Comparing and Contrasting Semantic Web Stack and XBRL Stack

Charles.Hoffman@me.com

November 26, 2017

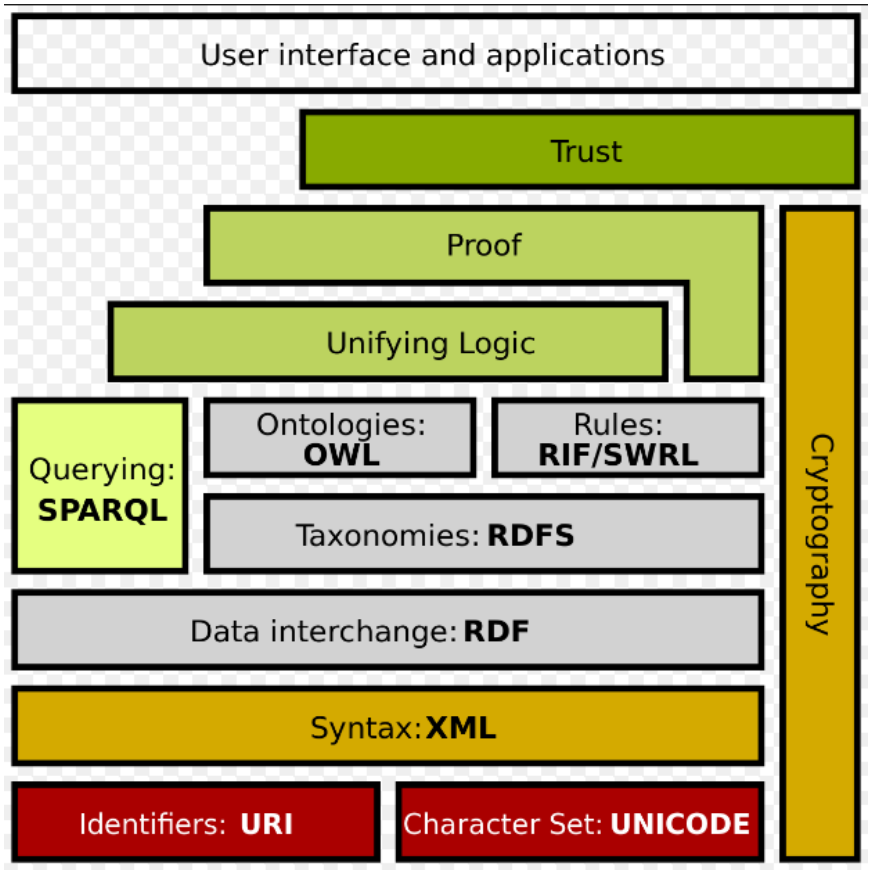
Semantic Web Stack (V1)

W3C reference model of architecture

NOTE: While you CAN create a multidimensional model with this stack of pieces; the multidimensional model is not really a fundamental part of the over-arching model.

See the **RDF Data Cube Vocabulary**, http://www.w3.org/TR/vocab-data-cube/

HOWEVER, the RDF Data Cube Model was specifically build to support SDMX (Statistical Data and Metadata eXchange) which is more an OLAP related model



Wikipedia, *Semantic Web Stack*, https://en.wikipedia.org/wiki/Semantic_Web_Stack

