



UNDERSTANDING THE IMPACT OF PDF'S HIDDEN CHALLENGES TO YOUR FIRM

*A comparison of MasterFile and its
OCR/PDF Evidence Cruncher to CaseMap
with Acrobat Professional*



MAY 2009

JUST THE FACTS
MASTERFILE'S WHITE PAPERS

CONTENTS

Understanding the impact of PDF to your office	3
MasterFile's OCR/PDF Evidence Cruncher and Acrobat Professional	5
MasterFile's PDF document repository compared to CaseMap's file folder links	6
Price comparison	8
CaseMap "Bates Numbers" and document naming conventions, CaseMap's PDF Loader and DocPreviewer	9
Our short take on electronic Sticky Notes	11

In their paper, "The Bell Curve & Documenting Indexing/Imaging", CaseSoft suggests that for 100,000 document cases, simply add Acrobat to CaseMap and you're good to go. In this article, we examine how PDF documents will impact your office and the challenges you'll face. We'll compare MasterFile and its OCR/PDF Evidence Cruncher to CaseMap with Acrobat Professional and see what the difference really means in a practical sense to you, the lawyer or litigator in a small or solo practice with many cases, several professional staff and a busy schedule.

UNDERSTANDING THE IMPACT OF PDF TO YOUR OFFICE

Now it may seem a little strange to start a comparison with this topic, but we need to be sure about just exactly what it is we expect these products to do.

Comparing cars is a relatively easy task because we already understand what to compare: size, safety, comfort, gas efficiency, and so forth. However, PDF technology has so many facets, we need to first make sure we know which attributes are relevant to lawyers, litigators or investigators, to understand what exactly needs to be compared. When software vendors state "supports PDF", or "allows PDFing", or "Bates stamps PDF", or "bulk loads PDF", or "converts e-mails to PDF", they are simply advising you that their product has PDF related technology. However, such statements are too vague and do not help to understand exactly how the product, or "just having" Acrobat, helps you or appreciate the very significant new problems they create for you. It would be like buying a tool box that just says "contains tools".

PDF files and related technology do not exist in a vacuum and rather than what bells and whistles or ad-hoc PDF utilities a particular software product has, how smoothly the software lets you deal with PDF related issues you'll face day-in day-out is **paramount**. This is especially critical to the lawyer, litigator or investigator with even just a dozen or two active files and several hundred PDF documents per case -- **that's several thousand PDF files which will be generated in very short order.**

A sampling of issues you'll be dealing with, therefore, include:

- **PDF conversion**

Managing PDF and other documents that need to be converted to PDF coming at you from myriad sources and in various sized batches. Some will be from your own work-product. Others will come as paper documents to be scanned or e-discovery from your client or opposing counsel on DVDS and "box-fulls". On top of this, expect several dozen paper documents everyday, including mail, expert reports, research, transcripts, etc., that will need to be OCR'd. And, of course, plenty of e-mail. All will need to be converted to PDF.

- **Coping with thousands of PDFs and other documents**
A healthy chunk of your PDFs will have several copies or versions. For example, the "virgin" PDF file, a redacted copy, perhaps a few copies with different Bates stamps on them for each time it was disclosed, etc. (In fact CaseSoft specifically recommends that you "make a back up copy of all PDFs in case you want to renumber"). Many PDFs will also have source document files that you also need to manage. So the total number of PDF files you'll be generating is, well, quite a lot, and multiplying like rabbits.
- **Working efficiently with PDF and other documents**
Your PDF documents will be used for many different tasks and working with PDF documents day in, day out mandates efficiency: Can you pinpoint specific documents instantly? Can you select documents based on issue, author, date, via full text search, etc. **and then process them.** That is, print or tag them with new issues, Bates stamp and/or add them to pending disclosure sets? Can you track disclosure history, reproduce disclosed document sets, and so forth? Then, there are other issues -- like locating the native source files, burning in watermarks, redactions and stamps so they can't be hacked and undone, etc.

