Searching for That Four-Leaf Clover in a Clover Field of Data? Streamline the Process With Enterprise Search

Article contributed
by dtSearch®

Have you ever sat down in a clover patch and immediately found a four-leaf clover? Me neither. Most folks don't have that kind of luck. Similarly, in the digital world, most people aren't fortunate enough to quickly happen upon the file or email that they are looking for amidst millions of other files and emails. But enterprise search can turn the table on luck, if not in a clover field, at least when it comes to locating data.
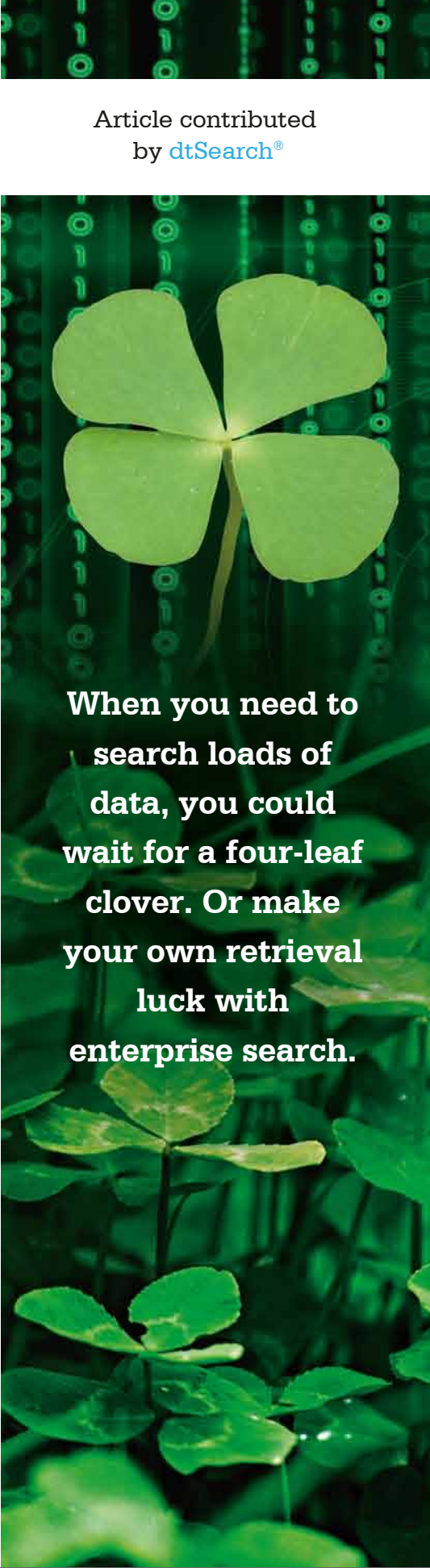
Enterprise search covers search engines like dtSearch® that delve into enterprise data, as opposed to search engines like Google or Bing that scan the public Internet. Before enterprise search can instantly search terabytes, it first needs to index the data. Happily, unlike locating that four-leaf clover, indexing is easy. Simply tell the search engine the folders, email repositories and other archives to index, and let the indexer take it from there.

No need to even tell enterprise search what types of data it is indexing. The indexer can on its own figure out if something is a PDF, an email, a Microsoft Office document format like Word, Access, Excel, PowerPoint or OneNote, etc. Enterprise search can automatically handle compressed formats like RAR or ZIP. It works seamlessly with multilevel nested formats, like an email with a compressed attachment with a Word document inside that itself embeds an Excel spreadsheet.

After indexing, 25 search features let you instantly find what you are looking for. Suppose you need to retrieve all files or emails with information on Clover Enterprises. You could do an "any words" or an "all words" or an exact phrase search. Or you could put together a more complex Boolean and/or/not and proximity search request: *Clover Enterprises and (takeover threat or poison pill) and not (cloverleaf w/2 texas)*. You can search across the full text of all data or limit your query to only files and emails that might contain the word *confidential* in certain metadata.

Let's say you are searching OCR'ed PDFs that may contain minor scanning errors, or emails that may contain

When you need to search loads of data, you could wait for a four-leaf clover. Or make your own retrieval luck with enterprise search.

misspellings. You can adjust fuzzy searching from 1 to 10 to sift through such typographical errors. That way, if an OCR program rendered *clover* as *clovar* or if someone mistyped *clovar* instead of *clover* in an email, you'd still find it.

