Playing hide and seek with your data? Here’s how to find that critical piece of information on your intranet – even from a tunnel with no online connection

Article contributed
by dtSearch®

No matter your job, it likely requires immediate answers from your enterprise Intranet. Indexed enterprise search lets you find what you need from anywhere, while cached indexed enterprise search lets you find what you need even when you are offline.

Together, these two components can help resolve what feels like a never-ending game of file hide and seek.

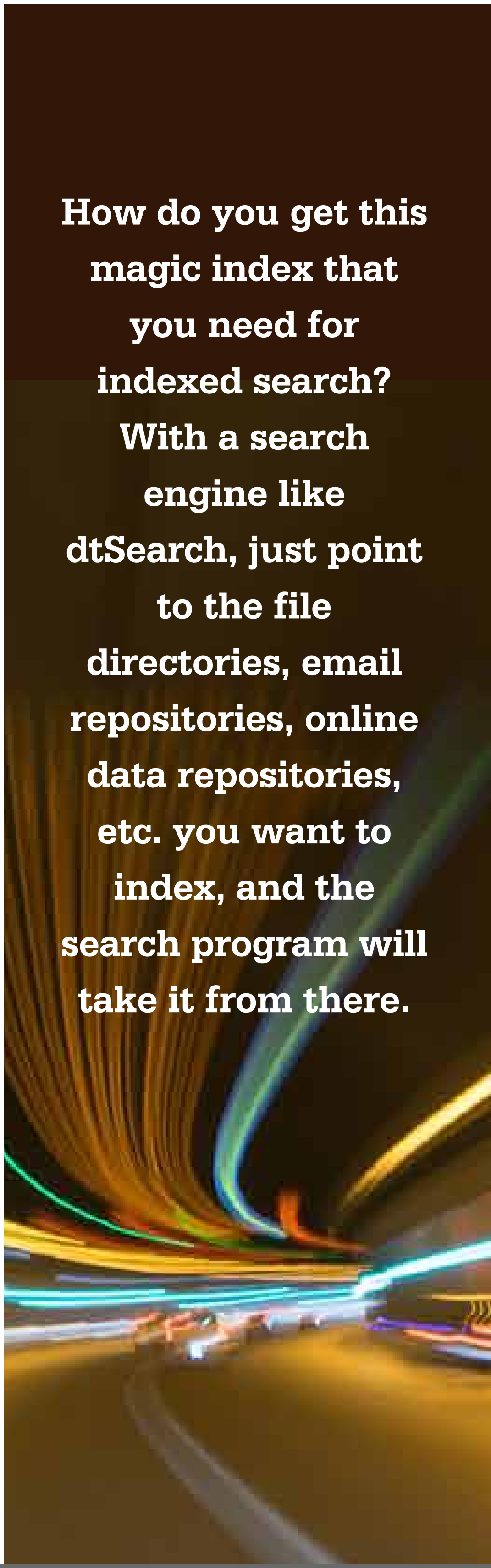
Indexed enterprise search. Suppose you are looking for information on a specific transaction. The timeline is a little hazy, but you think it was about 3 years ago. You could look at each PDF in the enterprise archives that looks like it might be relevant based on the file date and the file name in a PDF viewer like Adobe Reader. You can also look at each word processing document that might be relevant in Microsoft Word.

For completeness, you would also need to check any presentation file that might be relevant in PowerPoint, any potentially relevant spreadsheets in Excel, note files in OneNote, emails in Outlook or Exchange, web-based format like HTML, database records in Access, SQL, NoSQL, XML, etc. You’d further need to fully sift through any ZIP or RAR archives from around that time, whether such archives are in a standalone capacity or in the form of email attachments.

The alternative to going through everything individually is to do an indexed search across all online and offline data repositories. Four things make the indexed search a lot more efficient:

1. Indexed search bypasses the need to view each potential file, email, etc. in its native application. Instead, indexing approaches all files, emails, etc. in their "resting" binary format, avoiding the need to pull each up in its associated application.
2. Indexed search offers federated searching across all data sources comprehensively.
3. Indexed search includes over 25 different search features. More on that below.
4. Last, but certainly not least, indexed search even across terabytes is typically instantaneous.

How do you get this magic index that you need for indexed search? With a search engine like dtSearch, just point to the file directories, email repositories, online data repositories, etc. you want to index, and the search program will take it from there. No need to even tell the indexer what types of files and emails it is working with; the indexer can figure that out for itself. (As a side note, the indexer can use the contents of each binary file itself, rather than the filename extension, to determine the applicable file format. It is all too easy to save a PDF with a Word extension and vice-versa.)



How do you get this magic index that you need for indexed search? With a search engine like dtSearch, just point to the file directories, email repositories, online data repositories, etc. you want to index, and the search program will take it from there.

Reprinted with permission of BetaNews, betanews.com

After indexing, you can instantly search everything across all indexed data. In fact, multiple users can instantly search the indexed data concurrently. When you search, you can of course just look for specific words or phrases. But you can also leverage over 25 other search features, including multilevel Boolean (and/or/not) search expressions, proximity search (in either direction or one direction only), metadata-specific search, other positional search requests, regular expression search, numeric and numeric range search, date and date range search (with automatic date recognition across any number of date formats), wildcard search, concept search, and fuzzy search to sift through minor typographical and OCR errors. Indexed search can even identify all email addresses or find any valid credit cards.

Indexed search can automatically support any of the hundreds of Unicode international languages, including right-to-left languages like Hebrew and Arabic, and double-byte languages like Chinese, Japanese and Korean. After a search, indexed search can then sort (or instantly re-resort) retrieved items by default relevancy or any number of other sorting options. Indexed search can also display a full copy of retrieved files with highlighted hits.

Caching. When doing a standard indexed search, the search program will return to each file or email in its original location to get the data it needs to display that file or email with highlighted hits. But sometimes returning to the original document to display it with highlighted hits doesn't work. Maybe you have a huge file that is taking forever to load from a remote server. Or maybe the file or email no longer exists at all in its original location. Or maybe you're sitting dead-center in the middle of that offline tunnel.

This is where caching comes in. You can set up your indexes to include the full documents themselves in a compressed format. The cached version is then "ready to go" in an instant with highlighted hits, even when the original is slow or unavailable.

Now suppose after scrolling through a bunch of documents with highlighted hits, you want to see a search report summarizing each hit in context. The search engine can generate on-the-fly a search report showing each hit with as many words of context as you want from the index itself. You can also give the search engine a head start using a different type of text caching, the sole purpose of which is to generate instant keyword-in-context search reports.

The downside of both document caching and other text caching is that because the index includes additional data, the index will be larger than it otherwise would be. However, with today's abundant data storage options, the trade-off is usually worth it, even without a long tunnel in the middle of your regular commute.

Time to end the game of hide and seek with your data. Enterprise search with caching enabled can get you what you need, even when you are (heaven forfend!) offline.

Article contributed
by dtSearch®

Sometimes returning to the original document to display it with highlighted hits doesn't work. Maybe you have a huge file that is taking forever to load from a remote server. Or maybe the file or email no longer exists at all in its original location. Or maybe you're sitting dead-center in the middle of that offline tunnel. This is where caching comes in.

Reprinted with permission of BetaNews, betanews.com