Once you begin to use PDF documents for even some of your work they quickly become "mission critical" and the last thing you need during a busy work day, faced with critical deadlines, is fiddling and tinkering with inadequate or multiple software packages cobbled together with myriad manual operations to "work as one" just to get simple tasks done. What you really need are industrial strength, professional power tools you can rely on -- which won't waste your time, nor let you down at a critical moment, nor make your work tedious and error prone.

It's also critical to realize that it will not be just you who'll be working the PDFs; your staff will too, and therefore, software requirements and related issues change quite dramatically. It's one thing to use PDF just on your own desktop but it's quite another to reliably administer thousands of PDF documents, spread over a couple of dozen active case files along with the related source document files, among several users accessing them simultaneously and all performing different tasks with them.

How will you and your staff manage all these PDFs and other document files? Hopefully not with disk file folders, because if so, the integrity of the folders is entirely dependant on each user correctly remembering and performing manual procedures. It's been done and being done; however, when you have just a few staff, it can get tricky and become highly error prone, looking neither pretty nor professional. For example:

- Copies will get deleted by mistake. PDFs get misfiled in the wrong directories. Naming conventions are not adhered to or mistakes are made.
- Files begin to get scattered as they are copied and moved for temporary tasks like keeping new documents apart for quick review, bulk operations, printing, etc.

- Each batch of disclosed documents for each case now needs its own directory. People must remember where these all are. They must be secured manually to prevent accidental deletion. Simple tasks like tracking disclosure history or finding all parties who've received a particular a letter from a specific individual, can be a headache.
- Full text searching is done in isolation from other software such as CaseMap. Therefore, you can't do simple things like select documents from the search results and tag them with a new issue, correct the document type, pass them to colleagues for review, and so forth. Conceptually simple tasks need many extra steps when file folders are used -- something you only realize after the fact.
- Going mobile -- to court, your client's offices or simply working from home -- compounds these issues. Trying to solve this problem for file folders requires yet more software; it adds complexity. One such product's system administrator's guide -- for multi-user set-up -- is 30 pages long! By contrast MasterFile's replication and synchronization is built-in. It requires no configuration. One click replicates and synchronizes one case file or 20 -- it doesn't matter. And it's built on Lotus Notes: perfected over 20 years, unmatched -- reliable and secure enough for the US Navy's global network.
- Security and confidentiality also becomes problematic. It's almost impossible on Windows servers without any kind of document system and even then still not easy, but now, with users copying evidentiary and other files to other directories, notebook computers, etc., or for temporary tasks, it's not feasible.

Let's see how these issues are addressed by MasterFile's OCR/PDF Evidence Cruncher and Acrobat Professional, and by comparing MasterFile itself to CaseMap.

MASTERFILE'S OCR/PDF EVIDENCE CRUNCHER AND ADOBE ACROBAT PROFESSIONAL

The OCR/PDF Evidence Cruncher is tightly integrated with MasterFile's PDF and document repository -- both of which were designed specifically for the PDF tasks of legal or investigative professionals. In particular the Evidence Cruncher lets you:

- Select documents to process individually or in batches, directly in the repository, based on issues they are tagged with, or by author, by date, via full text search, etc.
- Automatically convert document files from over 30 widely used applications, such as Word, Excel, Outlook e-mail, Lotus Notes e-mail and/or

databases, etc., into PDF format. PDFs are automatically full text indexed and matched and loaded with the source documents back into MasterFile.

- OCR image documents; generated OCR text is automatically stored in MasterFile so critical information and facts can be immediately culled into MasterFile's unique Extract Repository and Fact Management system.
- Prepare documents for disclosure ensuring redactions are burned in so they can not be tampered with; redacted information is no longer present in disclosed files without requiring additional software to remove it.
- Bates stamp disclosed pages and track document disclosures so you know exactly what documents were disclosed, when, and to whom.
- Create a MasterFile briefcase of disclosed documents, complete with all redactions and Bates stamps, for subsequent reproduction or reference.

