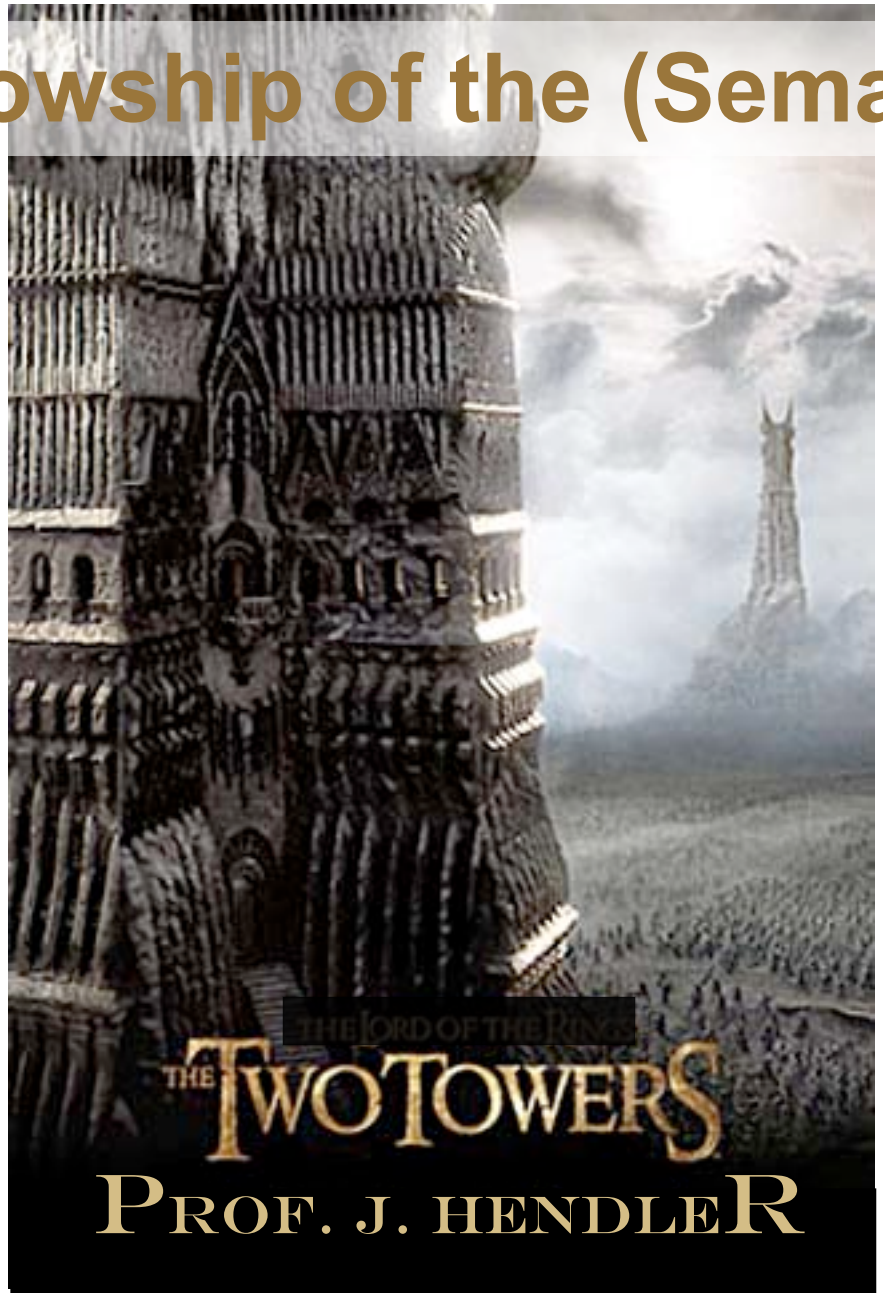
The Fellowship of the (Semantic) Web



Prof. J. Hendler

RPI

Hendler @ cs.rpi.edu

Ontological Conundrum

- The progress of the Semantic Web has been hampered by significant confusion as to what an ontology, and especially a Web ontology is.
  - Two separate visions (or perhaps two end points on what are a continuum) have caused significant confusion
- And the confusion blurs an important message
  - Both uses have proven valuable in the real world!!
- My goal in this talk is to try to reduce this confusion
  - Which could be a real first for a college Professor!

Outline

- Intro (now done)
- some imprecise analogies for motivation
- The kinda technical stuff
- Some shameless self-promotion
- Boffo Conclusion
  - http://www.thefreedictionary.com/boffo

