



# The Problem with Keyword Searches: Not Everyone Can Spell

December 10, 2012

It has been over a year since the verdict in the trial of Casey Anthony, a mother who was accused of murdering her two-year-old daughter, Caylee. The acquittal was widely covered in the media due to strong emotional responses by many of the people following the case. Recently, new evidence was found after an Orlando based television station, WKMG, discovered that computer investigators who conducted the keyword searches on the computer used by Casey Anthony overlooked vital data that could have changed the outcome of the case.

In general, keyword searches are used during the discovery process to help find documents that are responsive to opposing counsel's discovery requests or are relevant to the case. For the past twenty years, keyword searches have been used by many litigants to reduce the number of documents that have to be manually searched. A keyword search is typically performed by indexing all the documents into a searchable database or using an existing search feature that is built into a program, such as Microsoft Outlook's "find" feature.

Once the files are indexed or organized, the reviewer enters a specific word/phrase in a search form and the program or database returns any documents that contain that word/phrase in its content. For more information regarding performing advanced Boolean searches, check our post titled, "Performing Boolean Searches in Adobe Reader Across Multiple PDF Files".

Keyword searches are used frequently in the discovery process for multiple reasons including: the simplistic nature of performing the search, the low cost associated with the feature since it is built into many of today's e-discovery platforms, and because courts regularly accept keyword search procedures in their discovery protocols. However, as technology becomes more advanced and the amount of data that is produced increases, litigants and courts are beginning to realize that keyword searches may not be the most effective way of filtering the data.

More specifically, keyword searches have an essential flaw: when a litigant uses keyword searches valuable evidence may be missed due to the ergonomic concept of human error, such as spelling. That is exactly what investigators discovered in the Casey Anthony matter. Investigators and prosecutors spent nearly three years collecting, analyzing, and preparing for the Casey Anthony Trial. That preparation included searching a computer hard-drive, which was often used by Casey Anthony.

Throughout the case, prosecutors argued that Caylee was poisoned with chloroform and then suffocated using duct tape. Unfortunately, her body was found about six months after she disappeared and was too decomposed for her cause of death to be determined. When trying to convict Casey Anthony, prosecutors produced only a few

vague entries created by the computer's Internet Explorer ("IE") browser, which included a search for information about making chlorophyll, but Casey Anthony's mother testified that she was the one who conducted that search.

The initial search conducted by investigators that brought to light the IE browser's files, however, overlooked over 1,200 entries that were created by the Firefox browser, another browser that was also located on the computer's hard-drive. Those overlooked entries included a Google search for the term "fool-proof suffication," a misspelling of "suffocation". Whomever conducted the misspelled search also clicked on an article about suicide, which discussed taking poison and putting a bag over a person's head. Additionally, the Firefox browser recorded activity on the social networking site MySpace, which was commonly used by Casey Anthony.

As you can see, keyword searches have their limitation; in particular, the fact that human error may contribute directly to the data produced. When conducting a keyword search it is imperative to use wild cards (e.g. "!" or "\*" at the end of a word) and to include common misspellings of words, as mistakes are often made and code names or slang are sometimes used for cover-up purposes. Without doing so, a litigant risks missing indispensable information that may contribute to the outcome of their case.

If you or your company is facing litigation that includes producing electronically stored information and are running to problems with conducting effective keyword searches, contact James G. Ryan at [jryan@cullenanddykman.com](mailto:jryan@cullenanddykman.com) or via his direct line at 516-357-3750.

*A special thanks to Sean R. Gajewski, a law clerk at Cullen and Dykman, for help with this post.*