

## Threes All of the Way Down to Typologies

by Mike Bergman - Monday, October 03, 2016

<http://www.mkbergman.com/1985/threes-all-of-the-way-down-to-typologies/>



### Further Details on the Design of KBpedia's Knowledge Ontology (KKO)

Every knowledge structure used for knowledge representation ([KR](#)) or knowledge-based artificial intelligence ([KBAI](#)) needs to be governed by some form of conceptual schema. In the semantic Web space, such schema are known as "ontologies", since they attempt to capture the nature or *being* (Greek ??????, or ontós) of the knowledge domain at hand. Because the word 'ontology' is a bit intimidating, a better variant has proven to be the *knowledge graph* (because all semantic ontologies take the structural form of a graph). In [Cognonto](#)'s KBAI efforts, we tend to use the terms *ontology* and *knowledge graph* interchangeably.

Cognonto uses the KBpedia Knowledge Ontology (KKO) as its upper conceptual schema. KKO is the structure by which Cognonto's hundreds of thousands of reference concepts and millions of entities are organized [\[1\]](#). This article presents an overview and rationale for this KKO structure. Subsequent articles will delve into specific aspects of KKO where warranted.

### A Grounding in Peirce's Triadic Logic

The upper structure of the KBpedia Knowledge Ontology (KKO) is informed by the triadic logic and basic categories of [Charles Sanders Peirce](#). If the relation of this triadic design to knowledge representation appears a bit opaque, please refer to the introduction of this series in the prior article, [The Irreducible Truth of Threes](#). The key point is that 'threes' are the fewest by which to model context and perspective, essential to capture the nature of knowledge.

Peirce's triadic logic, or trichotomy, is also the basis for his views on [semiosis](#) (or the nature of signs). The three constituents of Peirce's trichotomy, what he called simply the *Three Categories*, were in his view the most primitive or reduced manner by which to understand and categorize things, concepts and ideas. Peirce's *Three Categories* can be roughly summarized as:

- **Firstness** [1ns] — these are *potentials*, the basic forces or qualities that combine together or interact in various ways to enable the real things we perceive in the world, such as matter, life and ideas. These are the unrealized building blocks, or primitives, the essences or attributes or possible juxtapositions
- **Secondness** [2ns] — these are the *particular realized things or concepts* in the world, what we can perceive, point to and describe. A particular is also known as an entity, instance or individual
- **Thirdness** [3ns] — these are the *laws, habits, regularities and continuities* that may be generalized from particulars. All generals — what are also known as classes, kinds or types — belong to this category. The process of finding and deriving these generalities also leads to new insights or emergent properties, which continue to fuel knowledge discovery. Insights arising from Thirdness enable us to further explore and understand things, and is a driving force for further categorization.

Understanding, inquiry and knowledge require this irreducible structure; connections, meaning and communication depend on all three components, standing in relation to one another and subject to interpretation by multiple agents. (Traditional classification schemes have a dyadic or dichotomous nature, which does not support the richer views of context and interpretation inherent in the Peircean view.)

Peirce argues persuasively that how we perceive and communicate things requires this irreducible triadic structure. The symbolic nature of Thirdness means that communication and understanding is a continuous process of refinement, getting us closer to the truth, but never fully achieving it. Thirdness is a social and imprecise mode of communication and discovery, conducted by us and other agents separate from the things and phenomena being observed. Though it is a fallibilistic process, it is one that also lends itself to rigor and methods. The scientific method is a premier example of Thirdness in action.

What constitutes the *potentials*, realized *particulars*, and *generalizations* that may be drawn from a query or investigation is contextual in nature. That is why the mindset of Peirce's triadic logic is a powerful guide to how to think about and organize the things and ideas in our world (that is, *knowledge representation*). Peirce's triadic logic and views on categorization are fractal in nature. We can apply this triadic logic to any level of information granularity.