Rensselaer

Tetherless World

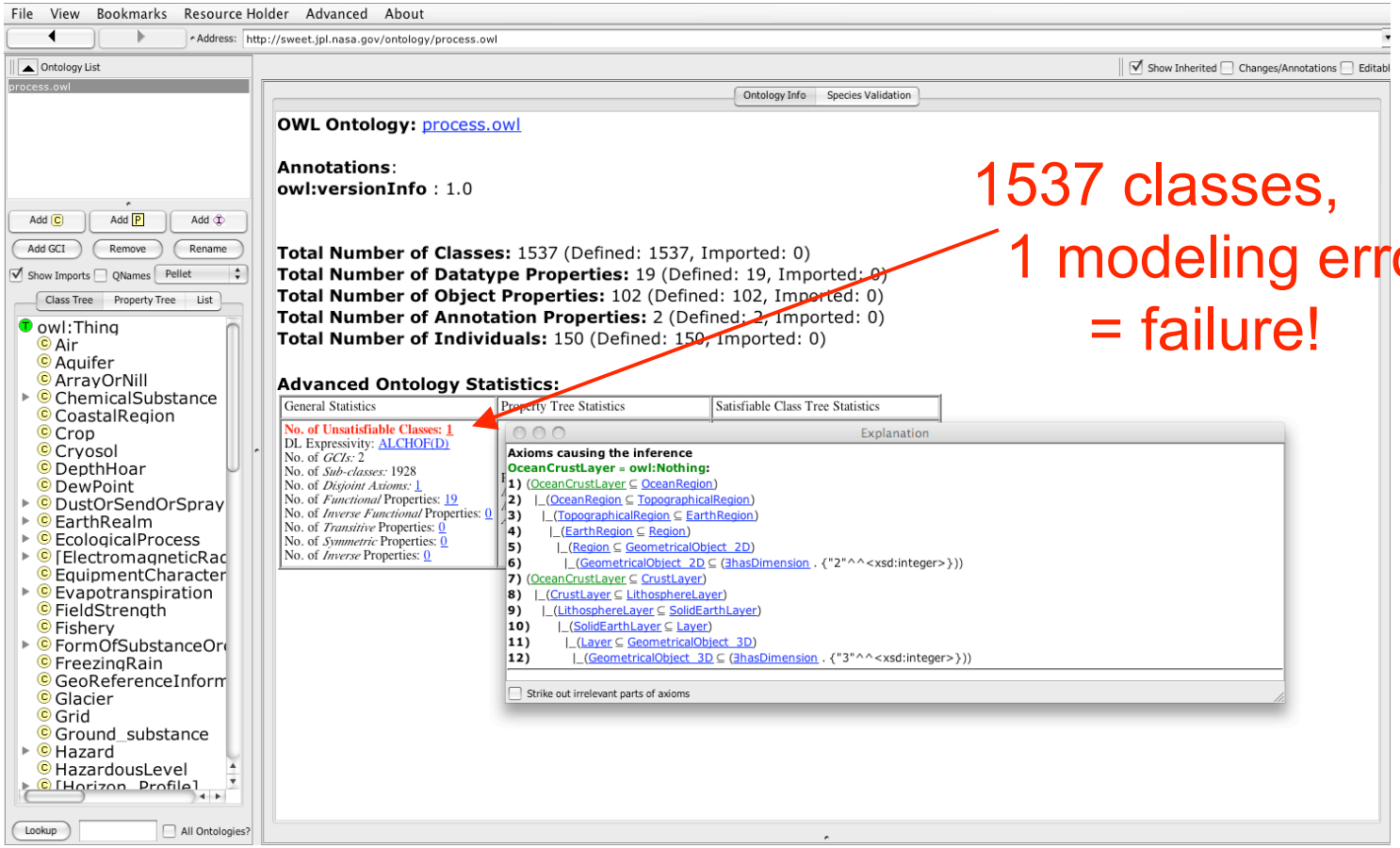
O*ntology*: the OWL DL view

- **O**ntology as Barad-Dur (Sauron's tower):
  - Extremely powerful!
  - ~~Patrolled by Orcs~~ Decidable Logic basis
    - Let one little ~~hobbit~~ inconsistency in, and the whole thing could come crashing down

Rensselaer

Tetherless World

Inconsistency is the bane of this view



(Swoop w/Pellet)

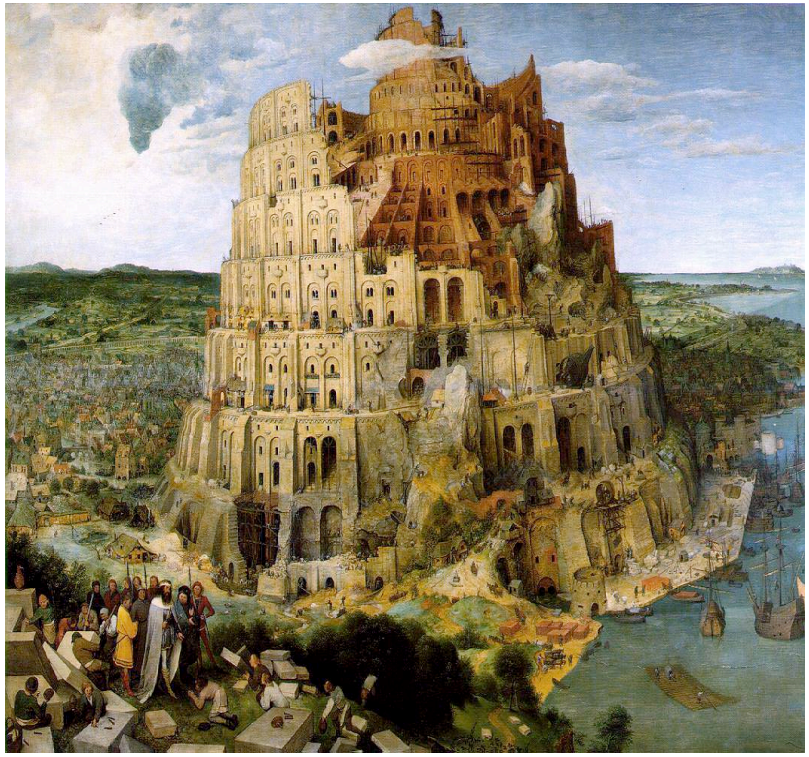
Rensselaer

Tetherless World

ROI: Reasoning over (Enterprise) data

- This "big O" Ontology finds use cases in verticals and enterprises
  - Where the vocabulary can be controlled
  - Where finding things in the data is important
- Example
  - Drug discovery from data
    - Model the molecule (site, chemical properties, etc) as faithfully and expressively as possible
    - Use "Realization" to categorize data assets against the ontology
      - Bad or missed answers are money down the drain

o*ntology*: the RDFS view



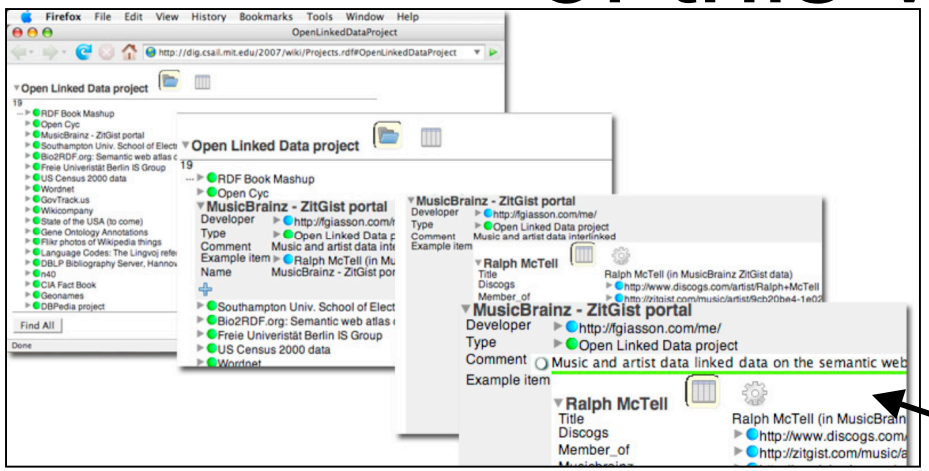
- **o**ntology and the tower of Babel
  - We will build a tower to reach the sky
  - We only need a little ontological agreement
    - Who cares if we all speak different languages?

Genesis 11:7 Let us go down, and there confound their language, that they may not understand one another's speech. So the Lord scattered them abroad from thence upon the face of all the earth: and they left off to build the city.