The OCR/PDF Evidence Cruncher is a powerful tool designed to efficiently handle high volume OCR and PDF conversion -- up to several thousand pages per day on dedicated computers. Think of the OCR/PDF Evidence Cruncher like a photocopier -- you need one in the office, but not for every user, nor is it used all the time. Similarly, only one Evidence Cruncher per 5-10 users is typically needed.

In contrast, Acrobat Professional is a stand alone product which does not include a document or PDF repository. Therefore, although the PDF document format is an ideal format for legal use and litigation, Acrobat itself was not designed for the specific tasks related to evidentiary documents that lawyers and litigators require.

Without a litigation support document/PDF repository, Acrobat Professional's ability to cope with the above three issues is missing and the process of managing your PDF files takes on all the complexities, error prone and multi-step, manual, procedures set out in the first section of this article.

MASTERFILE'S PDF DOCUMENT REPOSITORY COMPARED TO CASEMAP'S FILE FOLDER LINKS

Since it is not possible to cover every advantage or disadvantage that flows from MasterFile or CaseMap, respectively, below we give just a flavour of the differences between MasterFile's integrated solution, where all the parts support each other and work as a "team", versus using file folders and separate software packages for facts, OCR, PDF, full text search, evidence management, etc.

MASTERFILE'S PDF/DOCUMENT REPOSITORY	CASEMAP'S FILE FOLDERS
MasterFile is a unique integration of OCR/PDF technology, Advanced Case Analysis and Fact Management, PDF and Evidence Management	CaseMap stores all your documents, or e-mails converted to PDF with its DocPreviewer tool, in file folders whose integrity is entirely

<p align="center">MASTERFILE'S PDF/DOCUMENT REPOSITORY</p>	<p align="center">CASEMAP'S FILE FOLDERS</p>
<p>(or Litigation Support as it is commonly known) and work product management (or Document Management), in a state of the art package.</p> <p>It stores PDF and other documents in its unique repository. The repository manages all your documents -- PDF or not, in or out of the office -- and gives you far more benefits over storing documents in directories, than can be described here. Some of these include:</p> <ul style="list-style-type: none"> • All PDF documents, all source documents, all versions of the PDF documents, everyone's notes and comments, all critical information that's been culled, all document administrative information (security information, what's been disclosed to whom, etc.) is stored in one place instead of being scattered across thousands of PDF files, hundreds of source document files and stored in dozens of directories in servers, desktop and notebook computers. • Locating any document, any version of a PDF document, etc., literally takes just 2-3 mouse clicks. • MasterFile's unique repository manages both your evidentiary and work product documents eliminating the need for two systems just to handle your documents. MasterFile integrates with almost every application. • MasterFile's Extract repository lets you capture and see all the important information you've found in all your documents at a glance, rather than opening each file to review its sticky notes or keeping regular printouts of each file's sticky notes and reviewing those, one by one. • Collaboration is simple and secure. MasterFile's doc-link technology lets you send documents, or even extracts, for review to colleagues and co-counsel without e-mailing complete documents. Sending a document to the wrong person by mistake is a thing of the past. MasterFile's security ensures only those who have access to the document will be able to access it. • Sophisticated full text indexing, with fuzzy and proximity search, is built-in -- not an 	<p>dependant on every user correctly performing manual procedures -- which as we've noted above, has many serious problems. But CaseMap also creates some additional issues:</p> <ul style="list-style-type: none"> • Referencing documents in your notes, comments or facts in CaseMap is done via a "short name". This quickly becomes futile after only a few hundred end up being created for the documents of only one case. Alternatively, CaseSoft recommends each file be renamed and given a "short name" based on a CaseMap specific ID number created by the CaseSoft "Bates" Stamper -- making this issue even more problematic. We'll touch on this in more detail below. • CaseMap depends on Acrobat's full text search capabilities, which is based on correctly selecting one or more previously created Acrobat "indexes" for a set of folders. This also limits you to working with search results of one case at a go even if you've got two CaseMap files open. If the contents of the folders have changed, by another colleague for example, these indexes will be out of date and results will be incorrect. Moreover, since searching is done in isolation from CaseMap, you can't use the results of the search for common tasks, like simply tagging the documents found with a new issue. • Going mobile with CaseMap is also problematic. Documents, PDF or other, are referenced in a CaseMap file as "linked files". This means that the documents can never be moved or the links break and CaseMap will not find the documents again. For mobile use, this simply means they are not available unless you copy them all to your notebook and ensure both the drive letter and path on your notebook matches that of your desktop or server. • CaseMap's replication and synchronization technology suffers from many of the deficiencies listed in our article on mobile technology -- in particular: documents are not replicated, you can only replicate once (if you forget, you'll have to redo your work), you must force everyone out of the case or off the server to replicate, synchronization is one way, etc. <p>The bottom line is that by using directories, besides many functional limitations and manual</p>