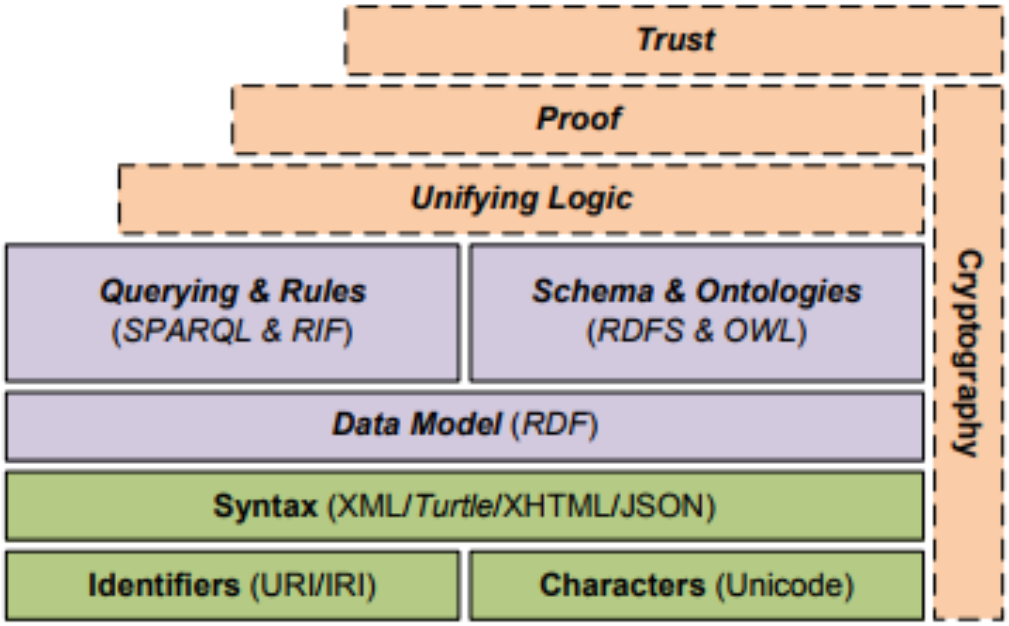
Semantic Web Vision: http://slideplayer.com/slide/2410665/

Understanding the Semantic Web Stack

- URI (Uniform Resource Identifier)
- Unicode
- XML (Extensible Markup Language)
- RDF (Resource Description Framework)
- RDFS (RDF Schema)
- SPARQL (Simple Protocol and RDF Query Language)
- OWL (Web Ontology Language)
- RIF (Rule Interchange Format)
- SWRL (Semantic Web Rule Language Combining OWL and RuleML) (see RIF and RuleML relation)
- SHACL (Shapes Constrain Language) said to be the preferred syntax for rules by some; others say RuleLog is the preferred approach to rules

Semantic Web Stack (V2)

Proposed alternative model

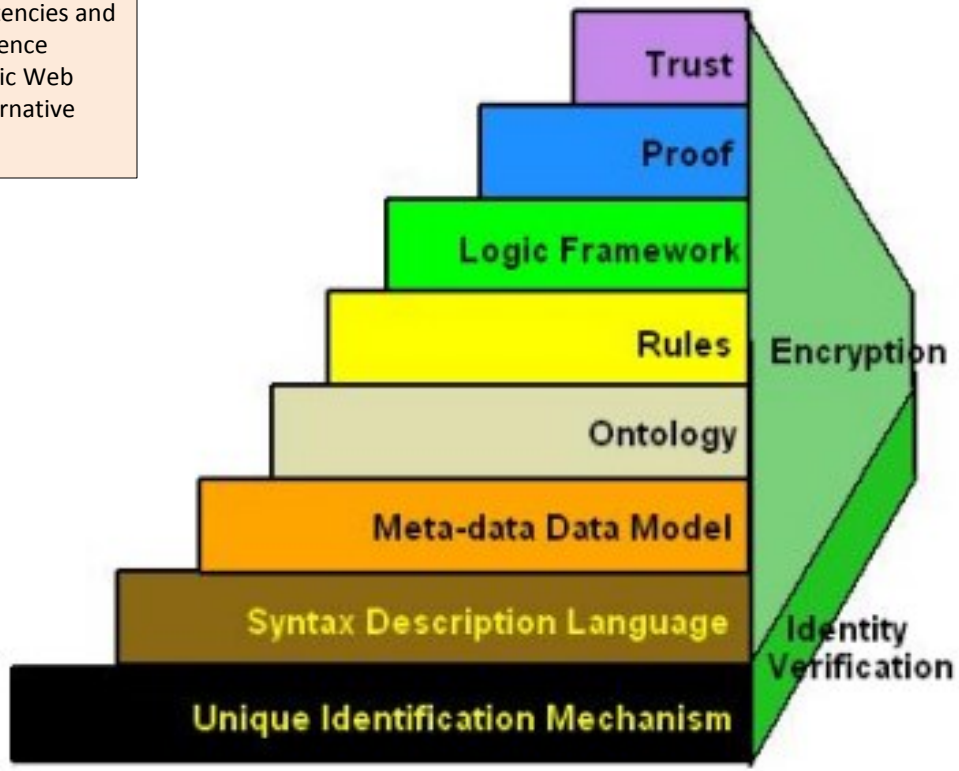


Aidan Hogan, *Chapter 1: Linked Data & the Semantic Web Standards*, http://aidanhogan.com/docs/ldmgmt_semantic_web_linked_data.pdf (provides an excellent explanation of the different layers)