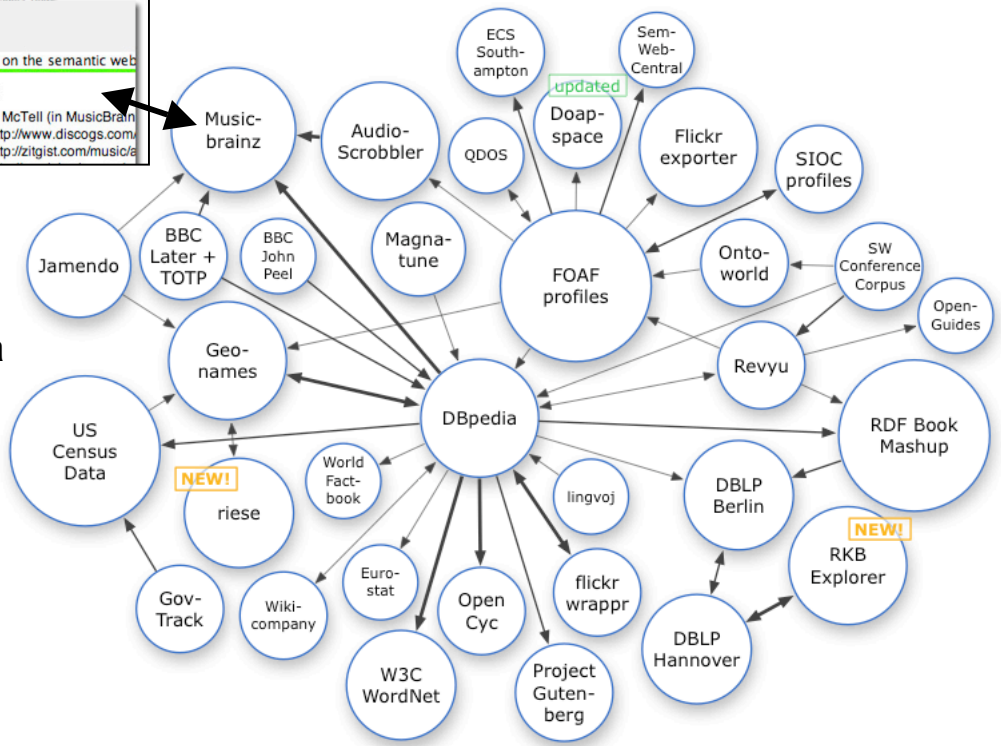
Rensselaer

Tetherless World

Boundaries are the bane of this view



Tabulator and Linked Open Data



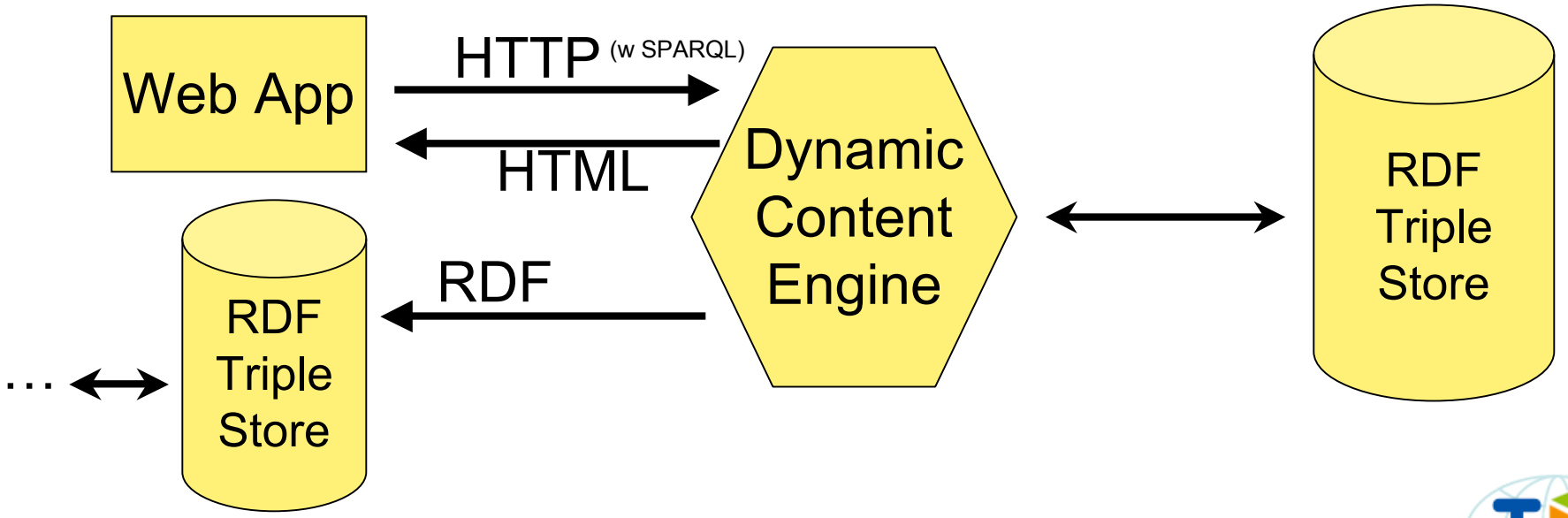
Rensselaer

ROI: Web 3.0

- The "small o" ontology finds use cases in Web Applications (at Web scales)
  - A lot of data, a little semantics
  - Finding anything in the mess can be a win!
- Example
  - Declare simple inferable relationships and apply, at scale, to large, heterogeneous data collections
    - *eg.* Use InverseFunctional triangulation to find the entities that can be inferred to be the same
      - These are "heuristics" not every answer must be right (qua Google)
      - But remember *time* = *money*!

