Ontology Summit 2014
Track A: Common, Reusable Semantic Content
Working Track A Synthesis

MikeBennett, GaryBergCross,
AndreaWesterinen

Summary

- Inputs:
  - Track A presentations from Jan 23 2014
  - Email dialogs
  - Co-Champion discussions & community input page
- Track A Goals:
  - Define/document:
    - Explicit conditions for and issues with reuse
    - Concepts/meta-ontology
    - Approaches to modularization and best practices
    - Specific design patterns and exemplary content
      - For content reuse in applied ontologies and semantic web/linked data, and for reasoning and big data
  - Expand tooling, such as OOR, to support defining and finding reusable content

Reuse

- Reuse issues not unique to ontologies/schemas
  - Parallels and differences with software and data model reuse
- Capture and understand range of conditions, contexts and intended purposes for which an ontology/linked data is "safely" and productively reused
  - Confirm/track that reused content works "as expected" in new contexts
- Understand dimensions of variability and affects on modularity and reuse
  - Variability across contexts (for ex, concept or property present or absent in different contexts/uses)
  - Variability over time (evolution of a module and need to take current trends and future directions into account)
- Separate reuse of classes/concepts, from properties, from individuals and from axioms
  - Easier to target what is possible to reuse and reduces amount of transformation and cleaning

Conditions for Reuse

- Specific items for consideration include:
    1. Content is accessible and can be found
    2. Content is "understood" from documentation etc.
    3. The re-user is motivated to find the content
    4. The content is in a form conducive to re-use or can be converted/transformed to a usable form
    5. The re-user knows how to do and has the tools to do the conversion/transformation
    6. The content is logically consistent with the micro-theories/ conceptual models of the re-user and this can be established
    7. The re-user trusts the content and its quality, and believes that this quality will be maintained

