

Super-Caffeinate Enterprise Search

Article contributed
by dtSearch®

Recommendations for the number of cups of coffee and tea to drink keep increasing. A major medical group recently raised that number to an eye-popping 5 cups of coffee per day. And I am personally committed to drinking every last drop. For 5 cups of coffee, I've got 5 ways to super-caffeinate enterprise search.

What is the first way?

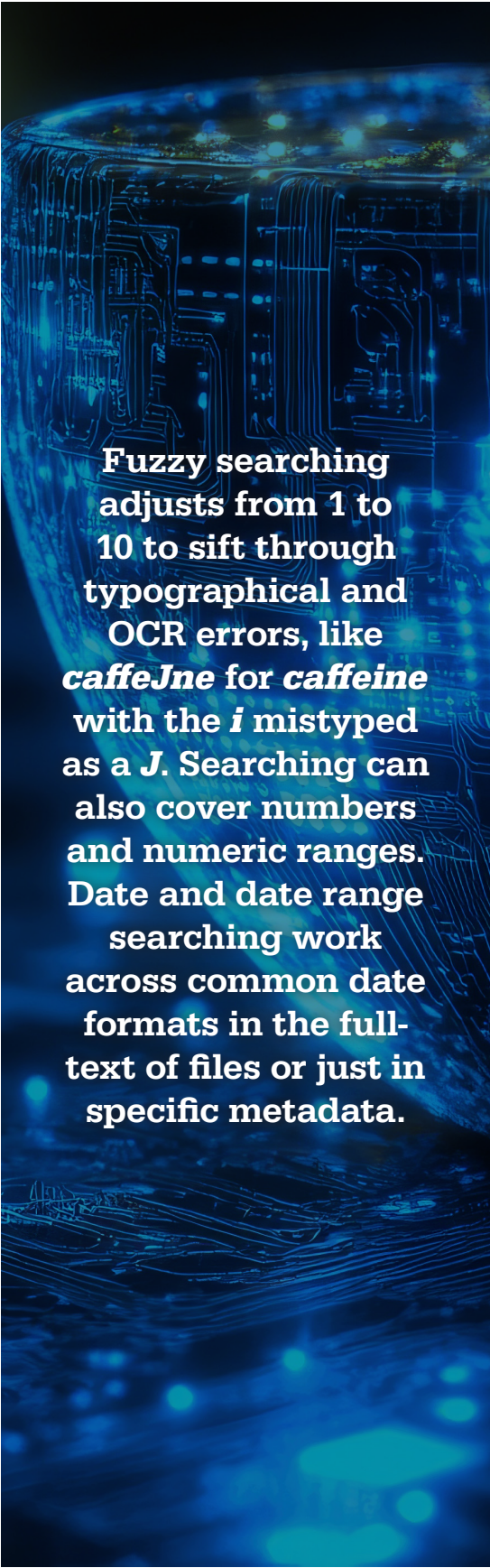
The first way is to make sure that you do indexed search. There are two basic types of enterprise search. The first is unindexed search, akin to a brute force march through data. This brute force march is effective but slow as enterprise search has to start from scratch in processing each search request. The second way is indexed search. Here, enterprise search first indexes the data. Indexing pre-collects each unique word and number across the data and records the position of each in the data.

Can you say more about indexing?

Indexing is easy to get going—even before your first cup of coffee kicks in. Just tell dtSearch the folders to index and the software will take it from there. The indexer automatically recognizes PDFs; Microsoft Word, Access, Excel, PowerPoint and OneNote; different email formats; web-based formats; etc. You can have a Word document with a .PDF extension or a PDF with a .DOCX extension and the dtSearch indexer can handle that correctly as it looks to the binary file itself to determine the file type. The indexer can also automatically work through multilevel nested formats, like an email with a ZIP or RAR attachment containing a Word document with an embedded Excel spreadsheet.

What about remote data?

So long as files appear in the Windows folder system, the files can be interchangeably local or remote, such as One Drive / Office 365 files, SharePoint attachments and DropBox files. Indexing is resource-intensive,



Fuzzy searching adjusts from 1 to 10 to sift through typographical and OCR errors, like *caffeJne* for *caffeine* with the *i* mistyped as a *J*. Searching can also cover numbers and numeric ranges. Date and date range searching work across common date formats in the full-text of files or just in specific metadata.

although a 64-bit multithreaded indexer can greatly speed up the process. A single dtSearch index can hold up to a terabyte of text, and there are no limits on the number of indexes that the software can create and instantly concurrently search. While indexing is resource-intensive, indexed searching is resource-light, enabling instant multithreaded searching across terabytes. And the indexer can automatically update indexes to reflect new files, edited files and deleted files without interrupting ongoing concurrent searching.