MASTERFILE'S PDF/DOCUMENT REPOSITORY	CASEMAP'S FILE FOLDERS
<p>add on or separate product -- and updates itself automatically as documents are added to the repository. You can immediately manipulate, process, cull facts and so forth from the documents returned in your search results, right in MasterFile.</p> <ul style="list-style-type: none"> • You can directly load e-mail from Outlook or Lotus Notes. • Going mobile is a pleasure: <ol style="list-style-type: none"> a. You have full unrestricted functionality on or off-line and all your documents with you. b. You're able to synchronize all your work bi-directionally with your office: you can update them and they can update you, whether you're working one case or a dozen cases, with just one mouse click, even over the internet. c. Databases, documents, notes, argument, etc. as well as communications are protected with transparent, on-the-fly industrial strength encryption. 	<p>processes to cope with, that can compromise integrity, you also end up creating many discrete document collections or "repositories" -- i.e. subdirectories or folders of documents for each case, cobbled together into a "system" by multiple software products. For example, you end up with:</p> <ul style="list-style-type: none"> • folders for linked documents, • folders for virgin copies of PDF files without CaseSoft "ID" stamps so you can properly Bates stamp them on disclosure, • folders for each disclosure set, • folders for the original source documents, • folders with temporary copies of documents just for printing, Bates stamping and other administrative jobs, which must be deleted afterwards, • folders on notebook computers made by mobile users, and so on.

PRICE COMPARISON

Since a MasterFile Evidence Cruncher is not needed for every user, we compare price based on one Evidence Cruncher per 5 users or one Acrobat Professional and CaseSoft Bates Stamper per 5 users.

	MasterFile -- (advanced case analysis, a PDF/evidence repository and document management) with the OCR/PDF Evidence Cruncher	CaseMap [®] -- (case analysis only) with the CaseSoft Bates Stamper and Adobe [®] Acrobat [®] 7 Professional
1 user	\$1,390	\$1,214
5 users	\$1,990	\$2,873
10 users	\$3,895	\$5,746

MasterFile includes built-in PDF capabilities so tasks such as creating PDF files from any application or performing PDF text extraction (not OCR), are available to every user. Four copies of Acrobat Standard adds that capability for the five CaseMap / Acrobat Professional users in our comparison -- at an additional cost of over \$1,000.

Price difference alone, however, does not reflect the true value of the MasterFile solution because MasterFile offers many advanced features and modules you would have to add to CaseSoft + Acrobat Professional, including:

- A litigation support document repository at about \$2,000 per user from Summation® or Concordance®.
- Comprehensive document management at about \$400 per user from Worldox®.
- Robust, secure, scalability to handle multiple users or offices and mobile teams.

Therefore, a complete solution built around CaseMap + Acrobat, comparable to MasterFile, would cost an additional \$2,000 for a solo user and over \$10,000 more for the 5 or 10 users of a small office.

CASEMAP “BATES NUMBERS” AND DOCUMENT NAMING CONVENTIONS, CASEMAP’S PDF LOADER AND DOCPREVIEWER

Before we end, we'd like to take a few moments to discuss these a little further. Often such "features" are simply attempts to overcome fundamental design restrictions, or attempts to use a feature for a task it was not intended -- creating risks and more problems for you.