Thus, KKO applies this mindset to organizing its knowledge graph. At each level in the KKO upper structure, we strive to organize each category according to the ideas of Firstness (1ns), Secondness (2ns) and Thirdness (3ns), as shown in the upper KKO structure below.

### Basic Structural Considerations

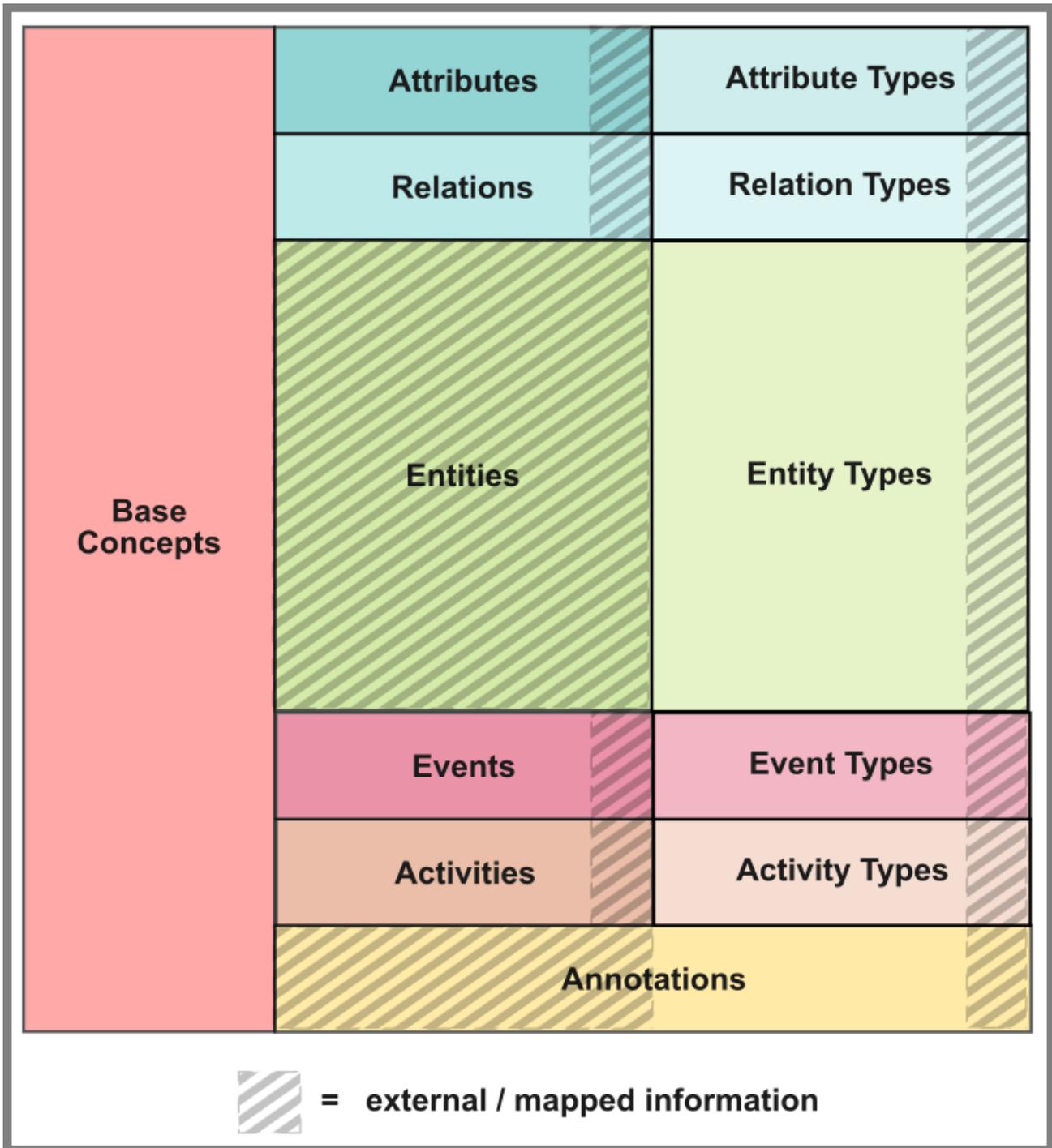
Now armed with a basic conceptual and logical grounding, what are the main kinds of distinctions we want to capture in our knowledge structure? Since our purpose is to provide a means for integrating knowledge bases (KBs) of use to artificial intelligence (AI), or [KBAI](#), the answer to this question resides in: 1) what conceptual distinctions are captured by the constituent KBs; and 2) what kinds of work (AI) we want to do with the structure.