Semantic Web Stack (CFL Model)

Comprehensive, Functional, Layered (CFL) architecture for the Semantic Web

Authors point out inconsistencies and discrepancies in W3C reference architecture of the Semantic Web Stack and propose this alternative model



Aurona Gerber, Alta van der Merwe, and Andries Barnard, *A Functional Semantic Web Architecture*, https://link.springer.com/content/pdf/10.1007/978-3-540-68234-9_22.pdf

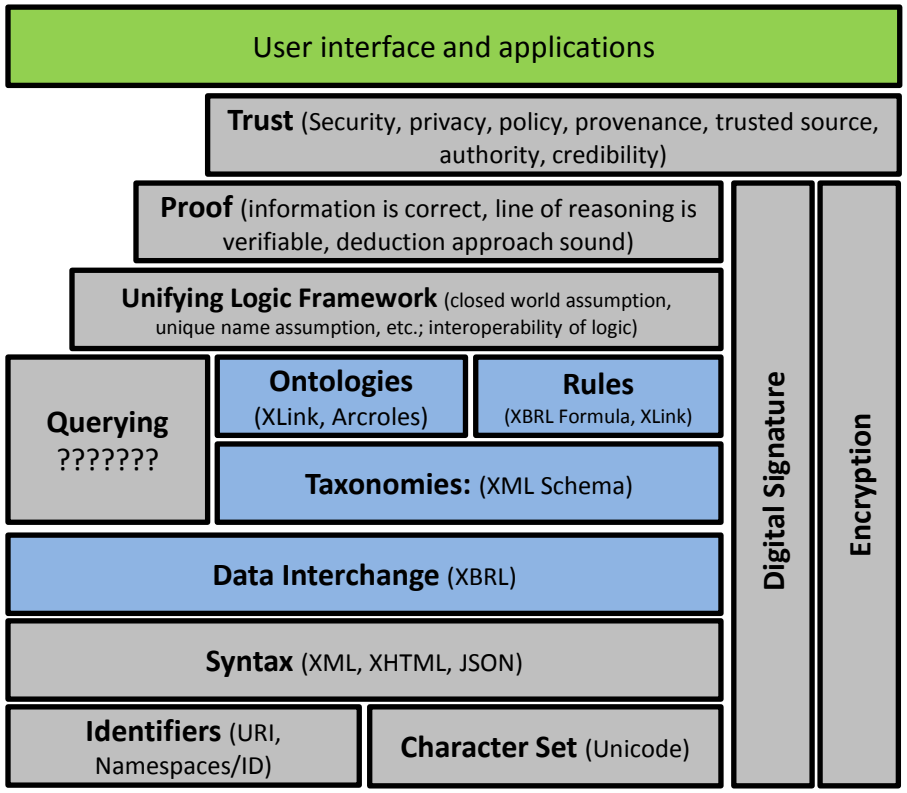
XBRL Business Reporting Stack (V1)

Inspired by W3C reference model and alternative model

NOTE: I want to explicitly put the multidimensional model in here but I am not exactly sure where to put it.

Further, the multidimensional model used by XBRL Dimensions is a true multidimensional model where OLAP can be used, but you are not forced to use OLAP.

