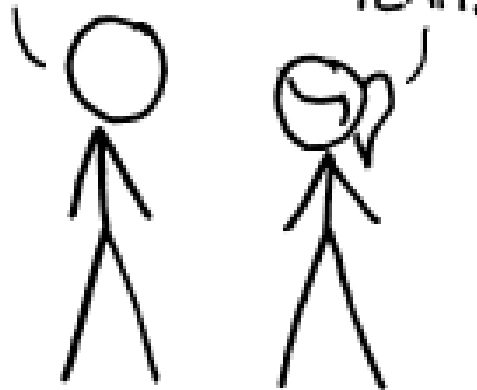
AIR ALT AOD AOT BI CCC CCPSS CCS CDT CHV COSTAR CPM  
CPT CPTSP CSP CST DDB DMDICD10 DMDUMD DSM3R DSM4 DXP  
FMA HCDDT HCPCS HCPT HL7V2.5 HL7V3.0 HLREL ICD10 ICD10AE  
ICD10AM ICD10AMAE ICD10CM ICD10DUT ICD10PCS ICD9CM ICF  
ICF-CY ICPC ICPC2EDUT ICPC2EENG ICPC2ICD10DUT  
ICPC2ICD10ENG ICPC2P ICPCBAQ ICPCDAN ICPCDUT ICPCFIN  
ICPCFRE ICPCGR ICPCHE ICPCJUN ICPCITA ICPCNOR ICPCPOR  
ICPCSPA ICPCWMAE ICPC5 LCH LNCAD LNCMMS30 MCM  
MEDLINEPLUS MSHOE MSHDUT MSHIN 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 PNDS PPAC PSY QMR RAM RCD  
RCDAE RCDSA RCDSY SNM SNMI SOP SPN SRC TKMT ULT UMD  
USPMG UWDA WHO WHOFRE WHOGER WHOPOR WHOSPA

# 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.



SOON:

SITUATION:  
THERE ARE  
15 COMPETING  
STANDARDS.

<http://xkcd.com/927/>  
Used by permission

**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?

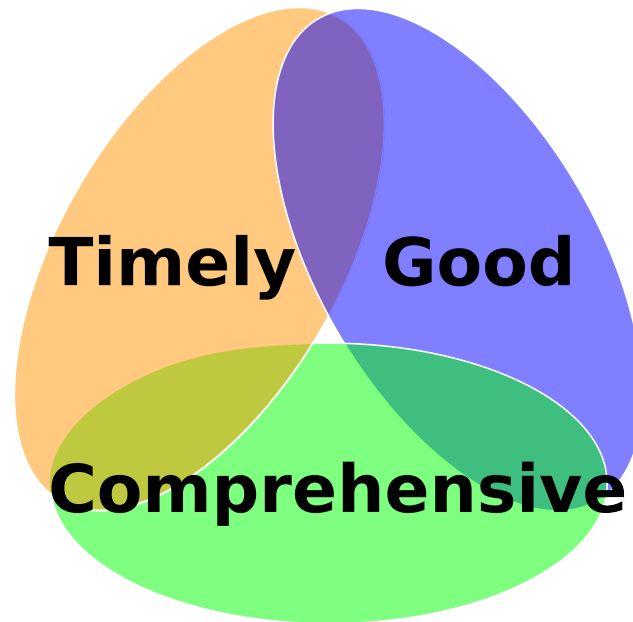
# 1. Problem complexity

- 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?