WHAT IS BATES STAMPING?

As we know, Bates stamping is simply a unique page number stamped on each page of each document to be used as evidence. At court, then, one simply has to say "turn to page 3278". Bates stamping, therefore, is used on the set of evidentiary documents that leave your office and generally not for documents that arrive at your office. However, although some law firms do stamp every document that does arrive at their offices with a unique document ID, let's be clear: this is not Bates stamping but simply a method to index documents.

CASEMAP DOCUMENT NAMING CONVENTIONS AND CASEMAP “BATES NUMBERS”

Document repositories which have integrated PDF technology for document production, like MasterFile, not only make true Bates stamping a simple task, but also allow you to track which documents have been disclosed to which parties, and manage the stamped documents so you can reproduce document sets (complete with Bates numbers, redactions, etc.).

CaseMap, as we've noted above, requires you to provide every document with a "short name" which is used whenever you need to reference that document in notes or commentary.

Well, it doesn't take much imagination to realize the big problems looming on the horizon with the "short name" system. First, these are soon going to be forgotten, especially once you've entered a few hundred for a dozen or two cases. Secondly, your colleagues are going to have an even harder time figuring out what the name for a particular document is -- one reason why the desktop and server search tools were developed. Naming conventions can be used but they are soon abandoned and also staff will inevitably make errors when hundreds or thousands of documents are involved.

Now, to "solve" these problems, CaseSoft recommends that you use the starting "Bates" number of a document as the "naming convention" assuming that thereafter, users will "know" the short name (i.e. the "Bates" number) as the file name for a document when they need to refer to it. Curiously, however, CaseSoft's Best Practices video webinar on document naming conventions, hosted by Dan Siegal, specifically advises you not to use the Bates numbers in any sort of naming convention. Instead, he recommends you use the date, document type plus the author as the naming convention. He then demonstrates this by retyping all this 3 or 4 times in various places to complete one document entry. For example, the date is retyped 4 times: in the document file name, in the date column on the documents spreadsheet and in both the long and short names for the document.

Nevertheless, as explained above, Bates numbering is used for those documents you've already decided are important for your case. In other words, Bates stamping is one of the last things you do with your documents after you've finished working with them in CaseMap or elsewhere. What CaseSoft is suggesting by recommending you use Bates numbers, is that you use these numbers when you start your case analysis -- but these numbers don't exist at that point! So CaseSoft actually had a problem: documents didn't have Bates numbers yet they were advising people to use Bates numbers to get around their "short name" problems. Enter their new "Bates" Stamper to create page IDs and their PDF Loader to link CaseMap to the corresponding PDFs.

It is doubtful that these unique page IDs created by CaseSoft "Bates" Stamper will be used as Bates numbers as rarely is every document used in analysis of a case disclosed and used as evidence. Therefore, a subset of the documents will be re-stamped with the true Bates numbers when they are disclosed which creates its own issues -- such as, for each of your cases:

- making sure you have a complete, un-stamped set of all your PDFs for re-stamping (as CaseSoft specifically recommends),
- from this un-stamped set, manually identifying (file by file), selecting and copying the specific documents to be re-stamped, and
- remembering all these new IDs after reloading the CaseMap file with the new Bates numbers.

Perhaps "CaseSoft PDF ID Stamper" would have been a better name -- but who needs more IDs? Now, we're not sure about others, but here at MasterFile when

we need to refer to a document in notes or comments, our staff have a problem trying to remember and provide ID numbers, let alone keeping track of them if they're changing!

The fact remains: any design based on using disk file names, short names, document IDs, naming conventions or whatever else we wish to call them, to locate and identify documents or simply reference them in your notes, is weak -- each method for it's own reasons.

THE CASEMAP DOCPREVIEWER

CaseMap 8 adds the DocPreviewer utility at an extra \$400. It includes the former Bates Stamper and PDF Loader utilities but now, also converts Outlook e-mails into PDF. As CaseMap tells you, you need to put these newly minted PDFs somewhere, which means you simply have more folders and their PDFs to manage as we've explained above. A key feature is e-mail attachments are embedded as PDFs within the PDF the e-mail was converted to. It looks tidy -- until you need to "send [some info] to CaseMap" from that attachment. At that point, CaseMap is quick to point out, you must save the PDF to your PC, load it in CaseMap and then, you can finally send that info from it to CaseMap. More manual work for you.