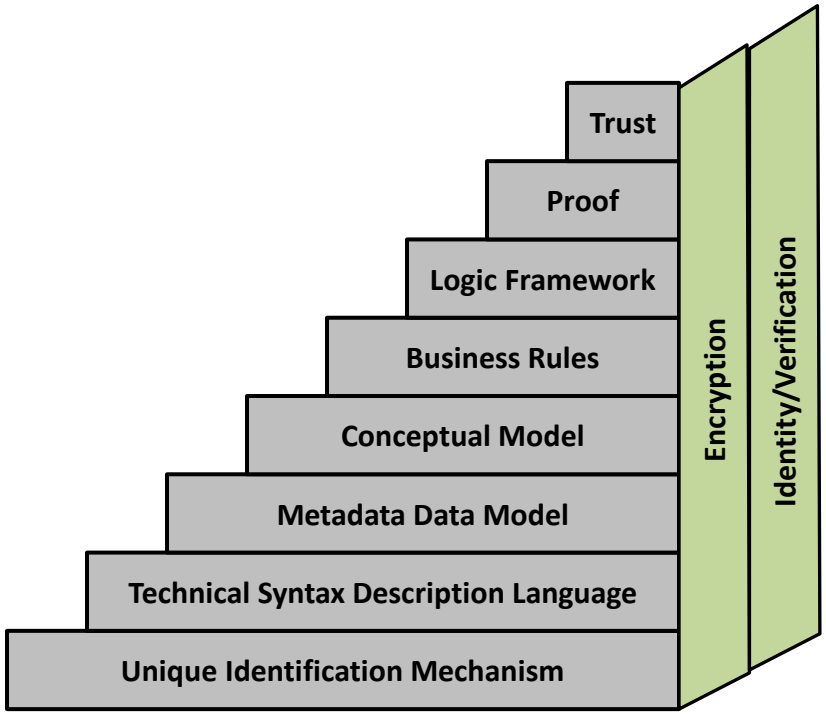
For more information please see, http://xbrl.squarespace.com/journal/2016/3/18/introduction-to-the-multidimensional-model-for-professional.html



This is inspired by the W3C Reference Model of the Semantic Web Stack

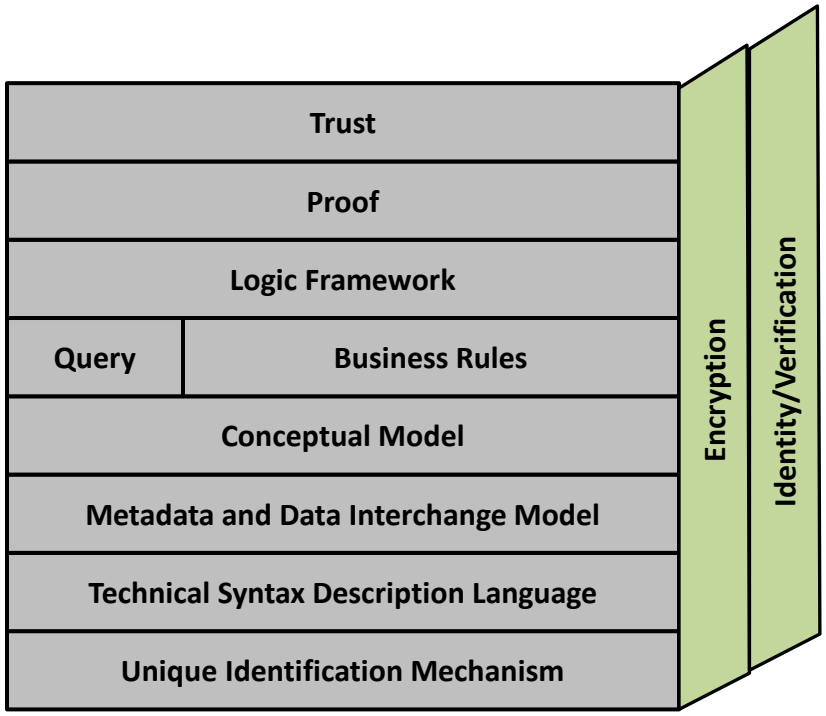
XBRL Business Reporting Stack (V2)

Inspired by alternative functional model



XBRL Business Reporting Stack (V3)

Inspired by alternative functional model and OSI model format



Wikipedia, *OSI Model*, https://en.wikipedia.org/wiki/OSI_model

XBRL Business Reporting Stack (V4)