The answers to these questions help us to define the basic vocabulary of our knowledge base, what Peirce called its *speculative grammar* [\[2\]](#). This base vocabulary of KKO is thus:

- **Attributes** are the ways to characterize the entities or things within the knowledge base; while the attribute values and options may be quite complex, the relationship is monadic to the subject at hand. These are intensional properties of the subject
- **Relations** are the way we describe connections between two or more things; relations are external-facing, between the subject and another entity or concept; relations set the extensional structure of the knowledge graph
- **Entities** are the basic, real things in our domain of interest; they are nameable things or ideas that have identity, are defined in some manner, can be referenced, and should be related to types; entities are the bulk of the overall knowledge base
- **Events** are nameable sequences of time, are described in some manner, can be referenced, and may be related to other time sequences or types
- **Activities** are sustained actions over durations of time; activities may be organized into natural classes
- **Types** are the hierarchical classification of natural kinds within all of the terms above
- The **Typology** structure is not only a natural organization of natural classes, but it enables flexible interaction points with inferencing across its ‘accordion-like’ design
- **Base concepts** are the vocabulary to the grammar and top-level concepts in the knowledge graph, organized according to Peircean-informed categories
- **Annotations** are indexes and the metadata of the KB; these can not be inferenced over. But, they can be searched and language features can be processed in other ways.

We will be talking about specific vocabulary items above in subsequent articles. One important distinction to draw for now, however, is the split between *attributes* and *relations*. In standard RDF and OWL ontologies these are lumped together as properties. In OWL, there is the further distinction of datatype and object properties, but these do not quite capture the difference we desire. In KBpedia, attributes are the descriptions or characteristics of a given entity (or its type); relations are the roles, connections or subsumptions *between* objects [3].

How these vocabulary terms relate to one another and the overall KBpedia knowledge structure is shown by this diagram:



### A Knowledge Base Grammar

Note that the three columns of this figure correspond to the three categories of *potentials* (1ns, left column), *particulars* (2ns, middle column) and *generals* (3ns, right column) described above. In terms of KBpedia, all of the instances of the knowledge structure, now numbering over 20 million in the standard version, are affiliated with the middle column (particulars). The classification aspects of KBpedia reside in the right column (generals). The reasoning aspects of KBpedia largely reside in the left and right columns (though reasoning work is also done on selecting and aggregating instances in the middle column) [\[4\]](#).

### Below the Upper Structure are Typologies

More than 85% of the classification structure of KBpedia resides in the generals, or types, in the rightmost column. These, in turn, are organized according to a set of *typologies*, or natural classification structures. Unlike the KKO upper structure, each typology is not necessarily organized according to Peirce's triadic logic. That is because once we come to organize and classify the real things in the world, we are dealing with objects of a more-or-less uniform character (such as *animals* or *products* or *atomic elements*).

There are about 80 such typologies in the KBpedia structure, about 30 of which are deemed "core", meaning they capture the bulk of the classificatory system. Another document presents these [30 "core" typologies](#) in more detail.

I have written elsewhere [\[5\]](#) about the basis for "natural" classification systems. These approaches, too, are drawn from Peirce's writings. Natural classifications may apply to truly "natural" things, like organisms and matter, but also to man-made objects and social movements and ideas. The key argument is that shared attributes, including a defining kind of "essence" (Aristotle) or "final cause" (Peirce) help define the specific class or type to which an object may belong. For Peirce, what science has to tell us, or what social consensus settles upon, holds sway.