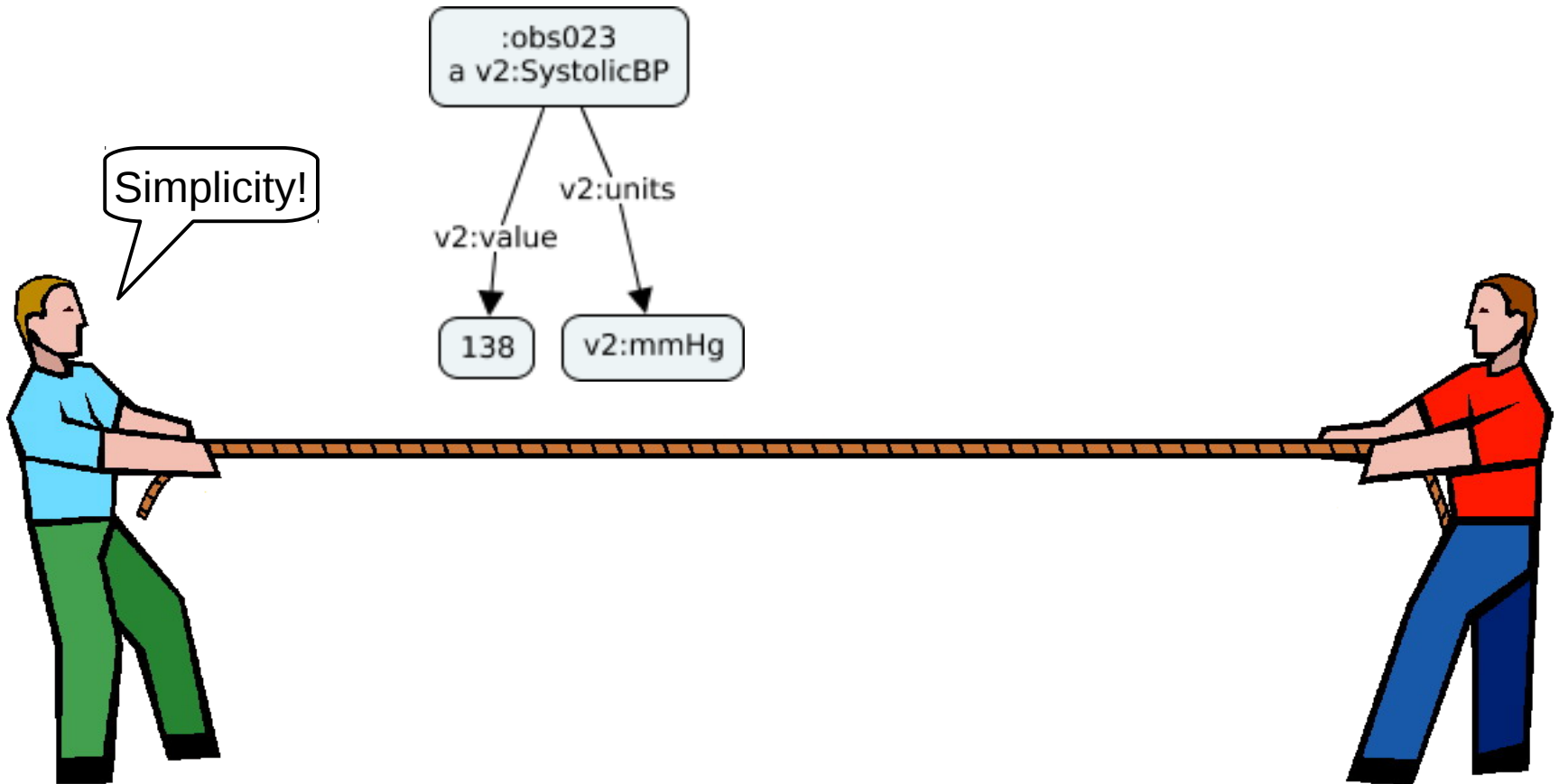
## 2. Diverse Requirements

- Different uses need:
  - Different data
  - Different granularity of data
- No such thing as a perfect standard