Further tuning of V3 model of XBRL Stack

#	Layer		Technical Syntax	Function
10	**Application**			Business application that makes use of information.
9	**Trust**			Security, privacy, policy, provenance, trusted source, authority and credibility of the source of information.
8	**Proof**			Form of proof that Information is correct; line of reasoning used is justifiable, deduction approach is sound, etc.
7	**Logic framework**			Defines the specific framework of rules of logic that are used by the business rules. For example, closed world assumption; unique name assumption. Uses the business rules technical syntax. Enables interoperability of business information between systems.
6	**Query**	**Business Rules**	XLink, Arcroles, XBRL Formula	Rules that describe logical, mathematical, mechanical, structural, and other relationships which assure that facts are consistent with the conceptual model.
5	**Conceptual Model**		XLink, Arcroles, XBRL taxonomy schemas	High-level model that describes things and relations between things to which all metadata and data must conform. Also known as an entity-relation model or a theory.
4	**Metadata and Data Interchange Model**		XBRL instance, XBRL taxonomy schemas which includes XML Schema and XLink	Physical document, stream, or set of technical syntax. Facts (data) and information about those facts (metadata) exchanged.
3	**Technical Syntax Description Language**		XML, XHTML, JSON	Technical syntax that is used to represent the things described by the unique identifiers.
2	**Unique Identification Mechanism**		URI, XML Namespaces/ID	Globally agreed-upon unique identifiers that describe things so that each thing can be identified.
1	**Character Set**		Unicode	Encode individual characters using a global standard character set.

Summary

- **Multi-stack architectures** are a reality of life; things change and people have different preferences
- Different stacks need to **interoperate** with one another, particularly in our internetworked world of today
- Interoperability can only occur at the **lowest level of logic** supported by a logic framework
- The “Logic Framework” layer of the Semantic Web Stack is a good target and would allow interoperability between the Semantic Web Stack and the XBRL Stack
- What would even be better is a clearly defined **Unifying Business Logic Framework**, a global standard logic framework for business represented using a controlled natural language with which is approachable by business professionals, be guaranteed safe and reliable, and be as powerful as possible

Unifying Logic Framework for Business

- a global standard logic framework for business
- represented using a controlled natural language format
- represent logic at the highest level possible such that the logic is understandable by business professionals
- rules created are approachable by business professionals
- built in but optional multidimensional model that does not force the use of OLAP, but usable with OLAP or OLTP
- enables interoperability between technology stacks
- enables interoperability between XBRL, GLEI, FIBO (Financial Industry Business Ontology), FRO (Financial Regulation Ontology), the US GAAP XBRL Taxonomy, the IFRS XBRL Taxonomy, etc.
- built based on the logic framework of the Semantic Web Stack which has been evolving for 25 or so years

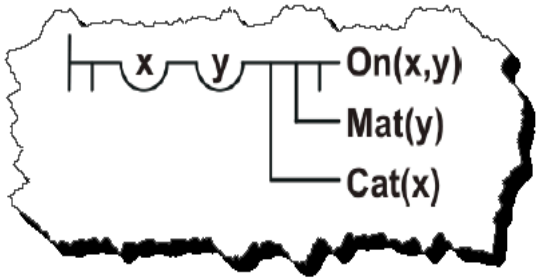
Defining a Unifying Logic Framework for Business

- Current alphabet soup of standard logic frameworks for rules:
  - ISO/IEC Common Logic (CL)?
  - OMG Semantics of Business Vocabulary and Business Rules (SBVR)?
  - W3C RDFS + OWL + RIF/SWRL syntax logic? (SWRL is not a recommendation, only a submission, RIF and SWRL seem to have issues)
  - W3C RDFS + OWL + SHACL syntax logic which specifies closed world assumption and unique names assumption?
  - Industry Initiative RuleLog which is designed to be appropriately expressive for supporting knowledge representation in complex domains and yet to be efficiently implementable?
  - Industry Initiative RuleML which allows for partially constrained logic profiles and fully-specified logic semantics?
  - XBRL Formula? (which has known deficiencies)
- Who chooses? Business professionals loose if there is no single global standard

Additional Details

Controlled Natural Language Format

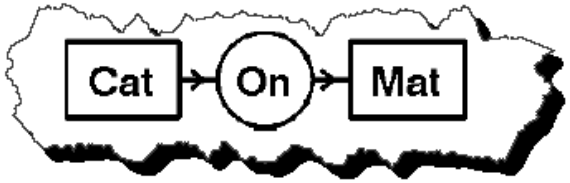
How do you say “A cat is on a mat.”



