

This Summer, Sever the Data Ball and Chain with a Search Engine

Article contributed
by [dtSearch®](#)


With cell phone service nearly anywhere on the globe, few of us can ever go fully off the grid. An assistant, a boss or a client can have an urgent question requiring a quick bit of data lookup even (and sometimes it seems especially) if we are at the beach, biking in the country or half-way up a mountain. So how do you retrieve the necessary information under those circumstances?

- 1) You could lug around enterprise data like a ball and chain. How bad anyway is it to have a few grains of sand in your laptop case? And what's a little extra weight when you're biking uphill or climbing?
- 2) Or you could cut the data ball and chain and use a secure browser-based enterprise search interface to find the answer. Enterprise search is not a scan-the-Internet-looking-for-lodging-deals search engine like Bing or Google. Rather, is a product category like dtSearch® that lets an organization instantly and granularly search across terabytes of enterprise data.

While enterprise search can take the load off when you're out-and-about, it does take some foresight to set up. Enterprise search can instantly search across terabytes of documents, emails and other data, but only after it first indexes everything. Indexing identifies each unique word and number, and records the location of each occurrence.

Although indexing terabytes does take some time, from the human perspective, it is a "no brainer." All you have to do is point to the folders and the like to index and the indexer will take it from there. The indexer can even handle cloud-based data like Office365 or SharePoint in a similar manner to how it handles other local files as long as the cloud data is accessible through the Windows folder system.

The indexer works directly with the binary format of each item, bypassing the overhead of pulling up each file in its associated application. Parsing specifications can be hundreds of pages long and vary dramatically



**You could lug
around enterprise
data like a ball
and chain.**

for different file types. Matching the right parsing specification to each binary format is therefore critical. Fortunately, the indexer can use the binary format itself to figure out if something is a PDF; a Microsoft, Word, Access, Excel, PowerPoint or OneNote file; an email format; etc.

Because the indexer uses the binary format for file identification and access, a PDF with a misplaced .DOCX file extension will not impede the indexer's work. The indexer can further accommodate nested files, like an email with a ZIP or RAR attachment encompassing a Word document with a fully embedded Excel spreadsheet. Text that blends in with the background in a file's associated application, like black writing against a black background, is easily accessible in the binary format. The same is true for obscure metadata and even tracked changes that remain in a file.

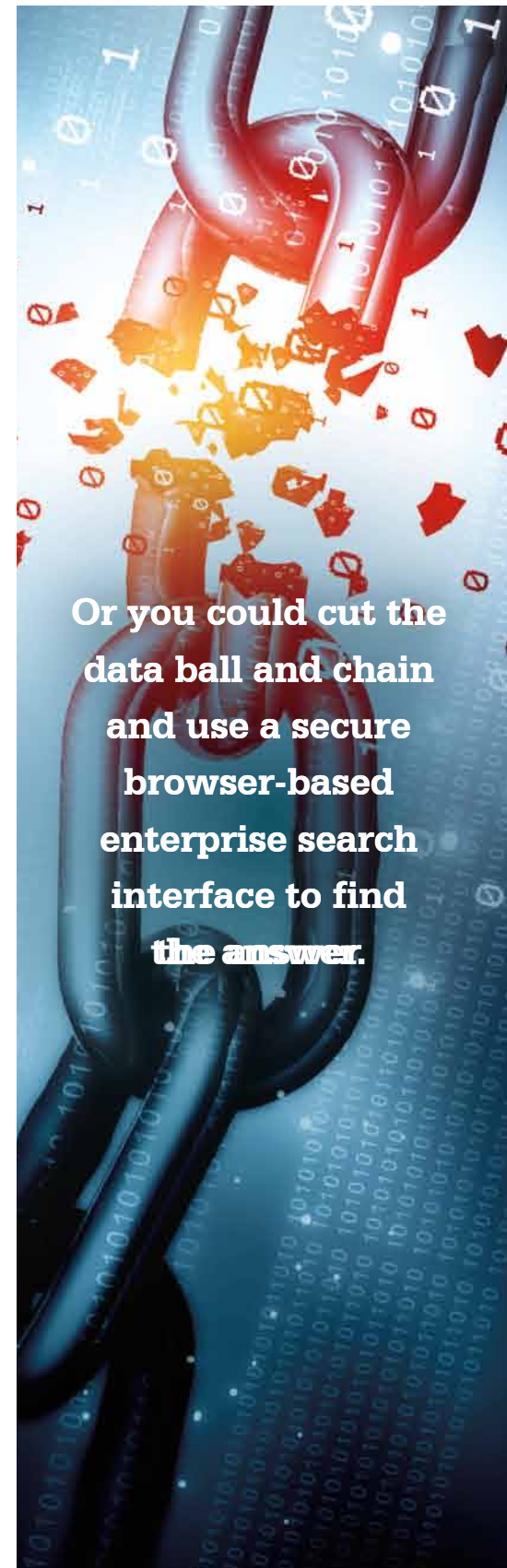
After indexing, search can run in a shared web-based environment for secure data access from the office, the beach, a back road or a mountain trail. And it is not just one search at a time that can proceed. Online search can run statelessly, with no built-in limit on the number of concurrent search threads, enabling everyone to simultaneously search.

Enterprise search offers over 25 different search options ranging from free-form unstructured "all words" or "any words" queries to precision word and phrase Boolean (and/or/not) and proximity expressions. Searching can cover all text and metadata, or target specific metadata only, by itself or in combination with full-text elements. Unicode support extends to hundreds of international languages, including double-byte Chinese, Japanese and Korean as well as right-to-left Hebrew and Arabic.

Fuzzy searching adjustable from 1 to 10 can sift through TYPOGRICKAL errors in emails and OCR'ed PDFs. Date and date range searching automatically extends across popular date formats like Jan 26, 2023 or 1/26/23 in text. A search can also find numbers or numeric ranges and number patterns, even identifying any credit card numbers in data. Lastly, enterprise search offers multiple options for relevancy-ranking and can display a full copy of each item with highlighted hits for convenient browsing.

So, this summer, leverage enterprise search to travel light, without the data ball and chain. And I'll race you up the mountain!

Article contributed
by [dtSearch®](#)



**Or you could cut the
data ball and chain
and use a secure
browser-based
enterprise search
interface to find
the answer.**