

EXPERT ADVISORY

SERIES

Advanced Discovery Quick Guide

DIY Searching – A Cheap Solution or an Expensive Mistake?

2018



DIY Searching – A Cheap Solution or an Expensive Mistake?

When involved in litigation and faced with a need for disclosure, it might appear that an in-house IT team are in the best position to search any identified or preserved data sets to cull the volume of data that may need to be supplied to an eDiscovery vendor for processing and review.

While this could be appropriate in certain specific circumstances, there are a number of significant potential issues that should be considered before proceeding on such a course. Getting things wrong at this early stage in the disclosure process could prove to be an expensive mistake.

REQUIREMENT FOR RE-COLLECTION

Even if the methodology followed is not criticized by the other side, should any additional keywords be required, or additional date ranges be needed, then entirely new searches will likely need to be performed.

This could be a long process to re-search all the areas that were originally searched. This could also result in a considerable amount of duplication if it is not possible to identify and exclude data which was already found to be responsive.

In addition, if data has been lost (for example automatically deleted by normal policy) then there could be allegations that the data was not properly preserved – potentially a considerable issue that could be exploited by the opposing party.

DUPLICATION

Depending on how the data is searched and exported, it is possible to end up with a significantly greater data volume to review. For example, if searching within Outlook, this may require the creation of a folder of emails for each search term that was run to save the responsive emails. This means that if 100 search terms are run (not unusual in a litigation case), then 100 folders would need to be created, with a potentially large amount of duplication as one email may be responsive to a large number of keywords, and therefore would appear in each of the respective folders.

LACK OF ADVANCED SEARCH FEATURES

Many programs which are used to perform searches have a limited search functionality. As an example of this, they may not be able to deal with Boolean logic (such as searching for “Apple” AND “Pear”, or alternatively “Banana” within five words of “Coconut”).

In addition, one of the biggest restrictions is the inability to run a list of search terms in a single search. This means that each term may have to be run separately, which, as above, can result in a large amount of duplication of data and could mean that the searches will take a considerable amount of time.

EXPORTING

Another of the benefits of using specialist software designed for use in electronic disclosure cases is that data can be exported in such a manner that all metadata is maintained. Our experience is that with other search methodologies, this is not typically the case. This means that the data received from the export for processing would likely have had the metadata altered from the original source data. This can then, in turn, raise questions over the process that was used and its acceptability for use in legal disclosure, as well as potentially opening up to allegations that the data has been manipulated.

As an additional point, when considering the export of data, if the plan is to search and then export data, for example to an external USB device, then any export or further copy of the data should maintain the original folder structures and should also ensure that the original source of the data can be determined from the exported data.

CHANGES TO INTERNAL METADATA

Some search methodologies may use the underlying programs to open the file in order to extract the text. Because this process actually opens the file, this can result in the internal metadata of that file being altered. This can lead to questions over the defensibility of the methodologies used, and potentially allegations of data manipulation.

MATERIAL NOT SEARCHED

Many search methods that are used by in-house IT, for example Outlook or Windows search will miss various types of data and may not report this (known as silent failures). As this is not reported, it may not be clear that such an error has occurred, and hence there will not be an opportunity to address the issue. This could be of particular concern because the other side may have the documents that were not searched and could demonstrate that they should have been included in disclosure and could therefore suggest that the disclosure provided is deficient.

Examples of items where searches failing silently could occur, resulting in the items not being correctly searched include;

- Email attachments
- Embedded files – e.g. an Excel file embedded within a Word document
- Files within container files such as ZIP or RAR files
- Emails stored within a PST archive (or other similar email container files)

- Scanned image files which have not been subject to OCR
- Directories or files to which the program being used does not have relevant permissions
- Untypically large files
- “Unusual” File Types (e.g. OpenOffice etc.)
- Files which contain Macros or other automation features (such as those included on many spreadsheets)
- Spreadsheets with text within “hidden” worksheets
- Text within comments or notes on documents
- Text within encrypted files

Failing silently can also be an issue for files which are considered to be “corrupt”. This could include, for example, PST email files which may easily be fixed. This could result in entire containers of email not being searched when a simple operation would result in the emails being searchable.