$$\Sigma_x \Sigma_y \mathrm{Cat}_x \bullet \mathrm{Mat}_y \bullet \mathrm{On}_{x,y}$$

$$\exists x \exists y \ \mathrm{Cat}(x) \wedge \mathrm{Mat}(y) \wedge \mathrm{On}(x,y)$$

```
(exists ((x Cat) (y Mat)) (On x y))
```




```
[Cat *x] [Mat *y] (On ?x ?y)
```

```
SELECT FIRST.ID, SECOND.ID
FROM   OBJECTS FIRST, OBJECTS SECOND, SUPPORTS
WHERE  FIRST.TYPE = "Cat"
AND    SECOND.TYPE = "Mat"
AND    SUPPORTS.SUPPORTER = SECOND.ID
AND    SUPPORTS.SUPPORTEE = FIRST.ID
```

```
A cat is on a mat.
```

Trust and Proof

- Proof
  - Proof that the answer found is correct
  - Transparency into the logic and line of reasoning used to derive information
- Trust
  - Transparency into the origin or provenance of the information
  - Information chain explaining source of information, authority and credibility of information source

N. Henze, *Proof and Trust Layers of the Semantic Web*, http://www.kbs.uni-hannover.de/Lehre/semweb07/10_proof_trust.pdf

Unifying Logic Framework

- **Agreement as to logic** being used, for example
  - Definition of meaning of “true” and “false”
  - Closed world assumption (or open world assumption?)
  - Unique names assumption
  - Default negation
  - Other expressivity details from the SHACL use cases and requirements, http://www.w3.org/TR/shacl-ucr/
- An **interoperability layer** that provides the foundation for combining these lower-level and having multi-stack architectures that can work together
- Logic frameworks can ONLY INTEROPERATE at the lowest level of problem solving logic of a logic framework
- Semantic Web Stack offers possibility of sophisticated but very safe reasoning with a combination of ontologies and non-monotonic rules, far beyond what is possible with todays XBRL Formula processors