ROI: Web 3.0

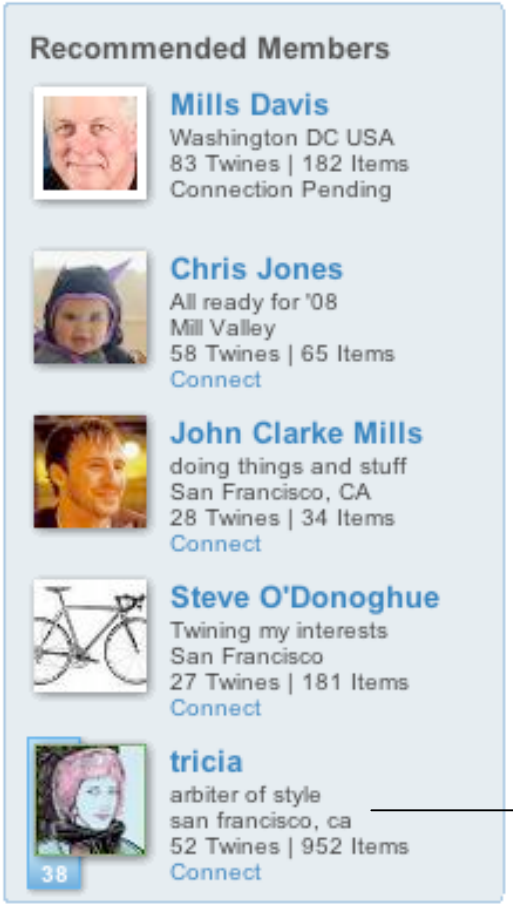
- ~2006: Web app developers discover the Semantic Web



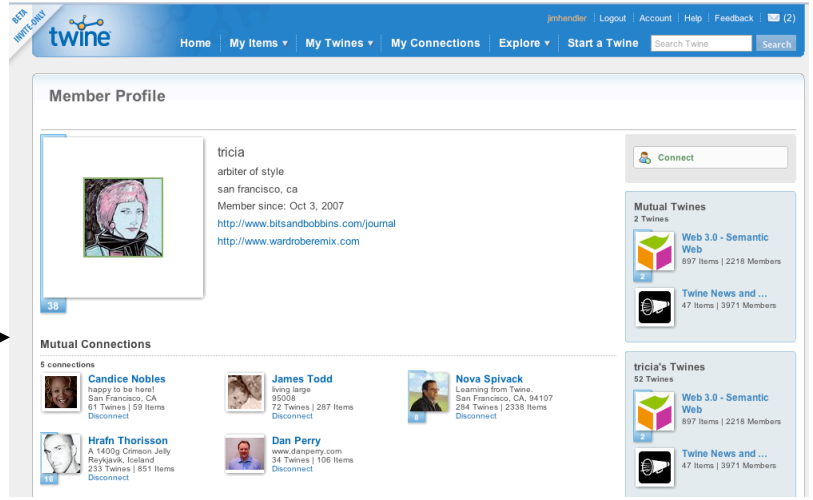
Rensselaer

Tetherless World

O asks o: how can you ignore soundness?



- Twine recommends some people I may want to connect to
  - What is correctness in this case?
    - If I find some folks I like this way, I use twine more. Surprises can be fun.
    - But if it does a "bad" job, I may go elsewhere

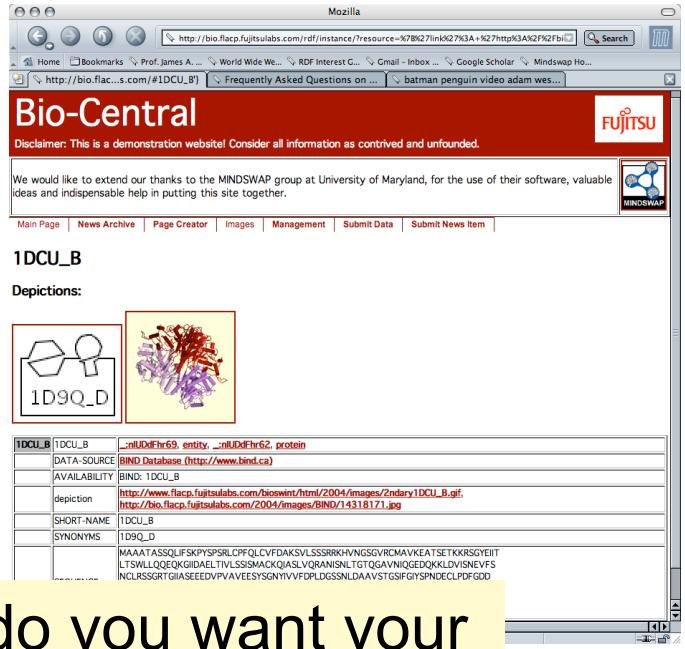


Rensselaer

Tetherless World

o asks O: Why do you need expressiveness?

- Often "folksonomy" isn't enough!



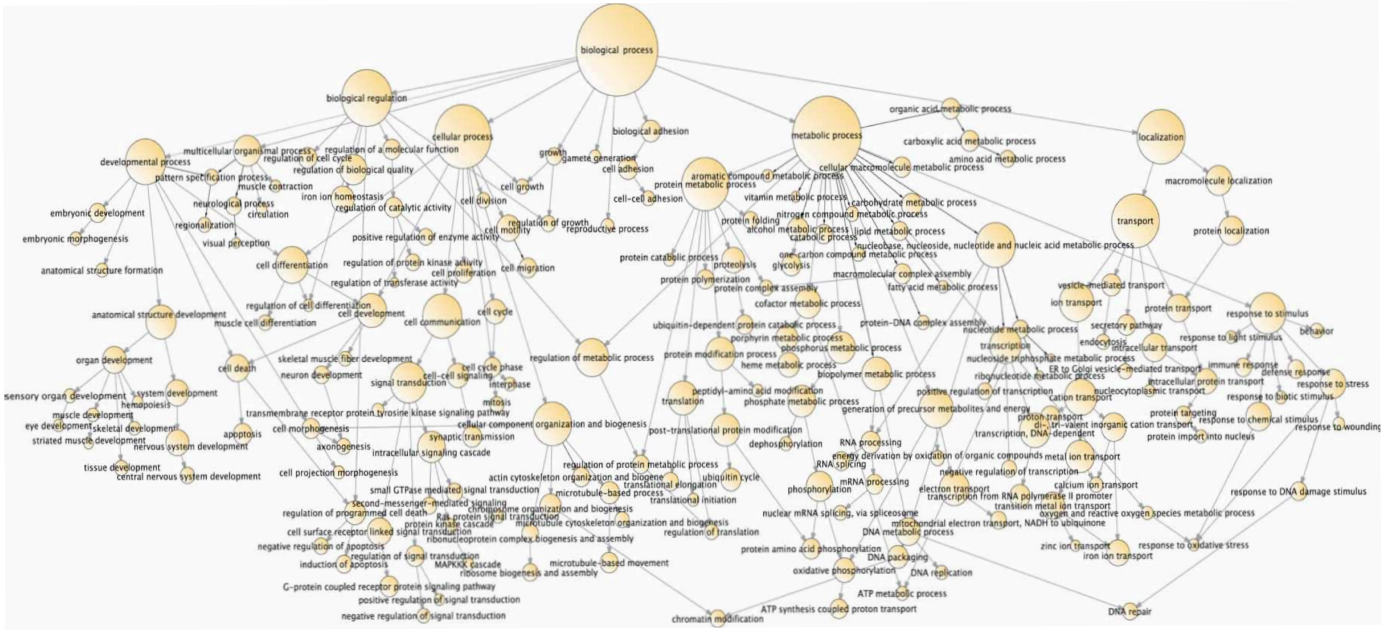
Which one do you want your doctor to use?