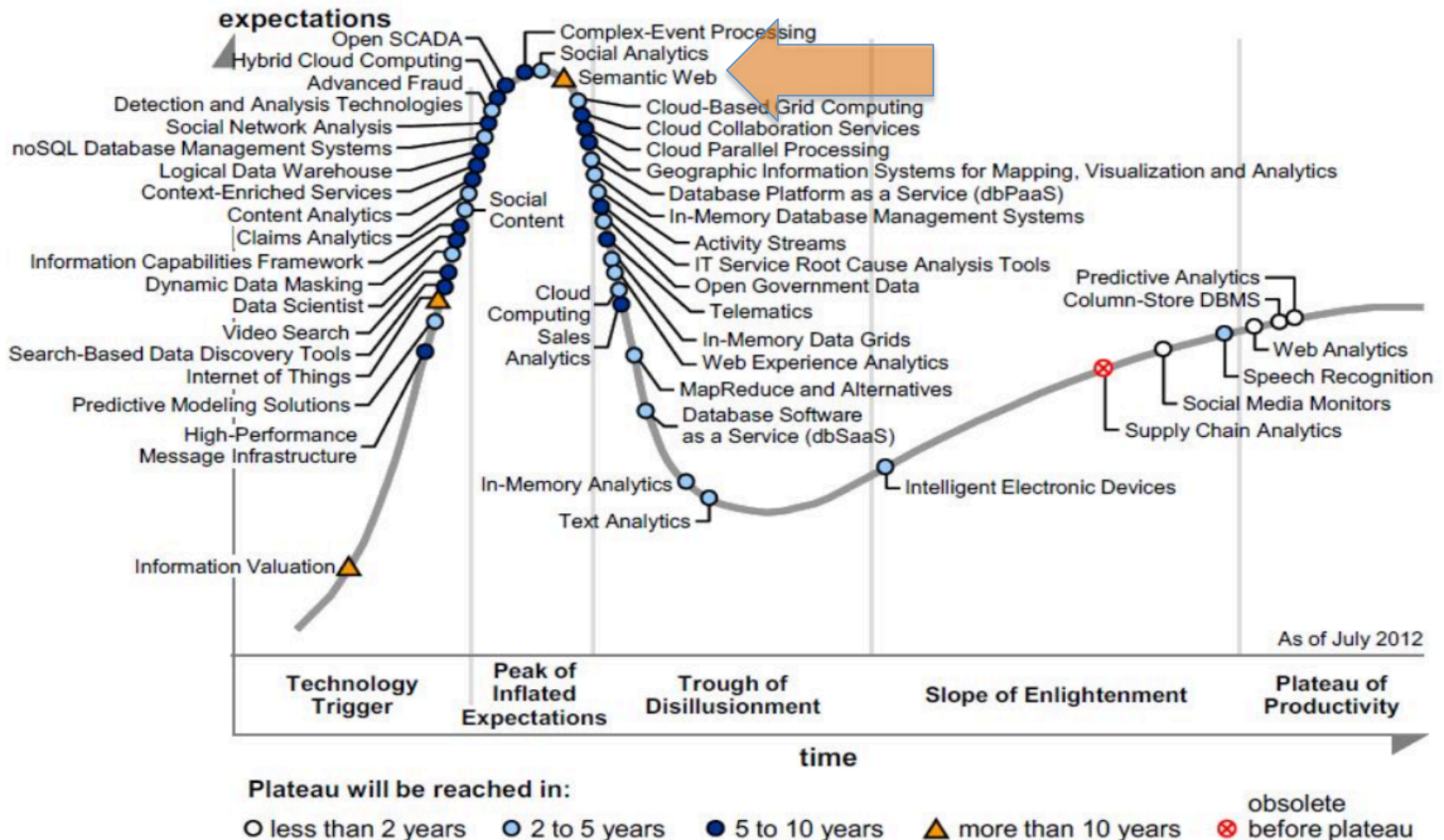# Ontology Design Patterns May be Useful

- "An ontology design pattern is a reusable successful solution to a recurrent modeling problem."
- So-called *content patterns* usually encode specific abstract notions, such as process, event, agent, etc. that are widely applicable and can be customized to needs

