

Set Up a Shared Search Repository for Your Organization So You Can Find What You Need – Even From the Beach

Let's suppose you are on the beach and your day is interrupted not by a great white shark, but a call from your co-worker. You just need to find that critical file to put an issue to rest. You could journey all the way back to the office. Or you could use a web-based search engine to securely find what you need as you leisurely sip your drink under an umbrella.

So how can you find what you need from the beach?

Before you head off to the beach, set up a secure web server with access to your organization's MS Office files, PDFs and even email files and nested attachments and install a search engine like dtSearch. dtSearch Web runs on any Windows IIS web server. The server can be sitting at your office or in the cloud say on Microsoft Azure. The Windows IIS web server would control security here. dtSearch Web doesn't require any development work to install, although it does have HTML5 forms that you can easily customize if you want.

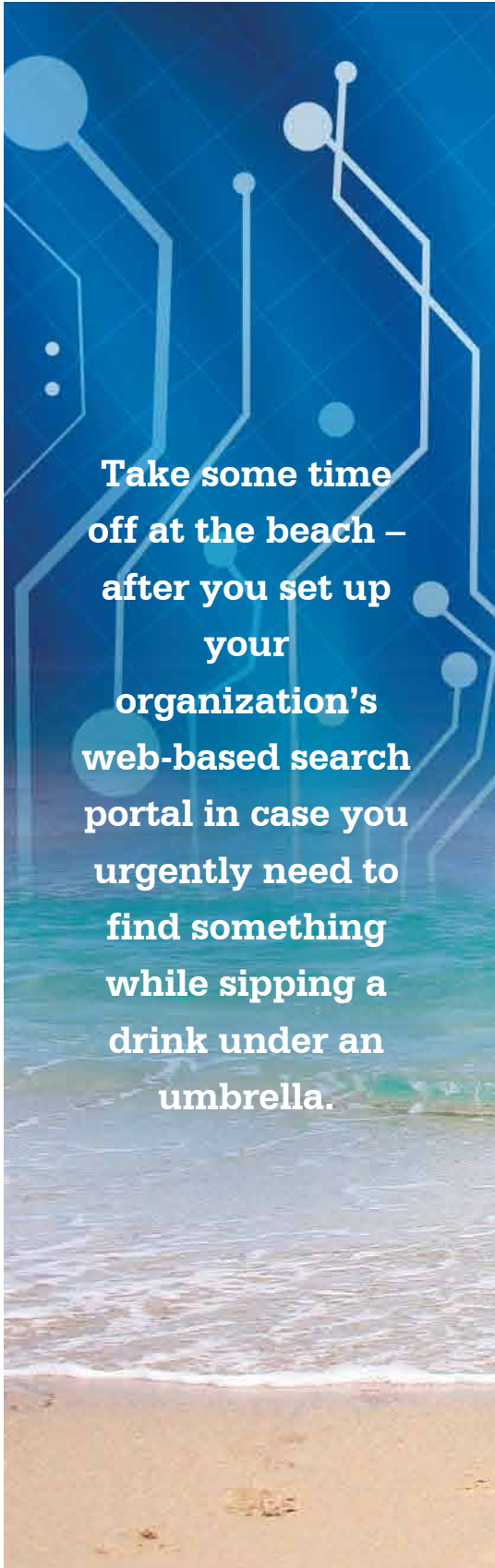
What if you are a developer?

Then you can use the dtSearch Engine SDK on a Linux, Mac or Windows server. The dtSearch Engine APIs also let you integrate dtSearch functionality with SharePoint, NoSQL or SQL data, including both database metadata and referenced file or BLOB data. The dtSearch Engine APIs offer faceted search and granular data classification options for security, all of which can work with any combination of database metadata, document metadata, document full-text contents or anything else you add to the index.

And can multiple people search the shared repository at once?

Both dtSearch Web and the dtSearch Engine support concurrent web-based searching. In fact, there are no limits on the number of search threads that the products can simultaneously process. Search time is typically instantaneous, even with multiple search threads and terabytes to search.

Article contributed
by [dtSearch®](#)



**Take some time
off at the beach –
after you set up
your
organization's
web-based search
portal in case you
urgently need to
find something
while sipping a
drink under an
umbrella.**

How does dtSearch search so quickly?

By first more slowly indexing all of the contents to identify each unique word or number and its location in the data. Indexing may sound like a lot of effort, but the search engine does all the work. Just click on the directories you want dtSearch to cover in its index and let the indexer do the rest.

Do you have to tell dtSearch what file formats you have?

No need to identify the file formats; the dtSearch indexer will automatically recognize popular file types like PDFs, MS Word, PowerPoint, Access, Excel and OneNote, as well as email formats including Outlook, Exchange and other popular email file types. The indexer can automatically go through ZIP or RAR archives. And it will also work through files embedded in other files. If you have an email with a ZIP attachment and inside that is a PowerPoint and embedded in the PowerPoint is a OneNote file, dtSearch will work with all of that automatically.

What if some of my files may have the wrong extensions, like PDF files with .DOCX extensions?