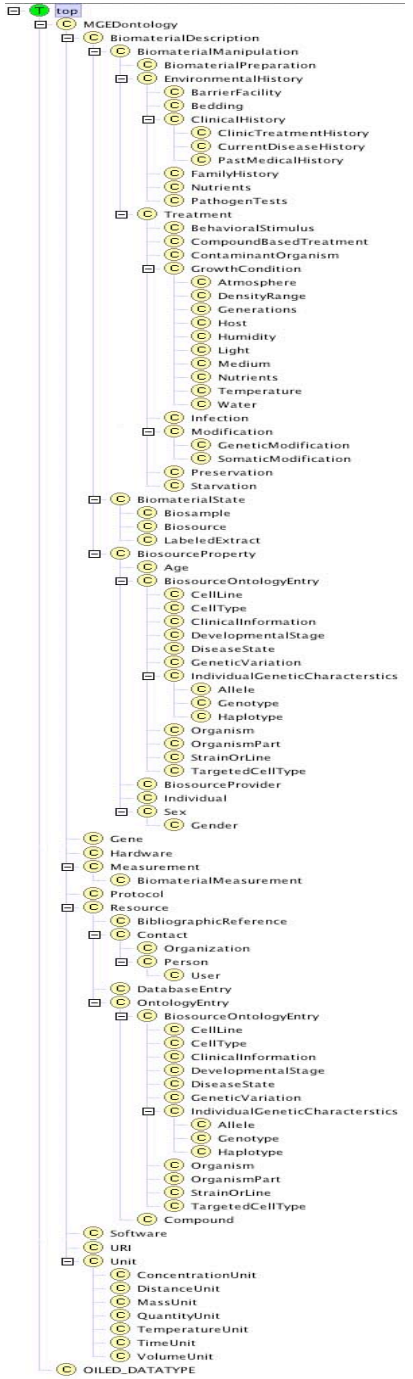
Rensselaer

Tetherless World

A big problem for O

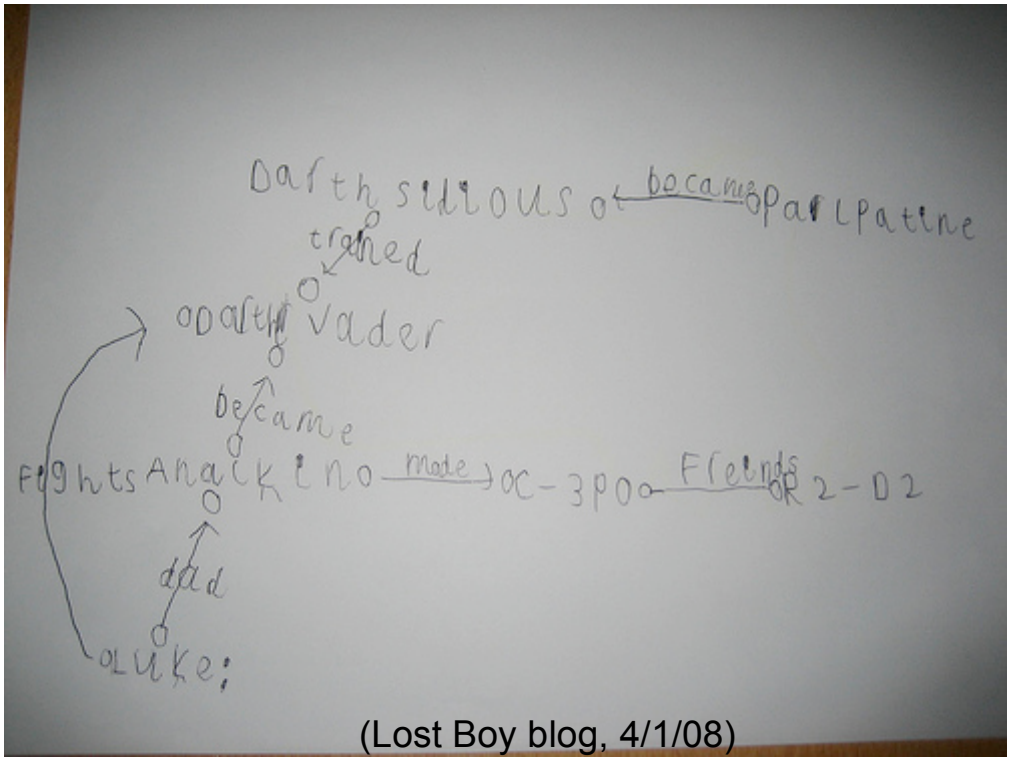
- Ontology mapping



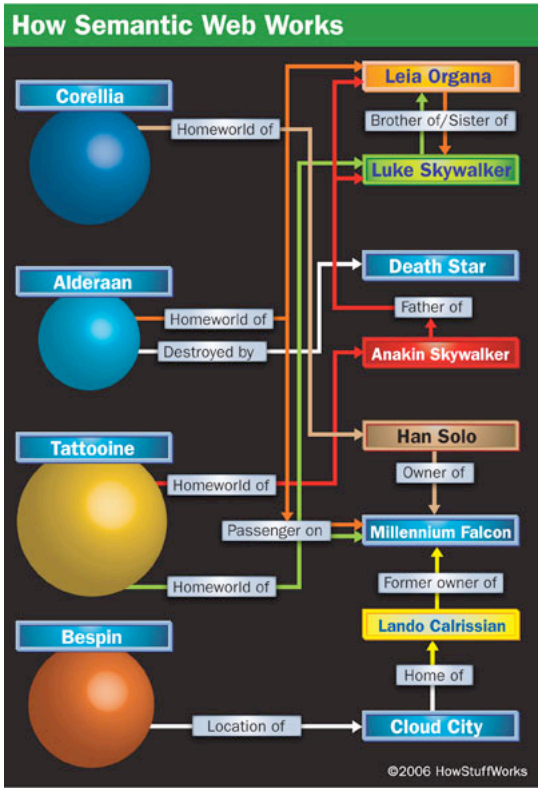


Rensselaer

Is not a big problem for o



(Lost Boy blog, 4/1/08)



