

Improved navigation and search and enabling better content discoverability are the key to enhance end-users experience. Semantic enrichment of content can offer extended user engagement and boost content usage. Semantic enrichment is the study of “meaning” behind the text, extraction of concepts from content and relating them for better knowledge discovery. Semantic enrichment value chain starts with insertion of conceptual codes in structured documents and extends to knowledge discovery and visualization using simple user query, which brings the real value of the content enrichment in facilitating research processes, increasing usage and driving more traffic.

With over 10 years of association with several leading publishers and information providers, Scope has recognised the need among publishers and information providers for semantically enriching their content. To meet this growing need for semantic enrichment Scope has developed *SemantiCz™*. The platform uses a combination of NLP rules, subject control vocabulary and subject matter expert (SME) to ensure high-quality output.

SemantiCz™ offers entity and concept extraction based on control vocabularies, emerging concept extraction, text classification, sentence extraction, parts of speech (POS) tagging, controlled vocabularies (CV)-based semantic tagging, generating relationships among concepts using Ontologies.

Service Differentiators

- Custom automation rules by developing a knowledge framework for any given domain, using SMEs
- Improved semantic enrichment through self-learning algorithms to understand the context and content of the input data, by comparing its existing knowledge repositories
- Domain experts to review and validate the extracted data to ensure relevancy
- The application does not require licensing or client side installation, and thus incurs no upfront cost
- System is highly scalable in features and volume; becomes increasingly robust while supporting rapid processing of content
- Customizable solution to accommodate clients’ unique needs
- Seamless integration of Scope’s processes with clients’ workflow

SemantiCz has the following Components:

Input handling: Scope can handle input documents in PDF, MS Word and XML formats, which will undergo pre-processing prior to the automatic extraction.

Automatic Extraction and Tagging Process: SMEs develop the knowledge framework, which is integrated into Scope's *SemantiCz™* platform for entity and concept extraction, text classification, sentence extraction, parts of speech (POS) tagging, controlled vocabularies (CV) based concepts, generating relationships among concepts using Ontologies, emerging/new concepts.

SME-Curation: SMEs validate and update the entities, concepts extracted using control vocabulary, emerging/new concepts based on NLP algorithm and relationships among the terms in each article; any new entities/concepts and relationships identified are fed back into the platform to enable self learning and updated in the control vocabulary.

Identification of Emerging Concepts: SMEs identify the emerging concepts through a combination of document level analysis and the timeline of the entire corpus for a concept.

Quality Audit: A quality audit is performed against quality parameters such as accuracy and comprehensiveness

Output: The final output can be delivered in RDF/XML formats, which can be used for query-based semantic search in search platforms.

For more information, please contact:

Richard Kobel

Vice President,
Business Development
USA
Scope e-Knowledge Center
Tel: +1 516 462 3555
rkobel@scopeknowledge.com

John Camarano

Associate Vice President,
Business Development
USA
Scope e-Knowledge Center
Tel: +1 301 538 6034
JCamarano@scopeknowledge.com

Deborah Harman

Associate Vice President,
Business Development
UK & Europe
Scope e-Knowledge Center
Tel: +44 770 203 6704
Deborah@scopeknowledge.com