What about international language text?

dtSearch automatically works with Unicode, covering hundreds of different languages in files. A single file can go from English to a different European language and alphabet to double-byte Asian text to right-to-left Middle Eastern text and then back to English and Unicode and dtSearch will track all of that.

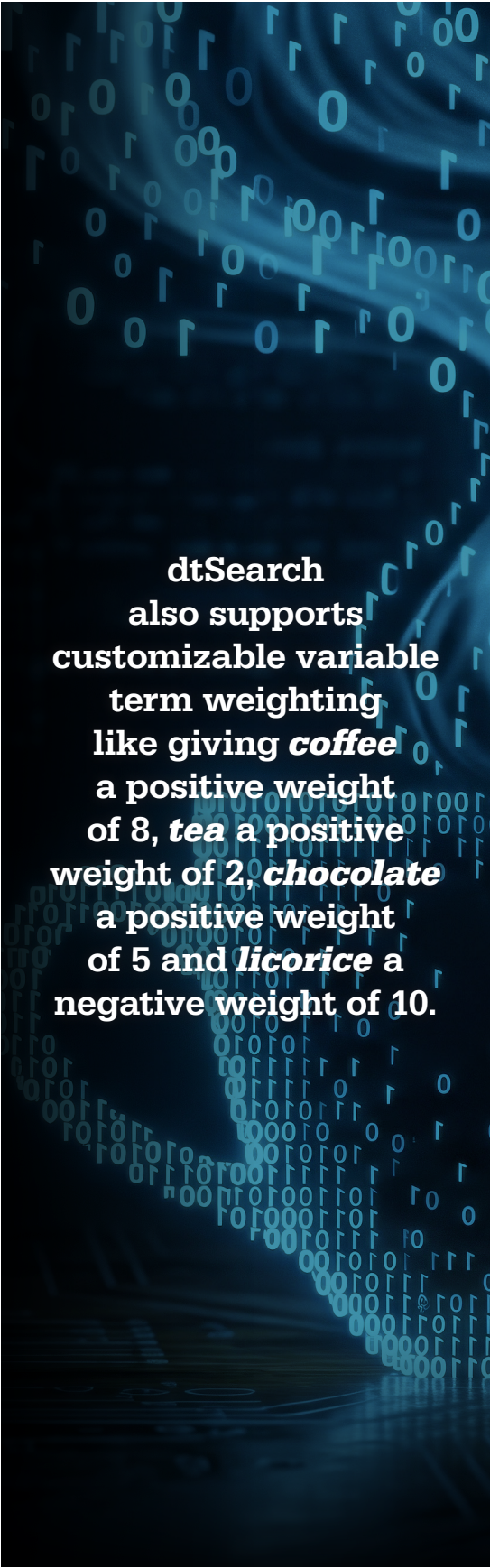
Beyond indexing, how else can you super-caffeinate enterprise search?

Choose from a range of precision search features to caffeinate your search requests. Sure, dtSearch supports basic search options like “all words” or “any words.” But for larger volumes of data, precision phrase, Boolean and proximity search options can really go that extra mile in getting you exactly to what you are looking for. Compare an “any words” search for *coffee* or *tea* to (“*cup of coffee*” or “*cup of tea*”) and (“*dark roast*” w/12 *caffeine*) and not *soda*. Concept searching lets you expand a search query to synonyms, like *java* for *coffee*. Further fine-tune a full-text search request with metadata requirements, like *subject metadata* includes “*population studies*” and not “*tea ceremony*”.

Any other ways to caffeinate search?

Fuzzy searching adjusts from 1 to 10 to sift through typographical and OCR errors, like *caffeJne* for *caffeine* with the *i* mistyped as a *J*. Searching can also cover numbers and numeric ranges. Date and date range searching work across common date formats in the full-text of files or just in specific metadata. A search for *date(9/21/26 to 11/1/26)* would pick up not only 10/12/26

Article contributed
by dtSearch®



**dtSearch
also supports
customizable variable
term weighting
like giving *coffee*
a positive weight
of 8, *tea* a positive
weight of 2, *chocolate*
a positive weight
of 5 and *licorice* a
negative weight of 10.**

but also *October 12 2026* and *Oct 12 2026*. dtSearch can even find valid credit card numbers across indexed data.