Slogan: A little semantics goes a long way

Rensselaer

Tetherless World

A big problem for o

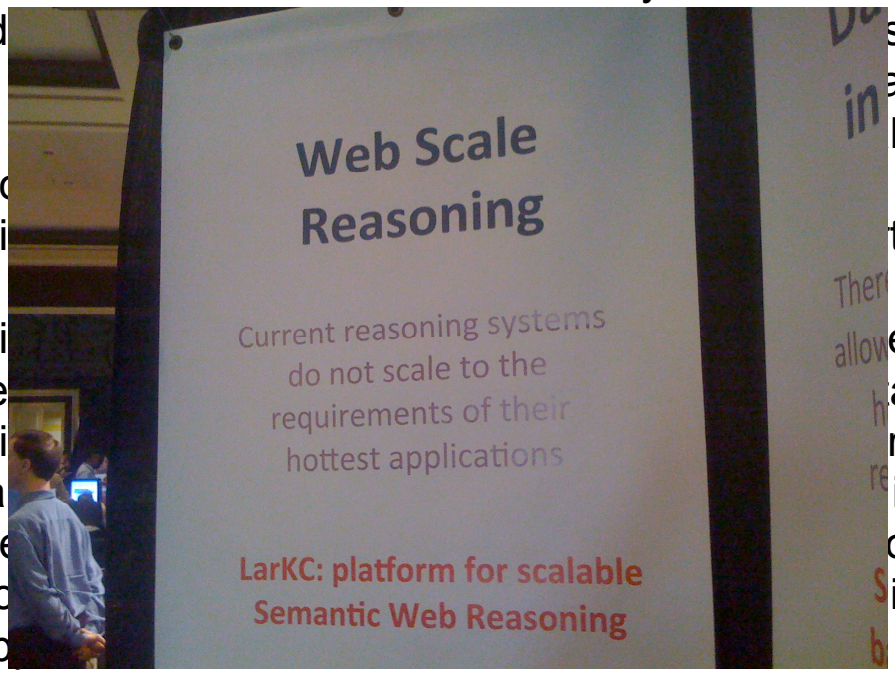
- What do we do with all this stuff?

* The primary goal is to for submissions to show how they add value to the very large triple store. This can involved s in the store via browsing, visualization, ation not directly queriable in the original l information could be tied to part(s) or the wh

* The tool or applicati tion of the data provided by the organizers.

* The tool or applicati ed to the target dataset, but there is still an expe a provided.

* The tool or applicati r application, as defined for the Open Track Cha is to demonstrate an interaction with the large owever, given the scale of this challenge, soluti ications, or as crucial to the success of future ap



(ISWC 2008 - Open Web, Billion Triple Challenge -

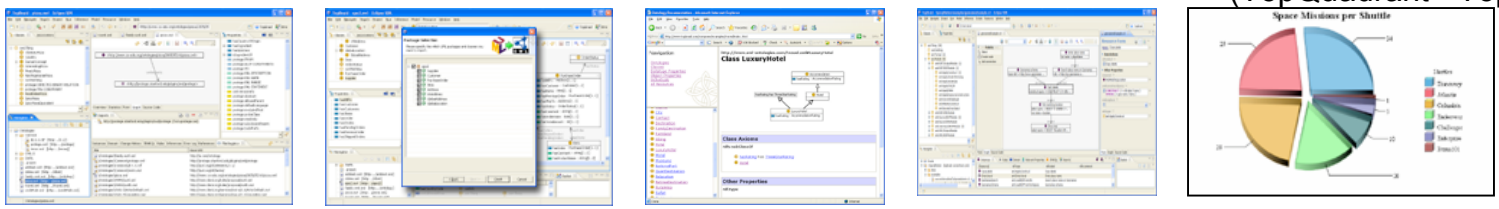
http://iswc2008.semanticweb.org/calls/call-for-semantic-web-challenge-and-billion-triples-tracks/

Rensselaer

Tetherless World

Is well understood in O

(TopQuadrant - TopBraid)



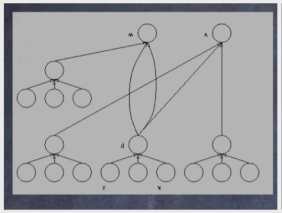
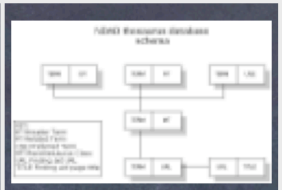
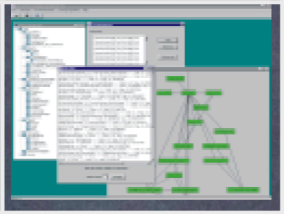
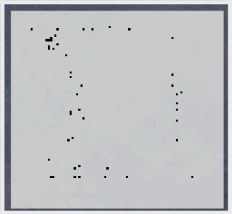
Slogan:Knowledge is power