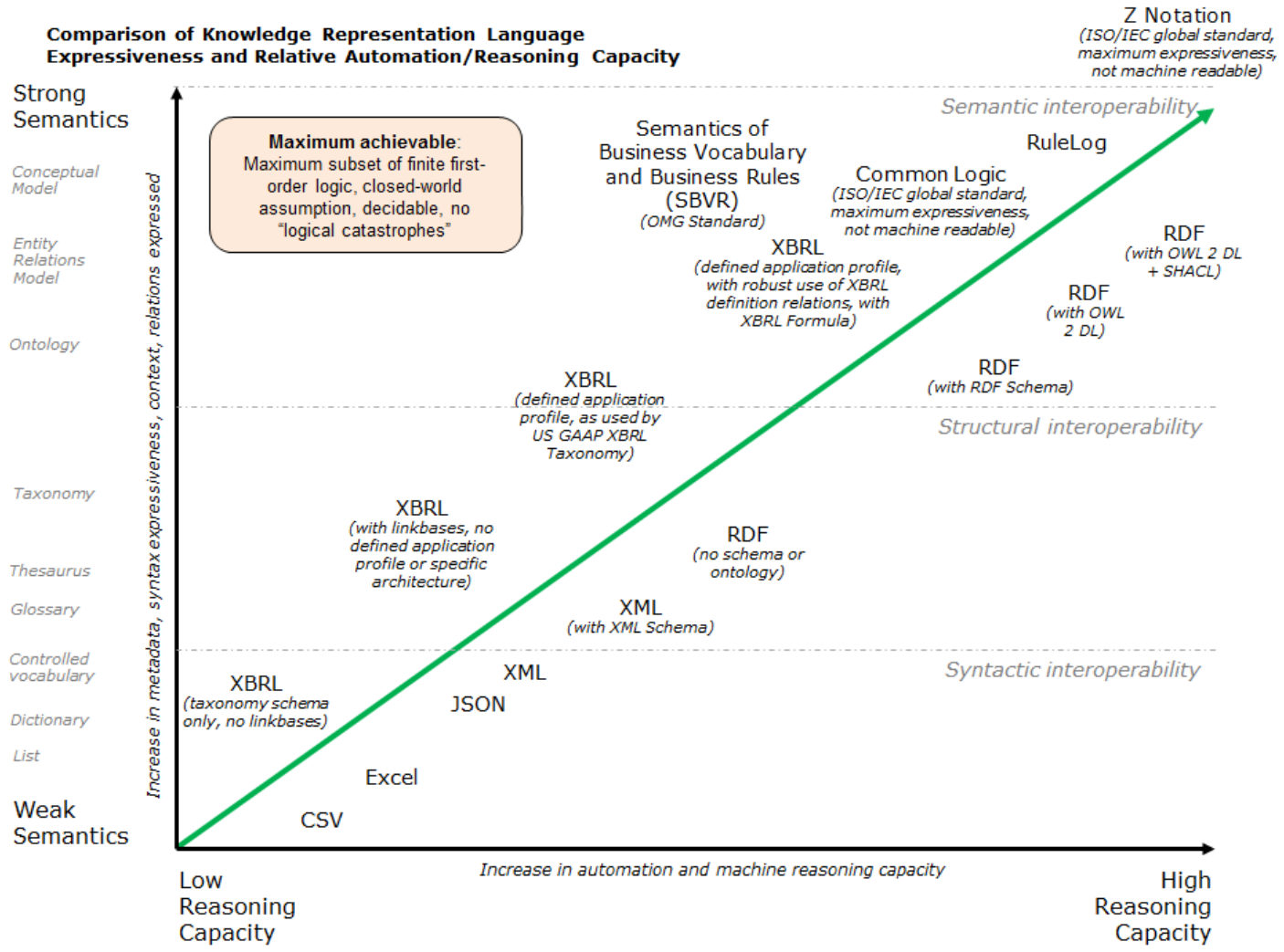
Michael Kifer, Jos de Bruijn, Harold Boley, and Dieter Fensel, *A Realistic Architecture for the Semantic Web*, http://www3.cs.stonybrook.edu/~kifer/TechReports/msa-ruleml05.pdf

Matthias Knorr, *Towards a Semantic Web Unifying Logic*, http://userweb.fct.unl.pt/~mkn/resources/ICCL2013.pdf

Identifiers

- **URI** (Uniform Resource Identifier; older approach which replaces URLs)
- **IRI** (International Resource Identifier; newer approach to URIs)
- **Namespaces** (Namespaces in XML; XML namespaces provide a simple method for qualifying element and attribute names used in XML documents by associating them with namespaces identified by URI references; essentially URIs in XML)
- **US GAAP** and **IFRS** (Financial reporting concepts that use XML Namespaces as identifiers)
- **GLEI** (Global Legal Entity Identifier; one important type of identifier)

Relative Expressive Power



Inspired by similar comparisons from *An Intrepid Guide to Ontologies* http://www.mkbergman.com/date/2007/05/16/ and *Semantics Overview* http://prezi.com/prwsxj8po3ln/semantics-overview/

OSI Layers Model Format Example

OSI Model			
Layer		Protocol data unit (PDU)	Function[3]
Host layers	7. Application	Data	High-level APIs, including resource sharing, remote file access
	6. Presentation		Translation of data between a networking service and an application; including character encoding, data compression and encryption/decryption
	5. Session		Managing communication sessions, i.e. continuous exchange of information in the form of multiple back-and-forth transmissions between two nodes
	4. Transport	Segment (TCP) / Datagram (UDP)	Reliable transmission of data segments between points on a network, including segmentation, acknowledgement and multiplexing
Media layers	3. Network	Packet	Structuring and managing a multi-node network, including addressing, routing and traffic control
	2. Data link	Frame	Reliable transmission of data frames between two nodes connected by a physical layer
	1. Physical	Bit	Transmission and reception of raw bit streams over a physical medium

Wikipedia, *OSI Model*, https://en.wikipedia.org/wiki/OSI_model

Comparing Semantic Web and XBRL Business Reporting Stacks