From Hitzler talk

5
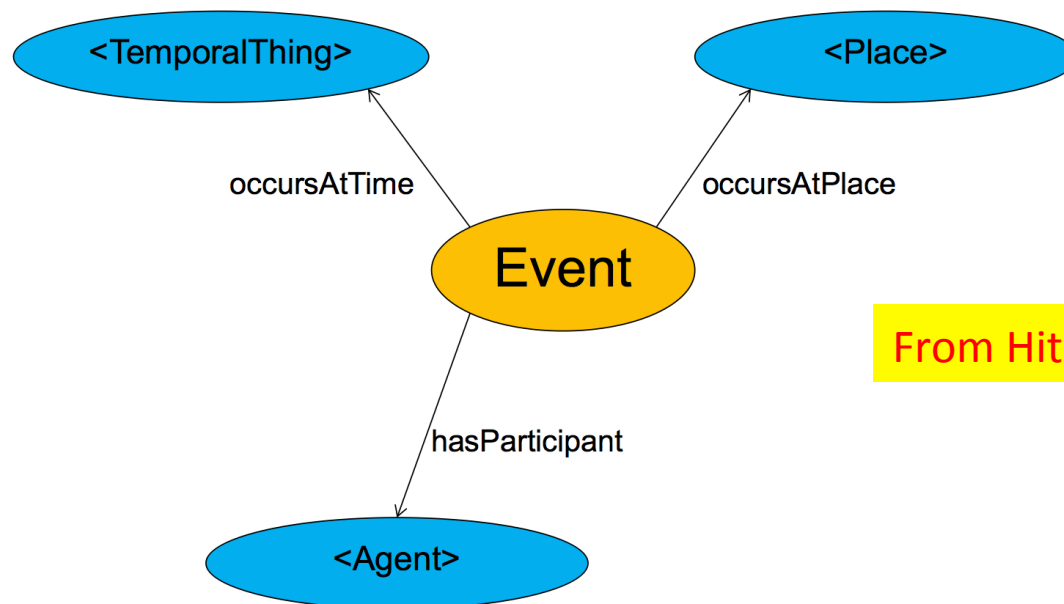
Big Data Landscape

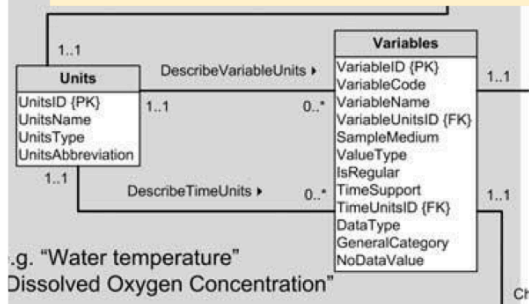Figure 1. Hype Cycle for Big Data, 2012

Big Data Vocabularies Need Semantics

Example fro Big Data Domain- Hydrology- Variables, Tags & "Ontology" Concepts

Observation DM uses RDB structure to
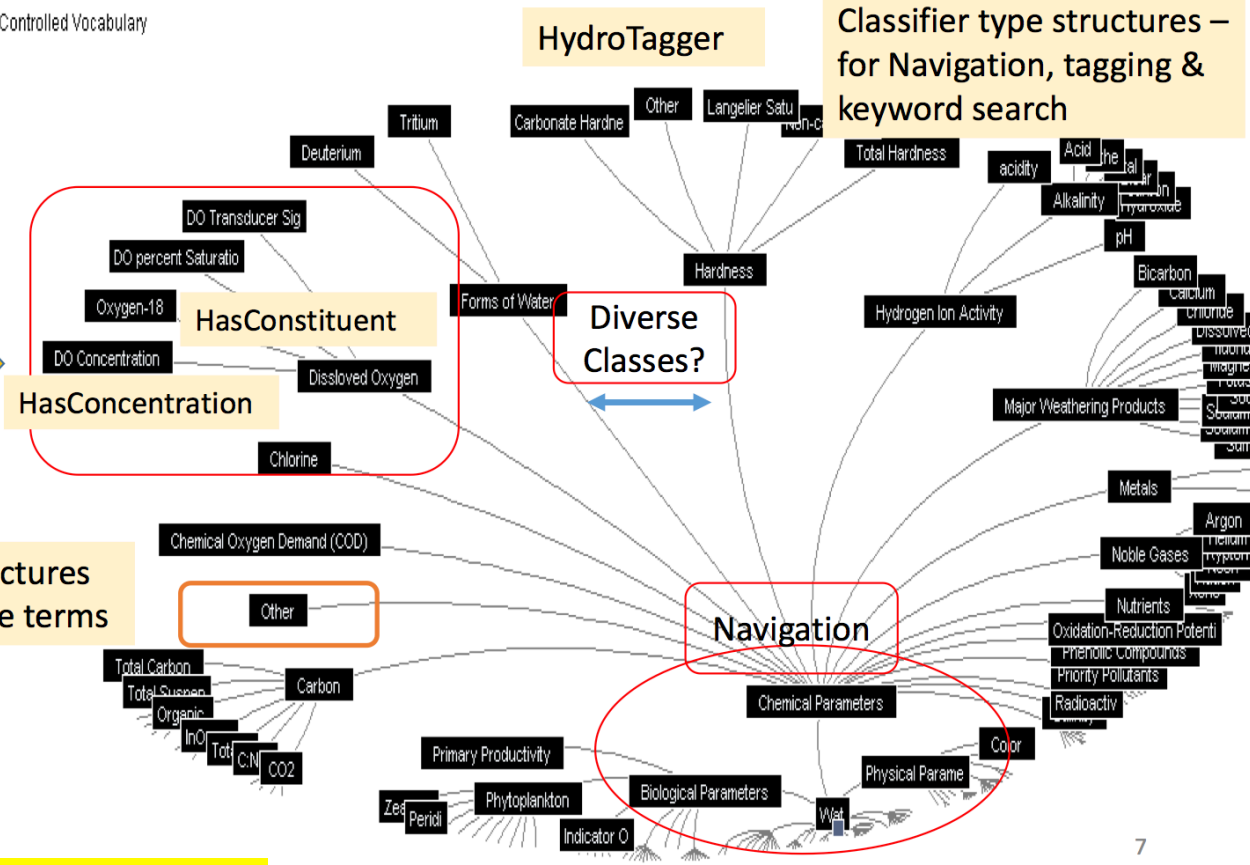Integrate files & handle heterogeneity, Good MD attributes -Limited semantics

CUAHSI Controlled Vocabulary

HydroTagger

Classifier type structures – for Navigation, tagging & keyword search

HasConstituent

HasConcentration

Diverse Classes?

