

The Ten Minute Guide to Native Document Review Technologies for Litigators

By: John Fuex, Platform and Product Development Manager, IE Discovery, Inc.

It is tempting to leave the technical details to the propeller-heads who live in cubicles decorated with action figures, but do so at your own peril when choosing the tools to drive the most costly element of the discovery management process. Planning a successful document review requires informed decisions on overlapping strategic, logistical and technical issues to optimize the process for cost, speed, accuracy and consistency. This technology primer will guide you through the most common approaches for native document reviews as part of a comprehensive discovery management effort.

What is so complicated about reviewing electronic documents?

In the early era of litigation support automation, a primary concern was getting hard copy documents and microform into the computer. The hassle of scanning documents and then running OCR software to make them searchable had many legal technologists naively anticipating the day when document collection was done from a hard drive instead of a file cabinet. Finally the big problem of getting discovery material into a searchable computerized format would not only be solved, but eliminated!

Unfortunately, the transition was not the expected panacea for the complexities of automated discovery management. In fact, native files created more challenges than it dismissed including:

No Standard Format: While scanning paper and microform was cumbersome, it did have the benefit of converging on a few standard file formats (TIFF/PDF) that could reliably be displayed by commonly available tools exactly as they looked on the hard-copy originals. Electronic documents, conversely, are a Pandora's box of chocolates. Native file collections consist of an arbitrary number of file formats, many of them proprietary. Building tools and workflows that wrangled this much variety is a daunting task.

Format Specific Issues: Tracked-changes, spreadsheet formulas, fast save data, and metadata are frequent topics among electronic discovery experts. All of these issues stem from a fundamental difference between hard-copy and electronic documents, the implied finality that is imbued on a document once the ink hits the paper. Granted, work-in process documents were always part of the discovery process, but electronic documents and hastily sent e-mails provide easier visibility into the formative aspects of a business and create a higher standard for due diligence that must be supported by the document review technology.

Quantity: The explosive rate that documents and other evidentiary artifacts can now be created, spam e-mail, and the increased ease in collecting electronic files have increased the size of document collections considerably. The need for efficient review technologies to sort through the mountain of data has become a necessity to keep litigation costs from spiraling out of control.

Why must litigators get involved in the technical details of the review tool?

When document collections consisted almost exclusively of images of hard-copy documents, it was safe for litigators to delegate most of the technical decisions to the technologists. However, the introduction of native files has intermingled the technical and procedural aspects of discovery management. Planning a document review that involves native files can now involve numerous decisions impacting cost, speed, and accuracy. Litigators without a basic understanding of review technologies may, without realizing it, be delegating strategic, logistical and substantive legal decisions to the computer guy.

Is a review tool even necessary?

It is not a given that every electronic discovery effort requires the implementation of a high end review tool. For non-trivial discovery efforts, these tools are indispensable, but don't let a vendor sell you a hammer to pound on a thumbtack when your thumb will suffice. If proper precautions are taken to protect files from accidental modification, a single reviewer with a spreadsheet can review a small collection in preparation for a native production. However, for more than a few thousand documents needing multiple reviewers or redactions, it is strongly advised to make use of a formal review tool.

The role of image conversion, or TIFFing, in modern document review technologies.

Reviewing and producing electronic documents using modern review tools is typically faster and more and cost effective than trying to force native documents through a legacy workflow by batch converting everything to image. However, it is still often necessary to include imaged versions of native files in a production set. In these situations it is advisable to push the conversion process as far back in the discovery management workflow to avoid needlessly converting documents that will ultimately be culled from the collection by de-duplication technology or relevance and privilege review. For document collections that contain a large amount of chaff, the ability of the tools to accommodate this modified workflow is critical for containing cost and meeting production deadlines.

Common approaches for displaying native documents in review tools.

Several technologies emerged during the evolution of discovery review tools to address the tactical issues around reviewing the menagerie of formats found in electronic discovery collections. The relative effectiveness of each varies based on the type and content of file being reviewed. Recognizing this, the most effective native review tools are built on a hybrid approach that allows the reviewer to choose the most appropriate display mechanism on-the-fly based on the characteristics of the individual document.