The dtSearch document filters which parse the files determine the relevant file type by looking inside the binary formats. The document filters do not rely on file extensions for this purpose. So if your PDFs have .DOCX extensions and your Word documents have .PDF extensions, it is not a problem.

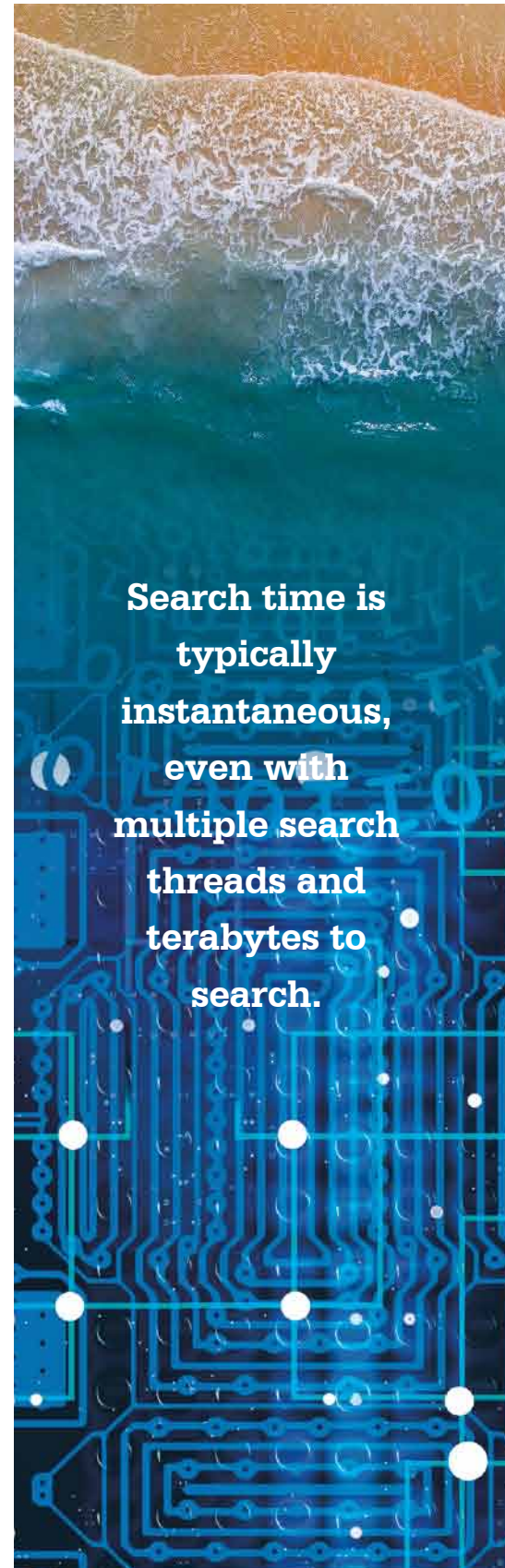
But what if I have Microsoft Office documents and emails that aren't web-ready?

The document filters can convert even non-web-ready content like Microsoft Word, Access, Excel, PowerPoint, OneNote, Outlook and Exchange files "on the fly" to HTML. That way, the search engine can display these in full with highlighted hits even if the files themselves are not web-ready.

But do the files have to be present to show them in full with highlighted hits?

Not if you activate caching when you build an index. With caching, you'll have a full copy of each file along with the index itself. dtSearch does not then need to return to the original files to display them with highlighted hits.

Article contributed
by [dtSearch®](#)



**Search time is
typically
instantaneous,
even with
multiple search
threads and
terabytes to
search.**

What search options are there?

dtSearch has over 25 different full-text and metadata search features. End-users can enter completely unstructured natural language search requests relying on relevancy-ranking based on hit term density and rarity to automatically go to the best matching files. Or end-users can enter more structured phase, Boolean, proximity, concept, etc. searching. dtSearch even has options to identify credit card numbers in data, or generate and search for file hash values.

What about typographical or OCR errors?

Let's say you mistype umbrella as umbrelta in an email. Or a blurry document that you OCR leads to a similar error. dtSearch fuzzy searching adjusts from 1 to 10 to sift through such misspellings. And fuzzy searching can work on top of other search options. Fuzzy searching even works with the wide range of international Unicode languages that dtSearch products support, everything from European languages to double-byte Chinese, Japanese and Korean text to right to left text like Hebrew and Arabic.

Anything else you'd like to add?

dtSearch has enterprise and developer products that run "on premises" or on cloud platforms to instantly search terabytes of "Office" files, PDFs, emails along with nested attachments, databases and online data. Because dtSearch can instantly search terabytes with over 25 precision search options, many dtSearch customers are Fortune 100 companies and government agencies. But anyone with lots of data to search can download a fully-functional 30-day evaluation copy from dtSearch.com

So take some time off at the beach – after you set up your organization's web-based search portal in case you urgently need to find something while sipping a drink under an umbrella.

Article contributed
by [dtSearch®](http://dtSearch.com)



**dtSearch has
enterprise and
developer
products that
run "on
premises" or on
cloud platforms
to instantly
search
terabytes of
"Office" files,
PDFs, emails
along with
nested
attachments,
databases and
online data.**