What about search sorting?

By default, dtSearch relevancy ranks retrieved files using the density and rarity of hits across indexed files. Take an “any words” search for *coffee tea chocolate*. If *coffee* and *tea* are prevalent across indexed data but *chocolate* only appears a few times, then *chocolate* files would score more highly for relevancy, with files with the densest *chocolate* mentions coming out on top. But you can super-caffeinate search results ranking as well.

How so?

dtSearch also supports customizable variable term weighting like giving *coffee* a positive weight of 8, *tea* a positive weight of 2, *chocolate* a positive weight of 5 and *licorice* a negative weight of 10. Variable term weighting can apply universally across indexed data, or with extra weighting to appearances in specific metadata or towards the top or the bottom of a file. Or do the search equivalent of blending your iced coffee in a shake by instantly re-sorting by a completely different metric like filename, file date or file location.

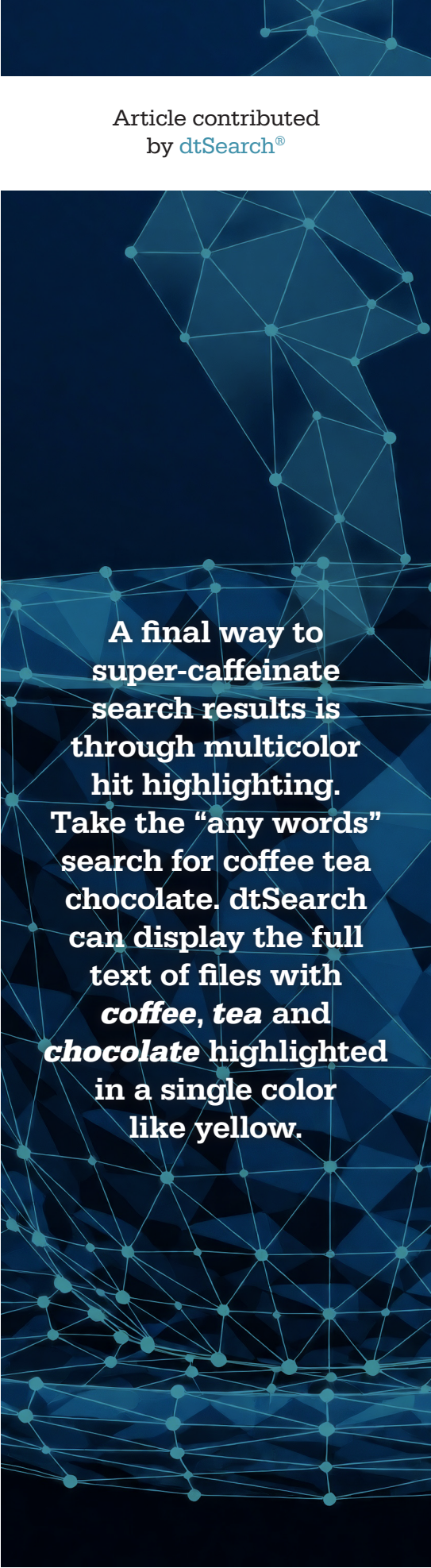
Any final way to super-caffeinate search results?

A final way to super-caffeinate search results is through multicolor hit highlighting. Take the “any words” search for *coffee tea chocolate*. dtSearch can display the full text of files with *coffee*, *tea* and *chocolate* highlighted in a single color like yellow. Or dtSearch can highlight different hits in different colors, like *coffee* in yellow, *tea* in light green and *chocolate* in light blue. Some find that multicolor hit highlights acts like that extra shot of espresso in supercharging search results viewing. And multicolor highlights definitely makes for a more interesting color palette in any case.

Final thoughts?

As you sip that extra cup of coffee or tea, go ahead and caffeinate your enterprise search as well. Download a fully-functional 30-day evaluation copy from dtSearch.com [dtSearch.com](https://www.dtsearch.com).

Article contributed
by [dtSearch®](https://www.dtsearch.com)



**A final way to
super-caffeinate
search results is
through multicolor
hit highlighting.
Take the “any words”
search for coffee tea
chocolate. dtSearch
can display the full
text of files with
coffee, *tea* and
chocolate highlighted
in a single color
like yellow.**