Rensselaer

Tetherless World

We use the same word…

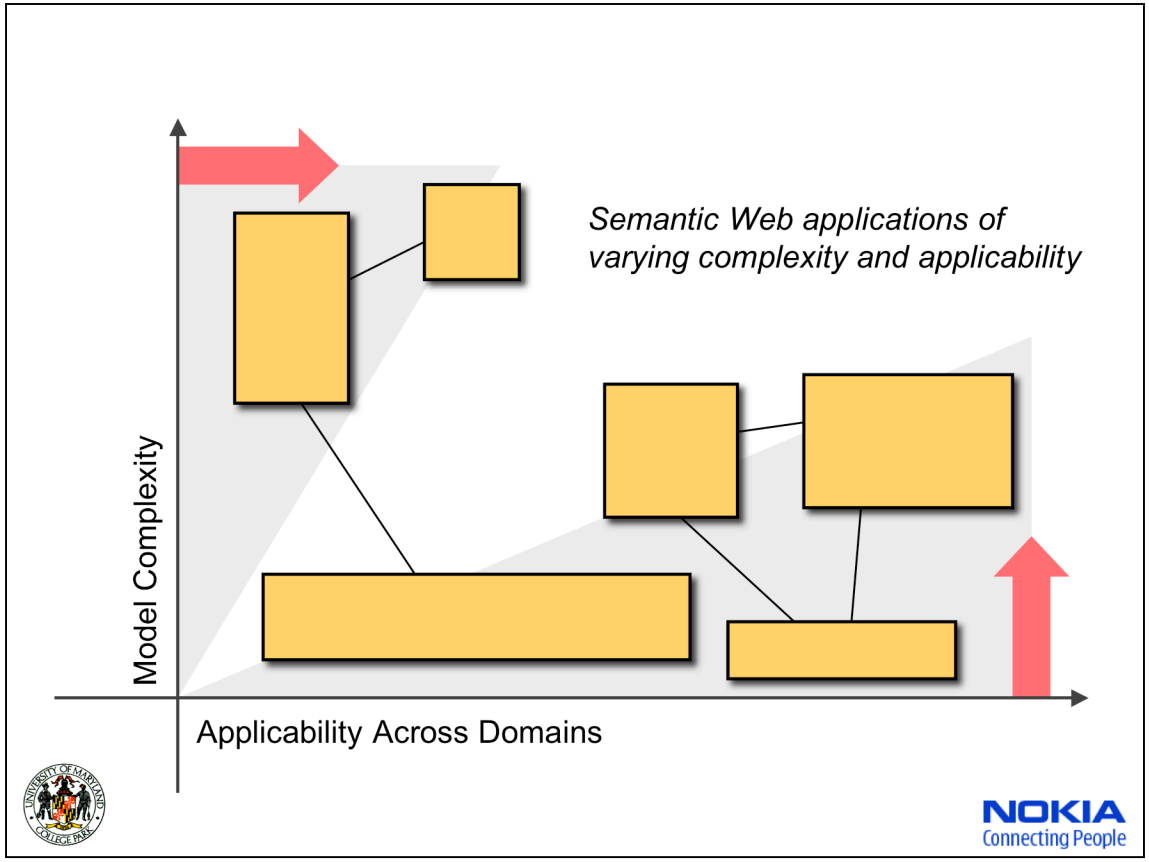
Sem Web Modeling

			…	
Graph Ontology	Labeled graph Data Dictionary Data Schema Ontology	Graph + limited logic Ontology	…	Logic Ontology
RDF	RDF Schema	OWL	…	KIF?

Rensselaer

Tetherless World

But O ≠ o



Rensselaer

Tetherless World

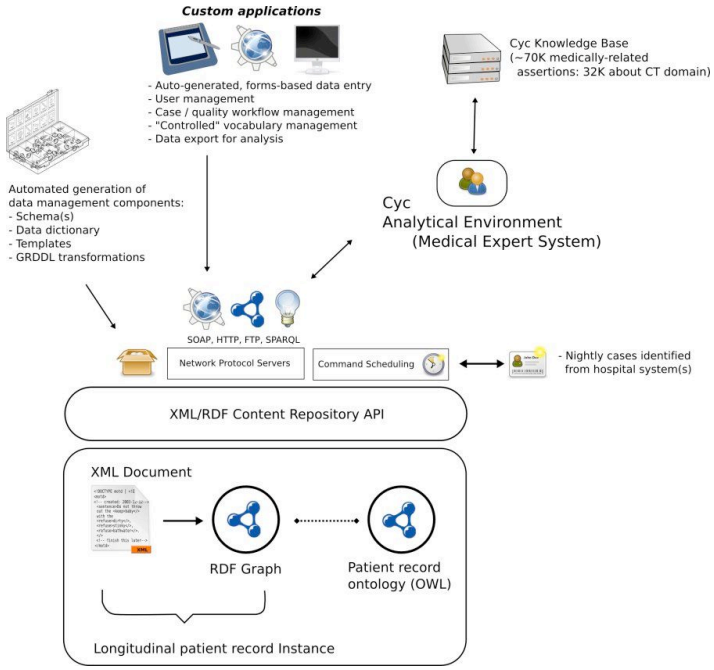
Why does this matter

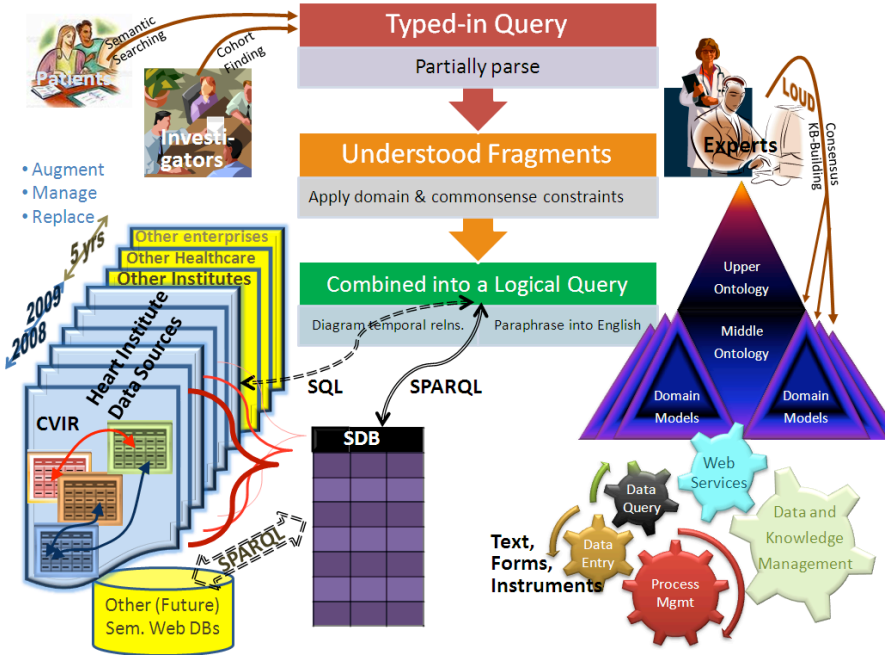
- Different issues of concern
  - Confuses messaging
- Effort is spent in different parts of the space
  - i.e. scaling vs. modeling
    - Leads to confusion in costs, esp. for interested parties
    - Starting out: You must know which O/o you're going after
- Different "first-concern" tools for the different models
  - Big O: ontology creation and modeling
  - Small o: triple store and SPARQL
- …

Tensions

- There are also some serious tensions between these models
  - Base in RDF (links) vs. XML (validation)
  - Soundness and Completeness
    - Big O: Mandatory
    - Small o: Impossible
  - Consistency impossible to maintain in large scale distributed efforts
    - Error, Disagreement, Fraud
  - Business Model
    - Enterprise v. Web Scale

