



WHITE PAPER

Concept Searching...

What It Is and Why You Need It

ICONECT Development, LLC
3601 N. Aviation Boulevard
Suite 2800
Manhattan Beach, CA 90266
(310) 643-8555
www.iconect.com

Content Analyst Company, LLC
11720 Sunrise Valley Drive
4th Floor
Reston, Virginia 20191
(888) 349-9442 Toll Free
www.contentanalyst.com

With the implementation of FRE 502, the use of advanced analytical software such as concept searching in your review platform becomes an important tool in preventing inadvertent waiver of privilege. Concept searching can also play a critical role in organizing large document collections as part of your EDRM workflow, and in accelerating the identification of responsive documents. So what is concept searching?

The Sedona Conference recently tried to define concept searching as “the combination of [a] query term and the additional terms identified by the thesaurus.” While it is admirable that Sedona tried to define concept searching in basic terms, their definition misses the mark. Worse, some of the more popular culling tools are confusing the idea of synonyms (or topics) with concepts, and may not be providing true concept searching.

Concept searching goes beyond synonym or keyword matching to include all documents that describe the same subject matter or concepts regardless of the specific terms or words used. For example, concept searching understands that the words “terminating” and “firing” and also the phrase “ending association with” are all describing the same idea or meaning. The ability to leverage the power of true concept searching and advanced analytics is critical to the ability to perform intelligent review and retain the clawback rights outlined in FRE 502.

To use another example, concept search engines understand that spam might refer to either unwanted emails or a canned meat product, and they understand which meaning is relevant to your document collection.

In some concept search schemes, a process called “natural language processing” is used. This process relies on outside “structures” like dictionaries, thesauruses and ontologies to find documents based on human-coded mappings of terms which can be subjective, narrow and incomplete.

Machine-based concept searching (e.g., using latent semantic indexing) derives meaning from documents through a rigorous mathematical analysis of the relationship between terms across documents. It requires no outside assistance from people or

dictionaries, and this enables machine-based concept searching to keep up as language evolves, whereas human-coded mappings require continual enhancements to remain current.

So if concept searching is about meaning, and keyword searching is about words, which search method should you use? The answer is that it’s not an “either/or” decision. Just like keyword searching, concept searching is another important resource to deal with an exploding volume of ESI (electronically stored information). You need both.

The integrated review tool is key. Today we all start with keywords: we know how a Google search works, and the courts—while becoming increasingly frustrated with its limitations—understand Boolean search. For

some cases, keyword searching may be sufficient to organize and review the material: it may be a small volume of documents, the case may be very straightforward, and the terms may be very specific.

Many other cases aren’t so straightforward for a variety of reasons: they have massive amounts of ESI; they involve complicated issues; they contain potential fraud or deliberately obscured terminology; or they might be in multiple languages. The review tool must be capable of combining keyword and concept searching, and organizing the search results effectively. For example, you must be able to use “find similar” capabilities to take one highly responsive document and plough through masses of ESI using concept searching to find all the other documents that have the same meaning and talk about the same concepts.



Concept Search as Preparation

Concept search and advanced analytic engines can also be used as part of a prereview analysis to organize related documents into concept-based folders. This enables the legal team to identify what issues or concepts are present in the overall collection, and your top reviewers can then focus on the key folders. This filtering capability reduces review time, improves review accuracy and provides significant savings on overall litigation costs.

Multinational cases add another layer of complexity and require technologies that enable you to search for documents in languages such as Chinese, Russian and Korean without first translating to English. If the concept search engine and associated review platform do not support Unicode based data, you are locked out of two-thirds of the world's market.

An advanced and proven review platform that can seamlessly integrate concept-based folders as well as support Unicode-based data is required to take full advantage of concept searching. When the review tool

is also web based, the power of concept searching is extended to branch offices and consultants for cost-effective collaboration. It is also important that the review platform includes integrations with e-discovery collection, processing and production tools so you have a streamlined end-to-end EDRM workflow.

Companies reviewing documents simply can't afford to have their best reviewers trudging through volumes of ESI when they could assign potentially non-responsive material to less-expensive contractors. The courts are also now demanding more accurate search strategies—they don't want to render judgments based on incomplete information because the parties used search tools and techniques that were inadequate.

This is where concept searching comes in. When properly employed and incorporated into an already robust review platform, concept searching offers a unique way to contain costs and manage time while improving the quality of the review. Few other technologies can offer a similar benefit, and fewer still put that benefit so easily within reach.

About Content Analyst Company and iCONNECT Development, LLC

Content Analyst Company is a provider of advanced search and document analytics software to e-discovery providers and the public sector; headquartered in Reston, VA, they can be reached at 888-349-9442, or info@contentanalyst.com.

iCONNECT Development, LLC provides award-winning litigation support and document review software trusted by Am Law 100 firms, corporate legal departments, Fortune 500 corporations and government agencies. Visit www.iconect.com or send inquiries to info@iconect.com.

iCONNECT and Content Analyst Integrated Solution

iCONNECTnXT, a Web-based hosting and review platform, is tightly integrated with e-discovery collection, processing, analysis and production tools for a seamless workflow, and now also includes an integration with CAAT, the concept search technology from Content Analyst.