Native File Review: A Primer

LexisNexis®

Introduction

There is a lot of talk in the electronic discovery industry about "native file review," "native file format," and "native file production." Many electronic discovery providers are scrambling to assure their clients that they support native file capabilities. The claim is that native file review is cheaper and faster, and that it is required by courts, government agencies and by the Federal Rules of Civil Procedure. But is this really true? Is it really less expensive? Is it more efficient? And more importantly, is native file production required under the law?

This primer will help attorneys and legal professionals find the answers to these questions by providing an understanding of the issues and considerations involved with using both true native file review and a provider to process data for document review.

Native File Review

In order understand native file review, let's first look at "true" native file review. True native file review does not require any kind of service—it is as simple as a client copying its files onto a disk and handing it over to counsel for review. Counsel then reviews the documents on its computers using the software applications used to create the documents. Native file production is made by providing a copy of the original data to the adversary. And in smaller cases with relatively few documents, it can be as simple as that.

But in substantial litigation and government investigations that involve thousands, if not millions of documents that exist over several servers, in different formats, and different locations, true native file review is impracticable. The services of an e-discovery provider are often vital for complex discovery where retrieval and review of information is much more complicated than copying electronic files to a disk or providing data to an adversary. So, although many e-discovery providers are claiming they provide native file review—because it appears to be inherently less expensive and time consuming than processing information into another usable format such as PDF—what they are touting is not true native file review at all, but rather, exactly like converted formats, the ability to review the documents as they appear natively and to then have a link, if necessary to the native file. Although this may appear to be a distinction without meaning, the difference is quite important because true native file review has many disadvantages that any good litigator will want to avoid.

The Disadvantages of True Native File Review

The most expensive part of any document production is the time it takes for attorneys to review the data and documents once they have been retrieved. Therefore, while true native file review may cost less and take less time in the beginning of a production (because the electronic files are simply copied to a disk without the need for third party assistance), these costs are offset by the lack of functionality of true native file review.

Limited Review Functionality

- In true native file review, there is no index across all of the documents because they are not stored in a central repository. Once the review is underway, a true native file review cannot capture the progress of the project across reviewers. Reviewing documents in different programs also precludes Bates numbering or branding the documents sequentially and tagging "hot documents" for others to see. True native file review also limits the amount of metadata that is available to the viewer; for example, a "bcc" field in an email can only be viewed as part of the sender's email. Similarly, emails cannot be viewed in context and any attachments to emails must be viewed in the program

in which they were created. Once the review is complete, privileged or redacted documents must be removed from the original CD, because true native file review cannot support redaction; nor can a privilege log be automatically created as the documents are reviewed.

Software Licensing/Purchasing Required for Review

- Counsel must also buy, license, and load all of the versions of all of the programs that support the true native file formats. Large document reviews can often involve fifty or more contract attorneys; each with a computer, each needing to access these programs. Review of this kind is further complicated by the fact that some documents will have been created using older programs that no longer exist or new versions of commonly upgraded software. Information is often lost when documents created in older programs are viewed in upgraded versions. Moreover, in order for the reviewer to access the metadata, each reviewer must be trained on the specifics of each version of each program.

Spoliation Risks

- The biggest risk of true native file review, however, is that when a document is reviewed in its native format, the metadata is automatically changed. While a copy of the document can be made before review, the copy itself changes the metadata (e.g. making a copy without taking necessary precautions can change the "date modified" field in some common programs). If a copy is not made, reviewers looking at original data run a real risk of spoliation. Any party using true native file review cannot certify to a court that the documents are in their true, original form which can leave parties and their attorneys open to tremendous liability.

In short, true native file review more often than not costs far more money in attorney review time and technical difficulties that what it saves on initial processing. This cost, in combination with the risk associated with spoliation, usually makes true native file review less than a desirable option in most cases.

True Native File Review vs. Converted Images

So, if true native file review is undesirable, why are some in the industry encouraging it? And if the native file review touted by some vendors is not true native file review, what is it?

What many vendors are calling "native file review" still requires the processing of data, no matter how the documents look or in what format they are reviewed. It is necessary to process data in order to create even the most primitive of discovery tools: an index. Additionally, critical metadata must be extracted and written to a database structure in order to take advantage of many of the basic functions of e-discovery review tools: the ability to search fields such as "date created" or to identify the author or custodian of a document.

Once the data is processed, some e-discovery providers then convert the data into a usable format such as PDF, while others maintain it as a text extract of the files, convert to TIFF (tagged image file format), or offer some sort of pseudo-native file format such as RTF (rich text format). The initial processing makes up about 80% of the overall cost of any e-discovery vendor service. By contrast, the cost of conversion is minimal but provides many benefits. Some e-discovery vendors argue that converting files to PDF drastically increases the costs of e-discovery

because all the files are converted up front. They compare this with other methods—which leave files in native format and convert to PDF or TIFF only at the time of production—but their arguments are misplaced. The confusion likely stems from a misunderstanding of the basics of e-discovery technical processes and confusion about the various pricing options that are available to meet today's e-discovery needs.