Not Irreconcilable Differences





Cf. Cleveland Clinic "Semantic DB" effort

OR ≠ XOR

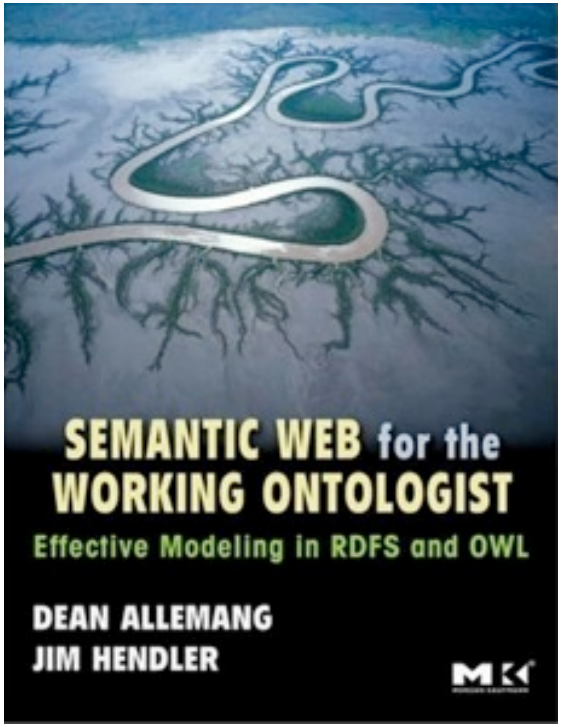
Rensselaer

Tetherless World

Which is why RDFS/OWL matter

- From the Original W3C OWL Faq*
  - Q. How is OWL different from earlier ontology languages?
  - A. OWL is a Web Ontology language.
    - Ability to be distributed across many systems
    - Scalable to Web needs
    - Compatible with Web standards for accessibility and internationalization.
    - Open and extensible
- Interoperability is lost if these two towers grow too far apart
  - It is important that RDFS and OWL remain WEB languages
    - Which is much harder in "O" than "o"
    - (W3C members: make sure your AC rep is watching the OWL 2 space with this in mind)

Where can I learn more?



http://www.amazon.com/Semantic-Web-Working-Ontologist-Effective/dp/0123735564

Rensselaer

Tetherless World

Or Ask the Experts

- Submit your questions on line for Tim, Nova, Wendy and Nigel
  - Vote on the ones you want to hear
- Watch the Webcast
- Or show up in person
  - We're buying the drinks!

http://tw.rpi.edu/launch/

Rensselaer

Tetherless World

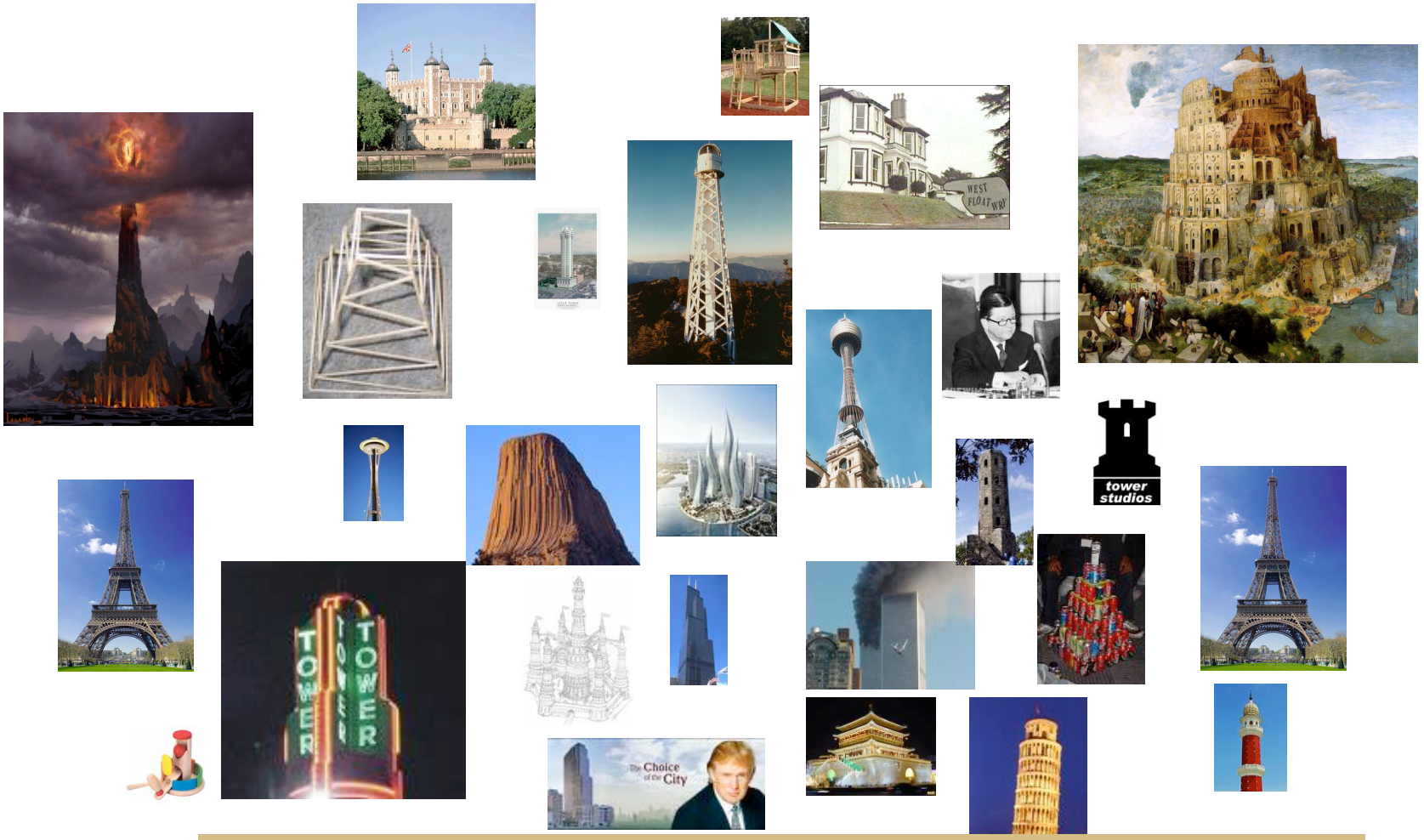
~~Boffo~~ Conclusion

- There's no ontology ontology
  - The term is used in many ways in our community
- This causes great confusion
  - And hurts our technology adoption
- The Web is a big place
  - Room for more than one vision
  - But they must play nice to get the network effect
- Which leads to…

Rensselaer

Tetherless World

… the Semantic **WEB** Vision



Interoperability trumps homogeneity ever time!

Rensselaer

Tetherless World