

New York Law Journal



E-Discovery Exceptions

When files get lost in the shuffle.



BY STEPHANIE GIAMMARCO
AND JENNIFER AIRA-VENTRELLA

Keyword searches performed on large volumes of electronically stored information (ESI or data) may yield ESI for review that includes password-protected files. Counsel may interpret the presence of these password-protected files to mean that all such files have moved through the electronic discovery process resulting in the complete identification of relevant files for subsequent review (i.e., any culled out password-protected files are not relevant). However, the actual reason for the inclusion of these password-protected files may have been a fortuitous accident. The file may have been attached to an e-mail that turned up in the search, for example. Thus, there may be other relevant password-protected files that were not selected because the file contents were not readily searchable and they were not attached to an e-mail. As a result, counsel may not review all relevant ESI and may make incomplete representations regarding the results of the search.

Password-protected files are one example of the many different types of files commonly called “exceptions”—those files that do not move through the electronic discovery process as one might expect. Understanding different types of excep-

Stephanie Giammarco is the national director of computer forensics and e-discovery for BDO Consulting, a division of BDO Seidman. Resident in New York, she holds certifications in public accounting, information technology and fraud investigations. **Jennifer Aira-Ventrella** is manager, e-discovery, at BDO Consulting, in the Los Angeles office.

tions and where e-discovery exceptions occur is important especially given the current legal landscape. Counsel’s knowledge of the options for handling exceptions may affect the ESI universe, case strategy and the accuracy of statements made regarding discovery.

Types of Exceptions

The most obvious areas where ESI exceptions can be identified are during the processing and post-processing stages. Processing is the stage where the ESI is prepared for loading into an e-discovery review platform, which includes de-duplicating—removing or marking duplicate records within the dataset—data filtering and keyword searching. Post-processing is the stage after the data has been processed and the results have been uploaded into the e-discovery review platform for review.

Exceptions in the processing phase include files that are password-protected, corrupted and encrypted. These documents have a common thread—they cannot be accessed by a person or process in the conventional manner (i.e., clicking on a document and viewing it). Password-protected files require a password, corrupted files require the user to “fix” the document, and encrypted files require a key. These same requirements exist for e-discovery review platforms (“e-discovery tool” or “platform”).

In most cases, in order to view, search or analyze these documents, the passwords must be obtained for password-protected documents, the corrupted documents must first be fixed and the encrypted documents need to be un-encrypted. Once these files are addressed, the data can be processed and loaded into the platform of choice. If these requirements are not met, then the ESI contained in these files is not being fully considered.¹

Other processing exceptions relate to file exten-

Thursday, June 19, 2008

sions, frequently referred to as unsupported file types. Although most of the available e-discovery tools are robust and accept hundreds of file extensions, there are certain file extensions that are not supported by the tools. This exception includes most database files, such as .mdb (Microsoft Access files) but can also include specialized programs such as .cad (AutoCAD files) and alternative e-mail files, such as .nsf (Lotus Notes). Therefore, if an ESI collection yielded a file with an extension not supported by the platform selected, that file will not be loaded to the platform for review.

A related exception occurs due to the ability for users to alter file extensions, often referred to as a bad signature file extension. For example, if a user changed the extension of a Microsoft Excel file from .xls to .xxx, and .xxx is not supported by the platform selected, this file would be considered an exception and not loaded to the platform. As a result, no determination could be made as to whether it falls within the parameters set for the case.

As discussed above, some files such as password-protected files, which are considered processing exceptions, may still load into the e-discovery tool; users simply cannot search or view them. These exceptions need to be addressed in the post-processing phase as described below. There are different reasons why this may occur—one common explanation is the document is part of an e-mail parent/child attachment relationship, also called a message unit, where at least one component fell within the parameters set for the case, such as a keyword hit result. Thus, the entire message unit is loaded to the platform, including the password-protected file. As a result, some e-discovery tools have a separate folder or area where these files are automatically populated and exceptions may sit there, never reviewed or considered as part of the case or investigation.

The exceptions that are recognized in the post-processing phase are not as easily identified, as the files may be searchable and/or viewable, but they still may not have moved through the e-discovery process as one would expect. E-mails with embedded objects are one example of an exception identified post-processing. Some e-discovery tools cannot support the extraction of an embedded object in its processing phase and most typically, human

reviewers identify these exceptions. As an example, an e-mail may appear as:

To: Smith, Jack
 From: Jones, Sally
 Date: May 12, 2004
 Re: Monthly Revenue Schedule
 Jack—

Here's the chart with the monthly revenue numbers.

As you can see, we went from \$6m to \$5m this month. Can we discuss this decrease in Friday's meeting?

Thanks,
 Sally

A human reviewer would note that there is a component to this e-mail—the embedded chart that Sally is referencing—that is missing.

Exception types that surface in the pre-processing phase may also appear in the post-processing phase. For example, as noted above, e-mails with password-protected attachments may be loaded into the platform even though the password-protected document remains unsearchable. Post-processing exceptions suffer the same fate as the processing exceptions—these documents will not be identified when conducting keyword searches on the platform's user-interface or when implementing content-based searching methodologies. Therefore, most of these documents, which may potentially contain pertinent information to a case, are often never considered or reviewed.