PDF is one of the most efficient ways of sharing and reviewing documents. PDFs can be reviewed or printed from any computer regardless of the type of software that was used to create the document. PDFs can be shared with anyone as long as they have Adobe® Acrobat®, which is available as a free download. Because PDFs have automatic file compression, they are one-tenth the size of a TIFF image, so they contain more information and are faster to send and download.

PDFs look exactly like the original documents as created in their native programs, and preserve all of the graphics, formatting, fonts and color. And although a PDF review format is far superior to reviewing native files because it provides secure access to the files and enables the documents to be stored, searched and categorized in one format in a single depository, most e-discovery applications allow files to be reviewed natively when it makes sense. This is especially helpful for reviewing the formulas in Microsoft® Excel® spreadsheets (the most troublesome file type to review in a converted format).

When considering the fact that up-front processing for search capabilities and metadata extraction make up the majority of the overall processing time and cost in e-discovery and the actual cost of conversion to a standard file format like PDF is minimal, choosing a provider with high volume processing capabilities and other important document review features can often provide the best of both worlds.

Native File Production

But what about the courts, government agencies and the federal rules? Aren't they all requiring native file review? The answer is no. While it may be true that certain attorneys will request native file production, adversaries have no say in how a party reviews its own material. *Thus, no one can require native file review.*

As for production, not one agency has promulgated any formal rule requiring native file production for litigation or government inquiry. In fact, the FTC discourages production in native and recommends that electronic documents be produced in searchable formats - including PDFs – as a best practice for merger investigation procedures. *See* http://www.ftc.gov/os/2002/12/bcguidelines021211.htm. The SEC has also indicated that it prefers production of electronic documents in formats compatible with LexisNexis® Concordance™ or Summation®, which require up-front processing and are therefore no longer in their true native formats.

Likewise, the Federal Rules of Civil Procedure do not mandate native file production. The rules require that parties to civil litigation try to reach an agreement on how documents will be produced. *See* Rule 26(f)(1), found at www.uscourts.gov. If an agreement cannot be reached, the documents must be produced either in the form in which they are ordinarily maintained or in a form that is reasonably usable. *See* Rule 34(b). Although an argument can be made that "form in which is it ordinarily maintained" means "native format," the comments to the amendments, published on July 25, 2005, explicitly note the disadvantages of native file production, such as "the inability to redact, leading to privilege problems, an inability to bates-stamp the 'documents' for purposes of litigation management and control, which is not an insignificant consideration, particularly in complex multi-party litigation." *See July 25, 2005 Comments* to then Proposed Rule 34. Even if the courts ultimately decide that this language is synonymous with native file production, a party will also have the option to produce

in a reasonably usable format, which would most likely include such widely used applications as PDF. In fact, the federal courts initiative known as the Case Management/Electronic Case Files project (CM/ECF) encourages litigants to file case documents electronically, and the only file format allowed is PDF. The CM/ECF project was initiated in 2001 and is being implemented in federal district and appellate courts nationwide. See www.uscourts.gov/cmecf.

In the majority of cases, PDF and TIFF productions are more than adequate to serve the needs of the parties and the courts. *See e.g., In re Priceline.com Inc. Securities Litigation*, 233 F.R.D. 88 (D. Conn. 2005)(court ruled that TIFF and PDF format are the "most secure format" for the production of documents in the case). Although there are particular instances where a party may need to produce specific documents in their native format, as in *Williams v. Sprint/United Management Co.*, 2005 U.S. Dist. LEXIS 21966 (D. Kan. 2005), the court's ruling was based on the fact that Excel® speadsheets can only display formulas in their native format. Thus, plaintiffs had demonstrated the relevance and necessity of the formulas which were part of the file's metadata. Absent a showing of relevancy, a party should avoid the wholesale production of metadata. *See Kentucky Speedway v. Nat'l Ass'n of Stock Car Auto Racing, Inc.*, 2006 U.S. Dist. LEXIS 92028 (E.D. Ken. Dec. 18, 2006)(court rejected Williams' rationale and ruled that metadata should not be produced absent a showing of relevancy) and *Wyeth v. Impax Labs., Inc.*, 2006 U.S. Dist. LEXIS 79761 (D. Del. Oct. 26, 2006) (refusing to order production in native, and applying District of Delaware's "Default Standard for Discovery of Electronic Documents" which requires showing of particularized need to warrant metadata production). Counsel should meet and confer early in the litigation to determine what format best serves the needs of the case.

Even though there is no evidence that courts or government agencies are requiring native file production with any regularity, any top tier vendor can produce data in its native format upon request, even if it has been first converted to PDF for ease and efficiency during the review process. Because conversion to PDF does not cost more than any other processed data, and carries with it tremendous advantages, best practices would dictate that counsel first review the documents in the most efficient and effective manner, and then handle documents in native file format only if required at the time of production.

Comparing the Options

The true native file review discussed above does not typically involve a vendor of any kind. Instead, attorneys and their clients simply copy data and open it in desktop applications for review. When an e-discovery vendor purports to offer native file review, there is other work to be done to make the files searchable and viewable in a commercial e-discovery application. It may be helpful to review a side-by-side comparison of the various stages of the case when considering whether native file review should be an option.

