

# Optimize Sifting Through Data As You Work Remotely

Article contributed  
by [dtSearch®](#)

You've spent weeks fine-tuning a quiet place to work, getting the perfect houseplant for your desk, and adjusting the curtains "just so" to strike a balance between letting in daylight while avoiding catastrophic screen glare. But you still need to optimize how you sift through data to maximize efficiency. Otherwise, simply figuring out how to get the critical information you need could end up consuming the bulk of your remote working day.

So, how do you instantly find the information you need so you can seamlessly leverage it to work more efficiently from your home office? First, for remote data, your main office can set up an enterprise or developer search engine running "on premises" or in a cloud environment like Azure or AWS. This way, remote workers can concurrently instantly search terabytes of "Office" files, PDFs, popular compression formats, emails with multilevel nested attachments, databases and Internet or Intranet data in one or more centralized repositories, displaying retrieved results with highlighted hits.

But this still leaves the issue of sifting through your own data on your own PC. A search engine can also enable remote workers to instantly search their own local files and emails. The key search engine mechanism that, whether remotely or locally, enables instant searching of terabytes is called indexing.

Indexing is easy. All you need to do is point the software at the folders and other data you want to index. Your search engine can then go off on its own to build the index, even figuring out for itself what type of files you have as it does so. The search engine can also sift right through your multilevel data, such as an email with a ZIP attachment that contains an MS Word file with a secondary MS Access database embedded right inside. After indexing, you can instantly "slice and dice" all the indexed data using over 25 different search options.

Now you've taken care of searching the files and emails already on your PC. But you may still have a halo of actual paper surrounding your desk. Ideally, you could see the pages just as they appear on the corner of your desk, while being able to instantly search the contents like you would any other "Office" documents.

For this purpose, Adobe Acrobat, for example, offers "searchable image" PDF as a scanning option. "Searchable image" PDF lets you see the full picture of the original page while using optical character recognition or OCR to turn the text into searchable words and numbers. Even if there are handwritten marks on the page that the OCR process doesn't pick up, such marks will still be visible inside the PDF document, making "searchable image" PDF extremely useful for remote work.

Another search engine tip for remote work. You know how sometimes you run across a PDF file where you try to cut and paste from the file and you can't? This is because the file may superficially look like an ordinary PDF, but it is really an "image only" PDF. When your search engine builds an index, it can flag all of those "image only" PDFs, allowing you to run them through an OCR application like Adobe Acrobat and make sure that they are fully digitized too. After OCR, you can cut and paste and instantly search the text along with the rest of your files.

One more tip for searching text with small OCR or typographical errors – like everyone always makes in emails. A feature called fuzzy searching can sift through minor deviations of letters in a word. This way, if Connecticut is mis-OCR'd or mistyped as Connexitcut, you can still find it with a small level of fuzziness in a Connecticut search.