If accomplished well, natural classification systems lend themselves to hierarchical structures that may be reasoned over. Further, if the splits between typologies are also done well, then it is also possible to establish non-overlapping ("disjoint") relationships between typologies that provide powerful restriction and selection capabilities across the knowledge structure. We believe KBpedia already achieves these objectives, though we continue to refine the structure based on our mappings to other external systems and other logical tests.

### The KKO Upper Structure

We now have the pieces in hand to construct the full KKO upper structure. Here is the upper structure of KKO with its 144 concepts:

**Predica** [1ns]

**Ontics** [1ns]

**Qualities** [1ns]

**Physical** [2ns]

**Being** [1ns]

**One** == Haeccity [1ns]

**True** [2ns]

**Good** [3ns]

**Form** [2ns]

**Structure** [3ns]

**Conceptual** [3ns]

**Absolute** [1ns]

**SimpleRelative** [2ns]

**Conjugative** [3ns]

**Attribuo** [2ns]

**Identity** [1ns]

**Haeccity** == One [1ns]

**Nature** [2ns]

**Beingness** [1ns]

**Real** [2ns]

**Matter** [1ns]

**SubstantialForm** [2ns]

**AccidentalForm** [3ns]

**Fictional** [3ns]

**Quiddity** [3ns]

**Intensional** [2ns]

**Conjunctive** [3ns]

**Quantity** [1ns]

**Values** [1ns]

**Numbers** [1ns]

**Multitudes** [2ns]

**Magnitudes** [3ns]

**Discrete** [2ns]

**Continuous** [3ns]

**Roles** [2ns]

**Typical** [3ns]

**Relatio** [3ns]

**Subsumption** [1ns]

**Similar** [2ns]

**LogicalConnection** [3ns]

**Unary** [1ns]

**Binary** [2ns]

**Conditional** [3ns]

**Particulars** [2ns]

**Entities** [1ns]

**SingleEntities** [1ns]

**Objects** [1ns]

**States** [2ns]

**Events** [3ns]

**PartOfEntities** [2ns]

**Members** [1ns]

**Parts** [2ns]

**FunctionalComponents** [3ns]

**ComplexEntities** [3ns]

**CollectiveStuff** [1ns]

**MixedStuff** [2ns]

**CompoundEntities** [3ns]

**Indices** [2ns]

**Indicators** [1ns]

**Associations** [2ns]

**Annotations** [3ns]

**Selectional** [1ns]

**Referential** [2ns]

**Directional** [3ns]

**Continua** [3ns]

**Space** [1ns]

**Points** [1ns]

**Areas** [2ns]

**2D Dimensions**

**SpaceRegions** [3ns]

**3D Dimensions**

**Time** [2ns]

**Instants** [1ns]

**Intervals** [2ns]

**Events** [3ns]

**Duratives** [3ns]

**Situations** [1ns]

**Activities** [2ns]

**Processes** [3ns]

**Generals** [3ns]

(==

SuperTypes)

**SignElem**

**ents** [1ns]

**AttributeTypes**

[1ns]

**RelationTypes**

[2ns]

**SituationTypes**

**Symbols** [3ns]

**Primitives** [1ns]

**Structures** [2ns]

**Conventions** [3ns]

**Constitue**

**nts** [2ns]

**NaturalPhenomena** [1ns]

**SpaceTypes** [2ns]

**Shapes** [1ns]

**Places** [2ns]

**LocationPlace**

**AreaRegion**

**Forms** [3ns]

**TimeTypes** [3ns]

**Times** [1ns]

**EventTypes** [2ns]

**ActivityTypes** [3ns]

**Manifestations** [3ns]

**NaturalMatter** [1ns]

**AtomsElements** [1ns]

**NaturalSubstances** [2ns]

**Chemistry** [3ns]

**OrganicMatter**

[2ns]