HTML Viewers

These viewers display a HTML version of the document that can be easily displayed by a web browser without additional software to be installed on the reviewer workstation. Although images and formatting are often lost in the translation, this approach has the benefit of requiring very

little advance preparation and licensing on individual reviewer workstations. For documents such as e-mail or word processing documents that don't rely on heavy formatting or graphics, the high performance of this technology can make HTML viewers a good first option.

Pros:

- Best performance.
- No client install on reviewer workstations.
- Does not require original application to be installed or licensed on reviewer workstations.
- Ability to highlight key relevance or privilege terms in the document display.
- Ability to cut-and-paste from document.
- Almost no risk of contaminating reviewer workstation with viruses or malware.
- Very easy to use, no additional training needed.

Cons:

- Poor fidelity to original document look and feel.
- Tends to mangle or lose formatting that can't be represented easily in HTML, images and diagrams may not appear that could contain useful information.

Universal File Viewers

As a practical matter, it is essentially impossible to create a viewer that could open and faithfully represent every file format, but these tools do a reasonably good job at reading and displaying the vast majority of files that are common in document collections. Though not quite as fast as the HTML approach, these viewers still perform very well. The primary benefit of this technology is that it is able to display most document types very faithfully without the licensing expense and complexity associated with using the original applications.

Pros:

- Very good fidelity to original document look and feel.
- Very good performance.
- Does not require original application to be installed or licensed on reviewer workstations.
- Almost no danger of contaminating client workstation with viruses or malware.
- Limited functionality of original application (column resizing in spreadsheets)
- Fairly easy to use, no additional training needed on viewer technology.

Cons:

- Requires installing a client on reviewer workstations (usually trivial).
- Requires installing client updates as support for new file types is added.

Native Application

It should come as no surprise that the viewer that can display the truest representation of a native file is the application that created it. Although it may require more extensive training, using the original application also provides the reviewer with the ability to peer further into the document and uncover hidden information such as spreadsheet formulas, previous versions of the document and revision markings. However, taking advantage of the additional functionality may require training on a number of different applications that may be prohibitive when trying to quickly build a review team and introduce time delays and additional cost. Also, the amount of time reviewers spend looking at the splash screens while the application loads for each document can really reduce the rate at which documents can be reviewed.

Pros:

- Best fidelity to original document look and feel.
- Enables expert reviewers to see hidden information such as previous versions, revision comments, underlying formulas, etc.

Cons:

- Poor performance.
- Requires reviewers to be trained on multiple applications that are sometimes complex.
- Original application may no longer be available or prohibitively expensive.
- Extensive preparation of reviewer workstations to install and upgrade configure software.
- May require more expensive reviewer workstations.
- High licensing costs for application software.
- Weak integration with review tool.
- Higher risk of macro virus contamination.

Recommendations and Best Practices

In the broadest sense, the key trade-off between these technologies is between review speed and how accurately the document can be represented by the technology. Clearly, both accuracy and speed are critical components of any document review effort and compromising on either is not ideal. This can be mitigated, however, by using a review tool that is flexible enough to allow the reviewer to use the faster viewing technology for the majority of documents where the HTML or Universal viewer are sufficient, but to fall back to the native application for a closer look at documents that require the additional functionality.

A successful document review requires informed decisions on overlapping strategic, logistical and technical issues, with an overall objective of optimizing cost, speed, accuracy and consistency. Understanding the benefits and limitations of the technology inside your review platform will allow you to match the proper tool to the job and will result in maximum accuracy even under the tightest deadlines.

About The Author

John Fuex is an industry veteran with over 13 years of writing and consulting experience in the field of discovery management. He currently serves as the Platform and Product Development Manager for the InfoDox.com at IE Discovery, Inc. (<http://www.ieDiscovery.com>). He can be contacted at jfuex@iediscovery.com.