The Legal Landscape

In matters involving many documents available for review, exceptions add to the complexity of the electronic discovery process. However, for a variety of reasons, not the least of which is the thoroughness of the e-discovery process, these exceptions may need to be addressed. Password-protected documents are by definition of higher security concern to the user. For that reason, depending on the nature of the case or investigation, files with password protection may be of particular interest. Similarly, if an individual purposely changes a file extension to avoid discovery, this file may also be of particular interest. Without addressing exceptions, these files of interest may not be considered in the discovery process—and yet, they may contain the case's smoking gun.

In litigation, the Rule 26(f) conference may be an appropriate venue to discuss how the parties may handle exceptions in the e-discovery process.² Many Rule 26(f) conferences already address database files, as it is known that most of the e-discovery tools do not support these files; however, there are other exceptions noted herein that can also be discussed.

One may argue that certain exception categories may not be reasonably accessible, as defined by Federal Rule 26(b)(2). Although Federal Rule 26(b)(2)

addresses production of ESI, not the preservation of such ESI or the obligation to discuss the preservation of this ESI, some recommend that this aspect should be discussed to make the meet-and-confer as productive as possible and limit future disputes or arguments.

Handling of Exceptions

If deemed necessary, many if not all of the above mentioned types of exceptions can be addressed. For processing exceptions, such as password-protected files and encrypted files, these files can be uploaded into a password cracking or recovery tool. The cracking or recovery tools simply run “in the background,” until the password or encryption is recovered. These tools can handle many files at once and require little human interaction.

The downside is that the tools may not be effective in ultimately cracking the files. Therefore, one may choose to employ a reasonableness standard to the time allowed for this process and if the password has not been recovered in that time, to discontinue the process. Similarly with corrupted files, one can attempt to fix these files, but there is no certainty that the files can be fixed. However, with corrupted files, the time and associated cost to fix these files could be substantial if each file has to be fixed individually. If it is deemed necessary to attempt to fix corrupted files, the time allotted to this process can be discussed prior to commencement of such an exercise.

As for file extension exceptions, those files with extensions not supported by the e-discovery tool may be segregated. A determination can be made if those files are to be reviewed in their native format, or alternatively, in a viewer, or simply not reviewed at all. As for the bad signature files, there are forensic tools that can compare the file extension to the file header/signature and identify those that do not match, thus indicating an inconsistency between the file type and the file extension. If warranted, these files can be “corrected” and reviewed.

For the post-processing exceptions, the concept of exceptions and how to identify them can be discussed and defined for the human reviewers prior to the commencement of the review. A workflow strategy can be developed and communicated to allow the reviewers to identify such exceptions (e.g., tagging) for subsequent follow-up, if required.

Once the agreed-upon exceptions are addressed, most e-discovery tools require these files be reinserted at the beginning of the e-discovery process for reprocessing and uploading for review, as necessary. If the e-discovery platform in place charges by volume, costs associated with re-processing these files, if any, should be discussed with the vendor.

In most cases, all exceptions are not addressed, as the time and money required to address all exceptions may not be reasonable. In addition, even if

the time and money are spent, some files simply remain unviewable or unsearchable. Therefore, it is often helpful to prioritize exceptions (e.g., by type or category, such as password-protected files, or by custodian). The determination of which exception files to address further needs to be made by counsel with input from the e-discovery team and/or service providers.

Since all exceptions may not be processed, the tracking and handling of exceptions are needed to document the lifecycle of ESI thoroughly. A data-tracking mechanism that records the processing steps and resulting volumes can be implemented at the commencement of a case. For example, if 100 GBs of data were collected, with 40 GBs of data being culled out during de-duplication and 10 GBs of data being culled out during data filtering, there should be a resulting 50 GBs of data loaded to the platform. If only 40 GBs of data are loaded to the platform, 10 GBs of data have not yet been accounted for and could be exception files. Additionally, a secondary mechanism can track how the exceptions are handled (e.g., successfully cracked and reloaded for review or unable to be cracked/no further work required).

Electronic discovery plays an integral part in most complex litigation or investigations. Thoughtful consideration of topics such as exceptions and their impact on the discovery process is important to serve clients properly and to minimize exposure to sanctions for law firms and their clients. Counsel can discuss the handling of exceptions with the e-discovery team and/or service providers. Similarly, in conversations or proceedings with adversaries, such as the Rule 26(f) conference, an understanding of how exceptions are to be handled may result in a more efficient and effective discovery process. Counsel who take a comprehensive approach to e-discovery—including exceptions—can make more fully informed decisions regarding available ESI and case strategy.

.....●●.....

1. Other types of ESI may not be searchable due to their nature (e.g., scanned images, faxes, basic non-text searchable PDFs, and foreign language documents for platforms that do not have foreign language capabilities.) Although this ESI may or may not be considered an “exception,” this limitation also affects subsequent review.

2. Counsel may be concerned that addressing complex questions and issues at the Rule 26(f) conference may not be appropriate. In a recent article by Moze Cowper and John Rosenthal, participants in the Sedona Conference—Working Group One on Electronic Document Retention and Production, “A Practitioner's Guide to Rule 26(f) Meet-and-Confer: A Year After the Amendments”—they state “...Our advice is that it is better, in most instances, to get the [difficult or complex] issue out on the table rather than ignoring the issue.” Page 7.