**OrganicChemistry** [1ns]

**BiologicalProcesses**

**LivingThings** [2ns]

**Prokaryotes** [1ns]

**Eukaryotes** [2ns]

**ProtistsFungus** [1ns]

**Plants** [2ns]

**Animals** [3ns]

**Diseases** [3ns]

**Agents** [3ns]

**Persons** [1ns]

**Organizations** [2ns]

**Geopolitical** [3ns]

**Symbolic** [3ns]

**Information** [1ns]

**AVInfo** [1ns]

**VisualInfo**

**AudioInfo**

**WrittenInfo** [2ns]

**StructuredInfo** [3ns]

**Artifacts** [2ns]

**FoodDrink**

**Drugs**

**Products**

**Facilities**

**Systems** [3ns]

**MentalProcesses** [1ns]

**Concepts** [1ns]

**TopicsCategories** [2ns]

**LearningProcesses** [3ns]

**SocialProcesses** [2ns]

**FinanceEconomy**

**Society**

**Methodetic** [3ns]

**InquiryMethods** [1ns]

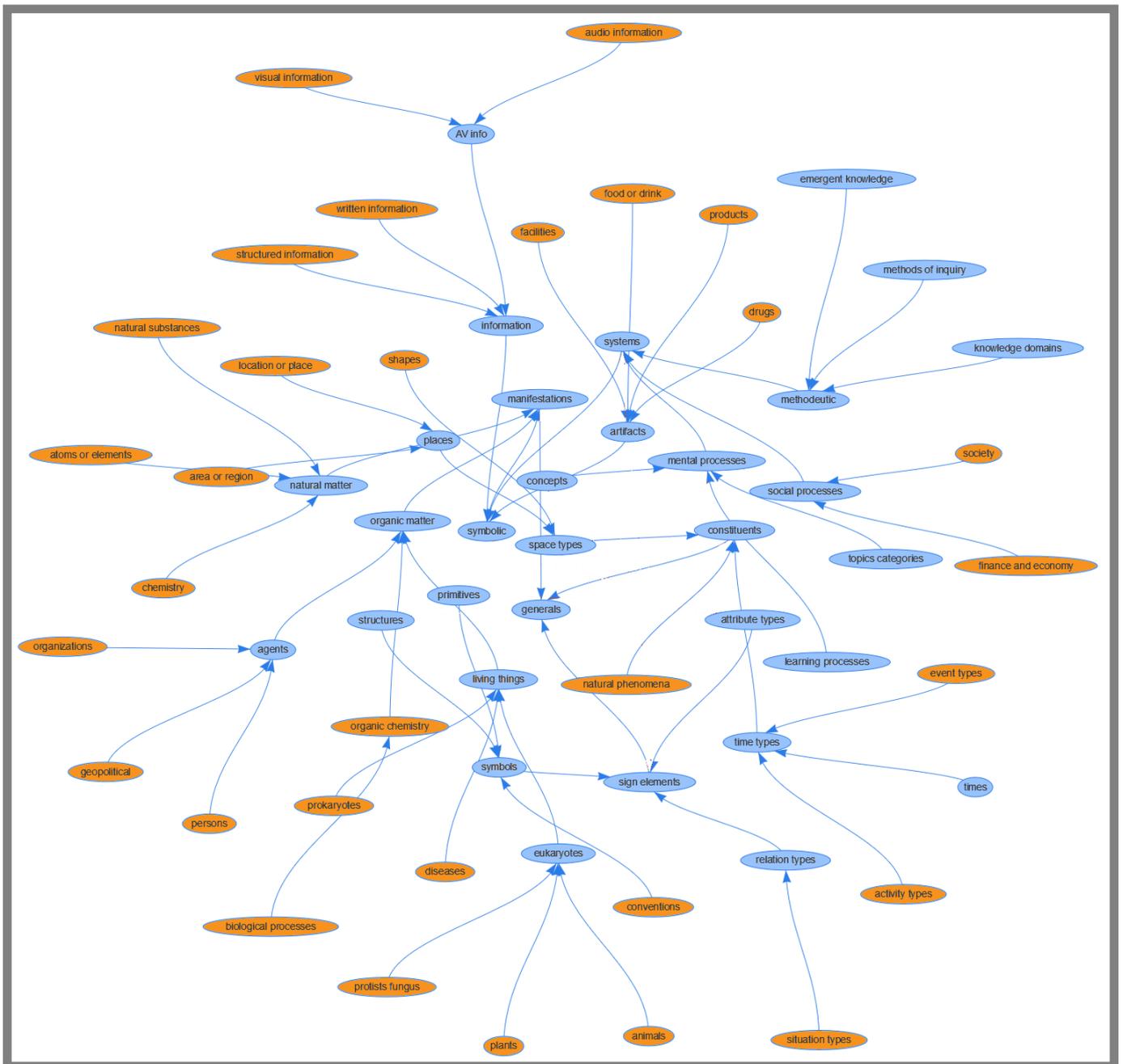
**KnowledgeDomains** [2ns]

