



Semantic Knowledge Source Integration

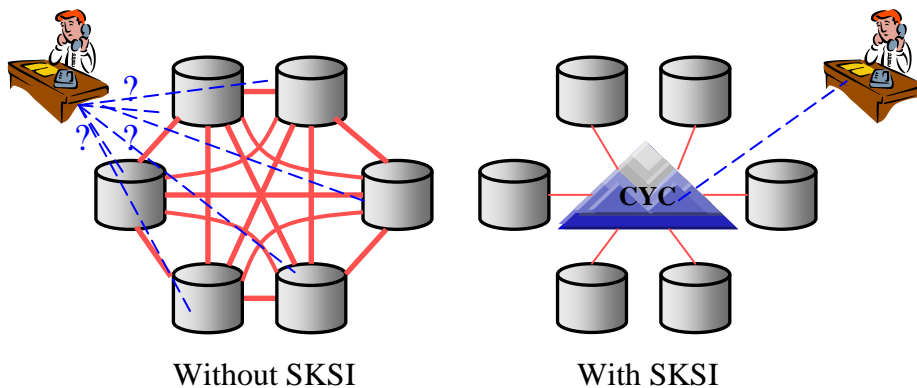
SKSI is a Cyc technology that enables information analysts to semantically integrate, access, query, and reason over disparate structured data sources such as databases, spreadsheets, websites, and triple stores. It combines data from many diverse sources, forming a coherent, uniform information nexus.

Highlights

- ◆ Seamlessly integrates knowledge and data bases.
- ◆ Helps answer questions that were previously answerable only by looking up data in each knowledge source and manually cross-referencing it.
- ◆ Provides a sophisticated audit trail and justification for its results.

At the heart of Cyc® is an ontology of over 300,000 general terms and a knowledge base (KB) of over three million facts and rules about those terms expressed in predicate calculus. The KB represents both broad human consensus reality knowledge and domain-specific knowledge – including fields such as intelligence analysis, counter-terrorism, and medical research, among others – in a semantically rich and precise manner.

Cycorp was incorporated in 1994 and is focused on developing and applying Cyc, which consists of the world's largest context-faceted knowledge base of common sense and domain knowledge, natural language I/O, and a state-of-the-art inference engine.



Problem: Disparate Heterogeneous Sources

Many analysts and decision makers in the government, military, and business sectors work amidst a sea of information from disparate sources and in incompatible formats. The analyst is responsible for identifying, extracting, and integrating the required information, often with minimal support from the individual systems. Thus, the analyst must spend valuable time cognitively sorting through the available sources to get at the truly relevant information.

Solution: An Intelligent Data Source Mediator

Semantic Knowledge Source Integration (SKSI) allows a knowledge base to act as an intelligent data source mediator. Users can issue queries without knowing in advance where the data contributing to the answer will come from. Thus, SKSI functions as a “schema-less query language”.

SKSI relies on a declarative mapping between external sources of structured knowledge and the Cyc Knowledge Base. Mapped knowledge sources can be treated as just another part of the Knowledge Base. When the SKSI system needs information to answer a user's query, it transparently and automatically identifies the appropriate source, generates and issues an appropriate request (e.g., SQL query to a relational database, HTTP request to a website, etc.), extracts the needed data, and semantically integrates the results.

Compatible with Most Structured Data Source Types

SKSI supports mapping SQL and SPARQL databases, CORBA APIs, and queryable web sites. By describing a data source using the Cyc/SKSI mapping vocabulary, its contents become part of a virtual, integrated knowledge base and are available for querying and reasoning by the Cyc inference engine and query mechanisms.

Unlocks the Power of Cyc Inference

Cyc contains a general-purpose inference engine plus a library of over 1000 special-purpose inference modules each optimized to make some common class of inferences (e.g., transitive relation reasoning) very efficient. Cyc is OWL-compliant, has a formal API, and can interface not only to structured external sources, but also (via clarification English dialogue) to human experts and unstructured (textual) documents.