MASTERFILE'S EASY E-DISCOVERY AND LITIGATION PLATFORM

In contrast to CaseMap's collection of ad-hoc utilities, MasterFile is a complete e-discovery and litigation platform. Therefore, loading case documents, e-mails, transcripts, etc., is uniform, straightforward and simple. OCR, conversion of hundreds of formats to PDF, document production, etc., is automatic. Any document can be pinpointed in its unique repository in 2-3 clicks -- there are no folders or anything to manage -- and documents, or extracts of key information or facts can be referenced via MasterFile's unique doclink technology where the case analysis needs. See the following article also for more detail: "*Using the OCR/PDF Evidence Cruncher and Express Loader – PDF conversion, OCR or loading documents and e-mail doesn't get any simpler*".

OUR SHORT TAKE ON ELECTRONIC STICKY NOTES

Adobe and others have suggested using Acrobat's electronic sticky notes for evidence mark-up. Before we end this article, we'd like to show you some inherent risks in this approach.

In the physical paper world, yellow Post It notes worked for evidence mark-up because besides holding your thoughts, they also served as bookmarks to critical evidence culled. You could see them sticking out here and there in a stack of documents, no matter how large. And since there was no searching needed to

find them, you could instantly put your finger on the critical information they flagged.

This bookmarking aspect is critical for evidence mark-up and does not exist with electronic sticky notes. In fact, only MasterFile's unique Extract Repository was designed specifically for evidence review to retain, and even improve, on this original bookmarking aspect of paper Post It notes.

Electronic sticky notes in Acrobat were not designed for flagging critical information in evidence -- and particularly not for thousands of pages of evidence. Instead they were designed for receiving feedback from others during the editing and review process of documents in draft stage. For example, you may draft a contract, send out a PDF copy for review, and ask others to provide feedback. Sticky notes specifically include status options such as "Accepted", "Rejected", "Completed", etc. and can be tracked with "Tracking tool". In other words, the sticky notes are added to a handful of documents that you are actively working with and tracking.

Very quickly for litigation the problem becomes: which PDF files have notes? You have to look file by file. Full text searching doesn't help as you have to remember what to search for and only partial results may be returned by each search. And, if you aggregate the notes, your index is out of date the moment a new one is added.

By using electronic sticky notes inside PDF files, critical evidence or information will be lost. Can you afford that? In contrast, MasterFile's live Extract Repository always shows you what's relevant in any or all your documents at a glance.



For more information about
MasterFile please contact us at:

MasterFile North America, Inc.
Unit 152 - 11782 River Road
Richmond, BC, V6X 1Z7
Canada

Phone: 1.604.331.0535
Fax: 1.815.550.1107
E-mail: info@masterfile.biz
Web: www.masterfile.biz

All prices are in US dollars as of December 2007. Prices and features are subject to change without notice and refer to CaseMap 7 and Acrobat 7.

All statements, technical information, and recommendations in this document and in any guides or related documents are believed reliable, but the accuracy and completeness thereof are not guaranteed or warranted, and they are not intended to be, nor should they be understood to be, representations or warranties concerning the products described. MasterFile publications may include technical inaccuracies or typographical errors. THIS DOCUMENT IS INTENDED FOR USE AS A GUIDELINE ONLY. THIS PUBLICATION IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NON-INFRINGEMENT. MasterFile reserves the right to make changes to the information described in this document at any time without notice and without obligation to notify any person of such changes. Please refer to our legal page for further information.

© 2009 MasterFile Software North America Inc. All rights reserved. MasterFile and Evidence Cruncher are trademarks used under license by MasterFile Software North America Inc. All other trademarks are the property of their respective owners. Information contained in this document is subject to change without notice.