Concept ID	Concept Name	Ontology Layer
41	Chemical	1
42	Organic	2
43	PCBs	3
1001	Homolog Groups	4

Classifier type structures Connect to variable terms

Navigation

1220 Deca_Chloro_PCB 5

Star Tree™ created with Inxight VizServer™

From Berg-Cross talk

Reasoning is a Research Question

Research Questions

From Katsumi & Gruninger talk

All of these observations raise the questions:

- Are Semantic Web ontologies able to support non-trivial reasoning problems?
- If not, why?
- Are the existing ontologies simply not designed with enough semantics to support these applications, or have they reached the limit of what Semantic Web languages can support?

Tooling

- For reuse, ontologies need tooling and governance
  - Search capabilities and governance in repositories (such as the Open Ontology Repository)
    - Including both topical ontologies and linked data schemes
  - Ontologies must include consistent, supporting metadata for query
    - Metadata includes context, use cases, labels, governance information, etc.
    - Possible definition is the Ontology Metadata Vocabulary
  - Reuse enhanced by feedback and user input
    - Support a recommendation system and feedback mechanisms in the repository
  - Governance needs a process and its enforcement
    - Process should include open consideration, comment, revision and acceptance

Best Practices (I)
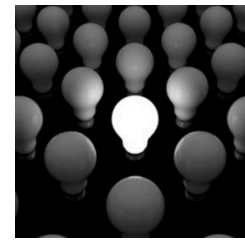
- Small, more modular ontologies and schemes
  - More possibilities for reuse due to greater focus and cohesiveness, and likely less dependency on the original context
- "Integrating" modules defined for an application or domain
  - Employing owl:equivalentClass and OWL axioms to map between the concepts, properties, etc. of the complete set of modular ontologies that address an application/domain
- Each module and its concepts, properties, axioms, ... well-documented via well-established labels and predicates
  - SKOS, etc.
  - A search for "primitives"
- Patterns of concepts separated from patterns of usage, analysis, traversal and diagnosis

Best Practices (II)
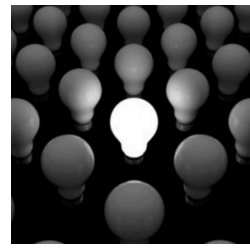
- Modularity viewed from the perspective of the user, not the creator
- Multiple domains represented such that the ontologies and schemes represent "common needs"
  - No single domain focus
- Plans for variability, extension and enhancement documented with the modules or Ontology Patterns
- Constraints or axioms distinguished as:
  - Definitive ("defining" the concepts that are necessary in the core module)
  - Pragmatic (related to the business uses or a particular domain)

BACKUP

Track Questions

1. How can we characterize or measure semantic content reuse, both between ontologies and by Big Data and Semantic Web communities?
2. What building blocks of common semantic content exists now to enable interoperability?
   - What additions are needed to move forward and how are these best achieved?
3. What is involved in reuse of Linked Data versus reuse of ontologies?
4. What is an example of a small set of semantic content that the community might propose for reuse?
   - Is there agreement on these or things like ODPs as building blocks?
5. What is an example of a large set that the community might propose for reuse?
6. Is it reasonable to expect reuse of an entire ontology like DOLCE and Semantic Sensor Network (SSN)?
   - If so under what conditions might this be reasonable?
   - Is it better to expect alignment rather than exact content reuse?
7. Is reuse about semantics alone or should it also address reasoning and data analytics?

Talks during Jan 23 Session

Briefings: (44B0)

- **Dr. GaryBergCross (SOCoP)** - **"Use and Reuse of Semantic Content: The Problems and Efforts to Address Them - An Introduction"** - slides   (44KU)
- **Professor PascalHitzler** (Wright State U) - **"Towards ontology patterns for ocean science repository integration"** - slides   (44B1)
- **Ms. AndreaWesterinen** (Nine Points Solutions) - **"Reuse of Content from ISO 15926 and FIBO"** - slides   (44B2)
- **Ms. MeganKatsumi & Professor MichaelGruninger** (U of Toronto) - **"Reasoning about Events on the Semantic Web"** - slides   (44B3)

Note a 2nd session is planned for March 6th and follow on discussion is expected.