A Content Indexing Solution

With dramatic increase of e-content, end-users rely on online sources to find the information of their interest. But a major issue faced by the information providers and their customers is discoverability. An inaccurate indexing would lead to irrelevant search results, while incomplete indexing would result in missing relevant documents for a user query. Thus, indexing becomes mandate for publishers and other information providers to ensure quick and accurate searching of content for end-users.

With over 10 years of experience in offering indexing solutions for several leading publishers and information providers, Scope has created the platform-enabled indexing services, *InDEXr*™. Scope's *InDEXr* offers publishers and online information providers a unique opportunity to enhance the discoverability of their content by providing accurate key concepts/entities for a wide range of textual (technical literature such as patents, journals, articles, books and book chapters, news articles, e-learning content and product data) and non-textual (images, multi-media such as audio and video) documents.

InDEXr uses a judicious combination of proprietary software, in the form of statistical, linguistics and Natural Language Processing (NLP) rule, and subject matter experts (SMEs). The platform auto extracts key concepts/entities from input documents based on frequency, co-occurrence and location heuristics. SMEs then validate the auto-generated index terms to ensure relevancy and accuracy.

InDEXr Differentiators

- Ability to handle varied input formats, such as TEXT, PDF, MS Word, XML ,HTML, images, audio and video, and deliver output in client preferred formats, including TEXT, MS Word or XML
- Capable of converting raw, unstructured and semi-structured information into structured information
- Indexing expertise in various domains, such as Science, Engineering, Technology, Medicine, Social Sciences, Economics, Humanities and Education
- SEO-friendly keywords to enable better findability
- Scalable to handle large volume of production in short time, thus reduces indexing bottleneck

www.scopeknowledge.com

A Content Indexing Solution

InDEXr has the following components:

Scope's indexing solution, *InDEXr*, has the following components:

Keyword Indexing

The proprietary platform automatically extracts keywords from input documents and ranks them. Scope's SMEs validate the auto-extracted keywords and based on their relevancy, create the final keywords list.

Controlled Vocabulary (CV) Indexing

InDEXr can integrate clients' CV (such as thesaurus), extract the existing terms, and rank them for SME validation and selection of the preferred terms. In addition, *InDEXr* can also facilitate displaying of broader and narrower terms for a CV term. Further, the SMEs suggest new keywords to be updated in the CV.

Indexing of Named Entities

InDEXr can extract details such as people, places, organizations etc. for indexing.

Geospatial Tagging

InDEXr is designed to tag places with their respective geological coordinates.

Image Indexing (Graphs, Tables, Photographs)

InDEXr can extract keywords from the captions of images; it further can index images using controlled vocabulary. Data from tables and graphs also can be extracted to enable technical data search.

Indexing Multimedia Content

Scope can transcribe multimedia content, such as audio and video, into searchable text and extract metadata and index the content with the aid of *InDEXr*.

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