**EmergentKnowledge** [3ns]

Where appropriate, each entry is also labeled with one of the *Three Categories* (1ns, 2ns, 3ns). Note that all of the typologies are shown under the main Generals (3ns) category. The main "core" typologies are shown in orange. (Note that **TopicsCategories** and **KnowledgeDomains** are big typologies, but are not disjoint in any way. **Shapes** is also a big typology, but about half of all entities have that type.)

## A Graph View of the Structure

We can differently view this structure through a graph view of the KKO structure, also showing the "core" typologies in orange. As might be expected, the KKO "core" ontologies tend to occur at the periphery of this graph:



## Upper Knowledge Graph with 'Core' Typologies Highlighted

### For Further Information

Of course, the purpose of all of this design is to provide a coherent, consistent, logical structure over which Cognonto may reason and link external data and schema. We will be discussing elsewhere the specific use cases and applications of this structure. For now, we wanted to set the stage for the design basis for KKO. This article will be a common reference to those subsequent discussions.

The KKO upper structure may be [downloaded](#) and inspected in greater detail.

---

[1] For an introduction to Cognonto and KBpedia, its knowledge structure, see M.K. Bergman, 2016. "[Cognonto is on the Hunt for Big AI Game](#)", *AI3:::Adaptive Information* blog, September 20, 2016.

[2] See M.K. Bergman, 2016. "[A Speculative Grammar for Knowledge Bases](#)", *AI3:::Adaptive Information* blog, June 20, 2016.

[3] Attributes, Relations and Annotations comprise OWL properties. In general, Attributes correspond to the OWL datatypes property; Relations to the OWL object property; and Annotations to the OWL annotation property. These specific OWL terms are not used in our speculative grammar, however, because some attributes may be drawn from controlled vocabularies, such as colors or shapes, that can be represented as one of a list of attribute choices. In these cases, such attributes are defined as object properties. Nonetheless, the mappings of our speculative grammar to existing OWL properties is quite close.

[4] As I [earlier wrote](#), "Description logics and their semantics traditionally split concepts and their relationships from the different treatment of instances and their attributes and roles, expressed as fact assertions. The concept split is known as the TBox (for *terminological* knowledge, the basis for T in TBox) and represents the schema or taxonomy of the domain at hand. The TBox is the structural and intensional component of conceptual relationships. The second split of instances is known as the ABox (for assertions, the basis for A in ABox) and describes the attributes of instances (and individuals), the roles between instances, and other assertions about instances regarding their class membership with the TBox concepts." In the diagram above, the middle column represents particulars, or the ABox components. The definition of items and the first and third columns represent TBox components.

[5] See M.K. Bergman, 2015. "[Natural Classes' in the Knowledge Web](#)", *AI3:::Adaptive Information* blog, July 13, 2015.