AUDIT TRAIL

One of the benefits of using specialized software for searching is not just the advanced search engines that can be utilized, but that searches are fully logged and an audit trail maintained. With many programs that we have seen utilized to perform searches, this same level of recording and audit trail evidence is not available. Should the other side query the exact searches that were run, having such a log can be invaluable.

EXPERTISE

If software with the capability to perform as required in a legal disclosure scenario is available in-house, it should still be ensured that the staff members performing the work have the necessary experience to do so. If not, then again this could be perceived as a weakness in the case and may be an area that the other side may look to exploit and question if the work was performed correctly.

INDEPENDENCE

If using an in-house resource, there could be a question over the independence of the individuals performing the searches. This may be of particular concern if, for instance, a member of the IT team may be involved in the matter. Even where this is not the case, members of the IT team may have a vested interest in the matter and their judgement could be called into question.

IN SUMMARY

The above gives an outline of the potential issues and drawbacks when using an in-house IT team to assist with disclosure for a litigation matter. While this may be appropriate in some cases, the above give an indication of the types of arguments that could potentially be raised by an opposing side. Because of the serious downstream issues that can result from this approach, serious consideration should be taken before deciding if you should utilize an in-house team or instruct a third-party expert in disclosure to support you.

ABOUT THE AUTHOR: DR TRISTAN JENKINSON

Dr Tristan Jenkinson has over ten years of experience working in the digital forensics and electronic disclosure field. He has provided expert evidence to the high court, and has been appointed directly, as well as being appointed as a single joint expert by the court. Tristan has advised and assisted corporate clients, law firms and government bodies on an extensive number of cases covering the UK, Europe, South America, Africa, and Asia. Previously, he led the UK based Digital Forensics and Incident Response team for a global consultancy.

Tristan has worked on different types of investigations across both the digital forensic and electronic disclosure disciplines, including investigations regarding fraud and asset tracing, allegations of data falsification and manipulation in relation to both files and emails, failures to meet disclosure obligations, intellectual property theft, contractual disputes, blackmail, rogue trading, regulatory investigations and breaches of competition law.

He has worked as the lead investigator on forensic investigations and as well as providing expert evidence to court his work has involved formal expert meetings and in court attendance at the request of lawyers and barristers to assist with cross-examination. His analysis of other expert reports focuses on highlighting report inconsistencies, omissions and forensically inaccurate statements. Clients have also sought his advice on wording of instructions for Single Joint Experts.

ABOUT ADVANCED DISCOVERY

Advanced Discovery is an award-winning and leading global eDiscovery and risk management company, partnering with law firms and corporations since 2002. Advanced Discovery and its global family of companies, Altep, Millnet Document Services and Legal Placements Inc., offer a complete suite of solutions to clients around the globe, including eDiscovery, information governance, Riskcovery® and compliance risk assessment, cybersecurity, legal operations consulting, data forensics and investigations, paper discovery and digital printing, as well as legal recruiting and placement. The company employs leading professionals in the industry, applies defensible workflows and provides patented and industry-proven technology across all phases of the eDiscovery and risk management life cycle. This devotion to excellence has earned Advanced Discovery inclusion in the Inc. 5000 Hall of Fame for being named to the fastest-growing companies list for six successive years and recognition as the highest-level Best in Service Relativity solutions partner for six consecutive years.

CONNECT WITH ADVANCED DISCOVERY

IN THE UK: +44 (0) 20 7422 8850

IN IRELAND: +353 (01) 532 9114

IN THE US: +1 866 342 3282

[Website](#)

[Contact Us](#)

[Blog](#)



v.041818.835A.DB