

Cognonto Snapshot

Artificial intelligence (AI) and machine learning are revolutionizing knowledge systems. Improved algorithms and faster graphics chips have been contributors. But the most important factor in knowledge-based AI's renaissance has been the availability of massive digital datasets for the training of machine learners.

Wikipedia and data from search engines are central to recent breakthroughs. Wikipedia is at the heart of Siri, Cortana, the former Freebase, DBpedia, Google's Knowledge Graph and IBM's Watson, to name just a prominent few AI question answering systems. Natural language understanding is showing impressive gains across a range of applications. To date, all of these examples have been the result of bespoke efforts. It is very expensive for standard enterprises to leverage these knowledge resources on their own.

Today's practices pose significant upfront and testing effort. Much latent knowledge remains unexpressed and not easily available to learners; it must be exposed, cleaned and vetted. Further upfront effort needs to be spent on selecting the features (variables) used and then to accurately label the positive and negative training sets. Without “gold standards” — at still more cost — it is difficult to tune and refine the learners. The cost to develop tailored extractors, taggers, categorizers, and natural language processors is simply too high.



So recent breakthroughs demonstrate the promise; now it is time to systematize the process and lower the costs. The insight behind Cognonto is that existing knowledge bases can be staged to automate much of the tedium and cost now required to set up and train machine learners for knowledge purposes. Cognonto's mission is to make knowledge-based artificial intelligence (KBAI) cheaper, repeatable, and applicable to enterprise needs.

Cognonto puts its insight into practice through a knowledge structure, KBpedia, designed to support AI, and a management framework, the Cognonto Platform, for integrating external data to gain the advantage of KBpedia's structure. We automate much of the tedium and reduce costs in many [areas](#), but three of the most important are:

- Pre-staging labels for entity and relation types, essential for supervised machine learning training sets and reference standards; KBpedia's structure-rich design is further useful for unsupervised and deep learning;
- Fine-grained entity and relation type taggers and extractors; and
- Mapping to external schema to enable integration and interoperability of structured, semi-structured and unstructured data (that is, everything from text to databases).

KBpedia is a computable knowledge structure resulting from the combined mapping of six, large-scale, public knowledge bases — [Wikipedia](#), [Wikidata](#), [OpenCyc](#), [GeoNames](#), [DBpedia](#)

and [UMBEL](#). The KBpedia structure separately captures entities, attributes, relations and topics. These are classed into a natural and rich diversity of types, with their meaning and relationships logically and coherently organized.

Each of the six knowledge bases has been mapped and re-expressed into the KBpedia Knowledge Ontology. KKO follows the universal categories and logic of the 19th century American mathematician and philosopher, [Charles Sanders Peirce](#). KKO is a computable knowledge graph that supports inference, reasoning, aggregations, restrictions, intersections, and other logical operations. KKO's logic basis provides a powerful way to represent individual things, classes of things, and how those things may combine or emerge as new knowledge.

KBpedia contains nearly 40,000 *reference concepts* (RCs) and about 20 million entities. The combination of these and KBpedia's structure results in nearly 7 billion logical connections across the system. About 85% of the RCs are themselves entity types — that is, 33,000 natural classes of similar entities such as 'astronauts' or 'breakfast cereals' — which are organized into about 30 “core” typologies that are mostly disjoint (non-overlapping) with one another. KBpedia has extended mappings to a further 20 other vocabularies, including schema.org, Dublin Core, and others; client vocabularies are typical additions. The typologies provide a flexible means for slicing-and-dicing the knowledge structure; the entity types provide the tie-in points to KBpedia's millions of individual instances (and for your own records). KBpedia is expressed in the semantic Web languages of OWL and RDF. Thus, most W3C standards may be applied against the KBpedia structure, including for linked data, a standard option.

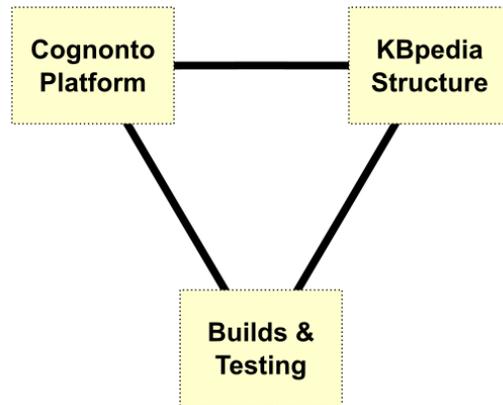
Cognonto has services to incorporate your own domain and business data, which are critical to fulfill domain purposes and to respond to your specific needs. We transform your external and domain data into KBpedia's canonical forms for interacting with the overall structure. Such data may include other public databases, but also internal, customer, product, partner, industry, or research information. Data may range from unstructured text in documents to semi-structured tags or metadata to spreadsheets or fully structured databases. The formats of the data may span hundreds of document types to all flavors of spreadsheets and databases.

KBpedia is purposefully designed to enable meaningful splits across any of its structural dimensions — concepts, entities, relations, attributes, or events. Any of these splits may be the computable basis for training taggers, extractors or classifiers. Standard NLP and machine learning reference standards and statistics are applied during the parameter-tuning and learning phases. Multiple learners and recognizers may also be combined as different signals to an ensemble approach to overall scoring. Alternatively, KBpedia's slicing-and-dicing capabilities may drive export routines to use local or third-party ML services under your own control.

Cognonto's modular technology is based on Web-oriented architectures. All functionality is exposed via Web services and programmatically in a microservice design. The technology for Cognonto resides in three inter-related areas:

- **Cognonto Platform** - the technology for storing, accessing, mapping, visualizing,

- querying, managing, analyzing, tagging, reporting and machine learning using KBpedia;
- **KBpedia Structure** – the central knowledge structure of organized and mapped knowledge bases and their millions of instances; and
- **Build Infrastructure** - repeatable and modifiable build and coherence and consistency testing scripts, including reference standards.



The Cognonto Web services may be manipulated directly from the command line or via cURL calls, or by simple HTML interfaces, by SPARQL, or programmatically. The Web services are written in Clojure and follow literate programming practices.

Cognonto's Web site offers a [demo](#) of a portion of its capabilities. Plus, the base KBpedia [knowledge graph](#) may be [explored interactively](#) with sample exports of its content.

Cognonto offers services across the full spectrum of KBAI applications, including:

- [Machine Learning Services](#);
- [Taggers, Extractors, Classifiers](#);
- [Content Harvesters/Publishers](#);
- [Data Prep and Staging](#);
- [Schema Mapping and Integration](#);
- [Knowledge Graph Services](#);
- [Dedicated SaaS](#); and
- [Platform Deployment](#).

Cognonto knows there are many important uses — and an upside of potential — for codifying knowledge bases for AI and machine learning purposes. Let us know how we can help you leverage your own domain and business data using knowledge-based AI.

About Cognonto

[Cognonto](#) (a *portmanteau* of 'cognition' and 'ontology') exploits large-scale knowledge bases and semantic technologies for machine learning, data interoperability and mapping, and fact and entity extraction and tagging.

Cognonto has been under development for more than two years, though the genesis of the venture reaches back nearly six years. Its principal founders have been leaders in the semantic technology space, with most recent efforts focused on KBAI. Cognonto LLC is based in Coralville, Iowa, USA, with offices in Québec City, Québec, Canada.