Now suppose you want to extend your search to related concepts. The built-in thesaurus has trefoil as a synonym of clover and shamrock as a related word, so a concept search could automatically extend to those as well. Or if *Clover Enterprises* had a previous name of *Fortune Enterprises*, you could make *clover* and *fortune* custom synonyms for purposes of this inquiry.
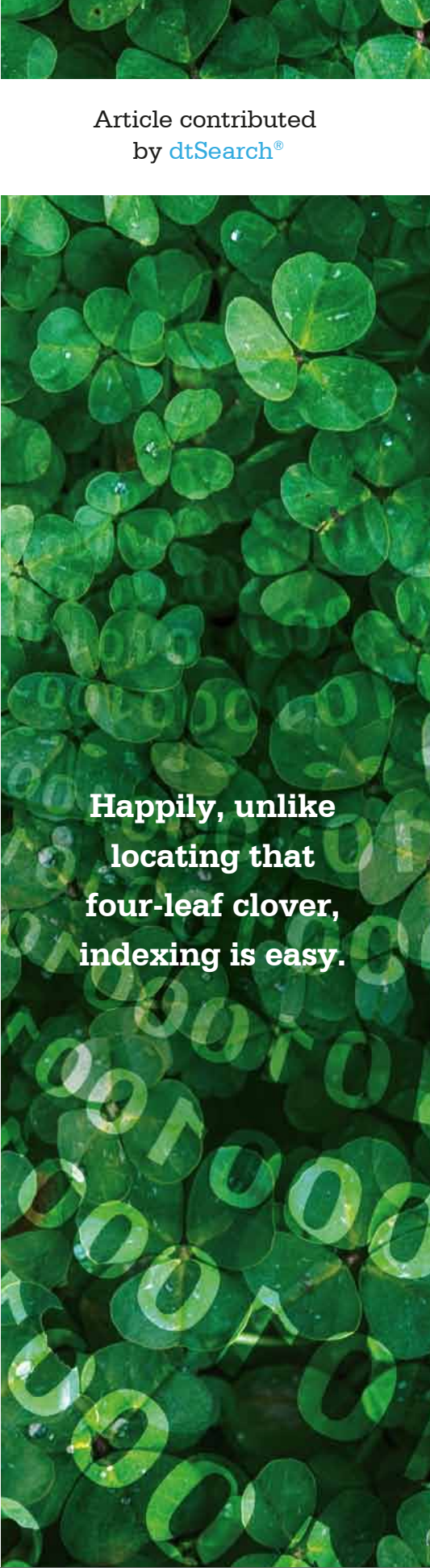
Beyond word and phrase searches, enterprise search can also search for numeric expressions. If Clover Enterprises is associated with case number *76323*, you could add a search for that. Or if you wanted to cover any case number from *76319* to *76367*, you could do a numeric range search extending throughout that range. You can also supplement that with a date range query covering *3/1/18* to *4/15/19*. Date range search can pick up not only dates in the XX/XX/XX format but other common date formats as well, letting you find a mention of *March 17, 2018*. Enterprise search can even identify any valid credit card numbers that may appear in data.

Search covers not only English text, but other international Unicode text as well, including European languages, left-to-right language like Arabic and Hebrew, and double-byte Chinese/Japanese/Korean text. Enterprise search also offers multiple options for relevancy-ranking search results. Or after a search, you can immediately re-sort by some completely unrelated criterion like file date or file location. Then browse the full text of retrieved files with highlighted hits for easy review.

Best of all, enterprise search doesn't just enable individual searching, it also enables multiple concurrent instant search threads so everyone at once can navigate the data. In an online or cloud-based environment, enterprise search threads can operate in a stateless manner, with no built-in limit on the number of concurrent search threads. Concurrent searching can further continue unaffected while one or more indexes automatically update to reflect new content.

In sum, when you need to search loads of data, you could wait for a four-leaf clover. Or make your own retrieval luck with enterprise search.

Happily, unlike locating that four-leaf clover, indexing is easy.