NATIVE FILE REVIEW	PDF FILE REVIEW	PREFERRED METHOD?
Up-Front Costs		
Per-unit charges from vendor will include initial document processing to capture the text and metadata so documents can be searched. Without this processing, there is no way to search across different file types or even between email messages and their attachments.	Per-unit charges include processing fees similar to native review processing. The bulk of the expense for either method is associated with indexing text and/or numerical data for searchability and sortability.	*Tied:* Both methods require up-front processing in order to make the data usable in discovery review.
Document Review		
Commercially-avaiable application viewers for native for native file formats enable the review team to see documents as they appeared in native file types. The ability to search between different file types or between emails and their attachments can be greatly limited, depending on the viewer selected. Native viewers do not enable the review team to share collaborative workspace for making notes and annotations, etc. With this method, there are no administrative rights for dividing documents into different groups for review, allowing only certain reviewers to redact, etc.	E-discovery applications enable viewers to utilize PDF as the common file format so functionality like Bates numbering, redacting, searching, sorting, accessing metadata, reviewing "parent-child" email messages in order and in context of a conversation, creating an automated privilege log, etc., can be employed from the beginning of the project. Documents are all stored in a common repository so the review team can share and access information as the project progresses.	*PDF:* The functionality and workflow benefits of PDF review far outweigh any perceived convenience of "staying in native" for the review process. The actual review process is at the heart of the case in e-discovery, so this is often the determining factor in making a decision about review format.
Document Production		
A subset of the documents reviewed in native format is typically converted to either TIFF or PDF for hosting in a litigation support database like Concordance or Summation, and for production to opposing counsel. Vendors charge to convert the subset of selected documents to the client's chosen file format. In only rare situations are native files actually produced to requesting party.	The documents designated for production are already in PDF, and can be handled in any number of ways. The production set can be stored in the online repository, burned to a CD, or transferred to a desktop litigation support database. The output format can be PDF, TIFF or native files, as selected by the client according to the needs of the case. Charges for output of the documents are only a fraction of the up-front processing costs.	*Tied:* For either review method, there may be some fee for creating a production set. The myth is that one would be somehow " paying twice" for the PDF option, when the reality is that the up-front charges and production charges are comparable with either method.

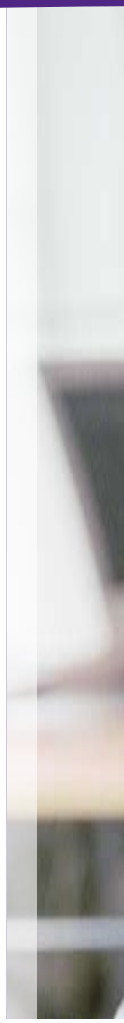
Conclusion

So why has native file review and production become such a hot issue? What it really comes down to is price. All vendors have to process data, which costs about the same from vendor to vendor. Differences in price in the industry, therefore, lie not with what type of format is chosen for document review, but in the services a vendor provides above and beyond just the conversion of data. Factors that go into pricing include: turn around time for processing, client services, 24 hour support, training programs, and supplemental services such as document retention and litigation preparedness consulting. Cost differentials can also be related to product features, such as real time redaction, creation of a privilege log at the click of a mouse, scrolling documents in real time, and ease of use: (clients say that PDFs simply look better and are easier to read than other formats which increases reviewer productivity).

All of these factors end up as an ultimate savings for the client because reviewers are able to review the documents much more efficiently. Some vendors provide à la carte pricing, where each line item appears reasonable, but the sum of the charges can be overwhelming. Other service providers bundle the majority of services they offer, so the "sticker price" may look higher, but the actual cost of the project is likely to be the same or even lower.

The best investment, however, is in an e-discovery service provider that has experience. With the current wave of spoliation cases resulting in huge jury verdicts, hiring a proven winner is an important safeguard in the discovery process.

The notion of native file review as a superior legal practice is nothing more than a red herring. The real issue behind these arguments lies in the cost structure of how an e-discovery project is managed. It is a myth that converting documents to a user-friendly, reliable file type like PDF necessarily has to be more expensive than reviewing in native format. In the end, the true costs must be compared side by side and the true value of any e-discovery provider will not be related to the type of formatting it offers, but rather to the usability and functionality of the data once it is in that format. And to the peace of mind the client receives in the process.



The Discovery Experts: Industry Relations

LexisNexis Discovery Services has the right consulting and technology choice for every discovery need. Top law firms, corporations and government agencies rely on the LexisNexis® products and services, Applied Discovery®, Concordance™, and Hosted FYI™, to meet their discovery obligations on time, accurately and cost-effectively. Services include records management consulting, data collection, forensics, media restoration, data filtering, data processing, review and document production in the format that each matter requires. The Industry Relations team works to educate the legal community on the continually evolving case law and technology of electronic discovery.

For more information or to contact the experts, please visit **lexisnexis.com/discovery**.

LexisNexis®