# WARNING: Data users are myopic

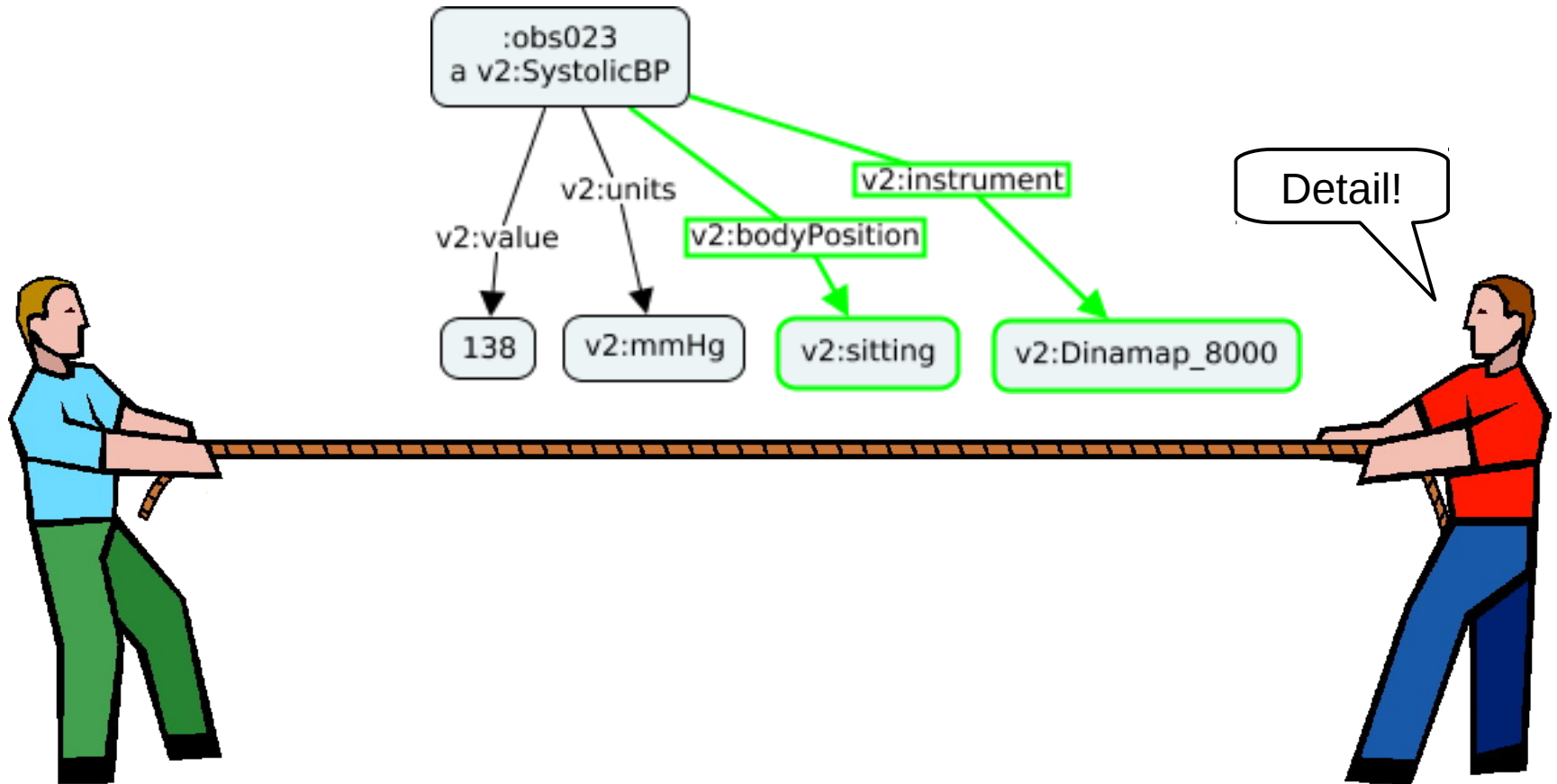
- Each user thinks his/her use case is the most important
  - Ignores other use cases

# The granularity dilemma



*Different uses want different granularities!*

# 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*