

Litigation Support Technical Standards

v1.8

Principal Author

Mark Lieb
Ad Litem Consulting, Inc.
PO 124873
San Diego, CA 92112-4873
(619) 546-7750
<http://www.AdLitem.com>

Editor and Sage

Browning Marean, III
DLA Piper Rudnick Gray Cary
401 B Street, Suite 2000
San Diego, CA 92101-4240
(619) 699-2700
<http://www.DLAPiper.com>

Table of Contents

Litigation Support Technical Standards.....	1
Preface.....	4
1.00 Introduction.....	6
1.01 For Vendors	6
1.02 For Firms.....	6
1.03 How to Use This Document.....	7
2.00 Business Standards.....	8
2.01 Outgoing Media Kit	8
2.02 Cost Codes for Litigation Support	8
2.03 RFQs	8
2.04 Quotes	9
2.05 Weekly Updates	9
2.06 Quality Control	10
3.00 Technical Standards	12
3.01 Media Labels.....	12
3.02 File, Folder and Volume Naming	15
3.03 CD Content and Organization.....	15
3.04 Organization of Sub-Folders.....	16
3.04.01 Images Folder.....	16
3.04.02 OCR Folder.....	16
3.04.03 Data Folder.....	17
3.04.04 Project Folder.....	17
3.04.05 Attach Folder	18
3.05 Bates Schemes	18
3.06 Data Files	19
3.07 Database Conventions.....	21
3.08 Native Files	21
3.08.01 Spreadsheets.....	21
3.08.02 Relational Databases.....	22
3.08.03 Deduplication, Keywords and Culling.....	22
3.09 Project Specifications Document.....	22
3.10 Bibliographical Coding Manual.....	23
3.11 Image Format.....	23
3.12 OCR	23
3.13 Slip-Sheets or Unitization Rules.....	24
3.14 Video.....	24
3.15 Synchronization	25
3.16 Transcripts.....	26
3.17 Delivery Media	27
4.00 Software Specific Requirements.....	28
4.01 Casesoft Suite.....	28
4.02 IPRO	28
4.03 Dataflight's Concordance and Opticon	30
4.04 Image Capture Engineering	32

4.05 RealLegal	33
4.06 LiveNote	33
4.07 Summation	33
4.08 iCONNECT	33
4.09 Trial Director	33
4.10 Sanction2.....	33
4.11 JFS Litigator’s Notebook.....	33
4.12 This Space Can Be Yours	34
5.00 Examples of What Not To Do	35
5.01 Media Labels.....	35
5.02 File / Folder / Volume Name Conventions	35
5.03 Database.....	35
5.04 Media Content.....	36
5.05 Load Files.....	36
5.06 OCR	36
5.07 Opticon Load Files.....	37
5.08 Image Format	38
5.09 Transcripts.....	38
5.10 General Errors / Issues	38
5.11 Real Experiences.....	39
Glossary	40
About Mark Lieb.....	41
Consulting for Firms	42
Consulting for Vendors.....	43
Litigation Support Department - The Book	44

Preface

Aside from the Introduction section, I tried to make every section generic enough that anyone can use it.

Please respect and include the copyright information. Creating this document has taken more hours than I care to admit.

I am also looking for feedback. Let me know if you thought the Standards document was helpful in any way. Do you feel certain titles or issues are missing? If your company wishes to be included in a future revision, please contact Mark Lieb at Ad Litem Consulting, Inc.

For The Law Firm and Litigation Support

This document is undergoing semi-regular adjustment. Originally this document dealt with basic standards for setting up Firm network folders and subsequently standardizing "load files". Your vendors should use the firm standard on all projects, whether they come from a paralegal or the litigation support department. Make sure all of your vendors get a copy of this document. You want to make sure they can deliver what you need the way you need it. This is also a great way to pre-qualify new vendors.

While this document outlines a unified approach to storage and dealing with discovery of all kinds, it doesn't really show you the even bigger picture: organizing the department and running it as a business.

To read a sample, please visit: <http://www.Litigation-Support.info>.

If your firm has a litigation support department, ask them how often, on average, they have to "correct" what the vendor delivered.

For Vendors

Litigation Support will be your biggest cheerleader if your product doesn't represent extra work for them. Any deviation from the standard results in more work for Litigation Support. As the customer, this makes Litigation Support dislike certain vendors. Litigation Support bills the client, just like the attorney or paralegal. This means time spent fixing vendor product may mean a higher bill passed along to the client for work the vendor should have done.

If you are interested in learning about how to cater to litigation support and make them your biggest internal cheerleader, please contact Mark Lieb.

These are the technical specifications Litigation Support vendors should match in order to provide the best work product. When these standards are not met, the product may require hours of Firm billable time or a follow-up delivery by the vendor. By providing this firm standard, it is our hope that we can lessen or remove this problem.

Irrespective of who initiates a project, the firm will always expect these standards to be met when the vendor creates product.

For All

I will endeavor to expand the document to include all major applications. I encourage software makers to contact me for inclusion. My thanks to those who have already contributed.

The latest version of this document is available by visiting <http://www.eDiscovery.org>.

1.00 Introduction

This document was initially designed to eliminate any discrepancy between firm technical needs and how the vendor created the technical aspect of their products. Litigation Support spends needless hours changing the vendor delivery. The firm pays for product that litigation support will have to modify. Today, the document covers as many technical requirements as possible for as many types of discovery and software as possible.

To get a good idea of the reason for these explicit directions, please visit the final section of this document entitled, "[Things not to do](#)". All of these examples are from real life. All of these examples caused headaches, delaying reviews, productions and more.

I hope that this document is helpful to you.

1.01 For Vendors

Litigation Support strongly encourages you to always contact us about new projects and bids. We can supply you with the latest standards and requirements. Please consider our requirements in your bids.

To contact Litigation Support:

Name _____ email _____
Phone (_____) _____ --- _____ Fax (_____) _____ --- _____
Name _____ email _____
Phone (_____) _____ --- _____ Fax (_____) _____ --- _____
Name _____ email _____
Phone (_____) _____ --- _____ Fax (_____) _____ --- _____

Delivery of products that does not meet our technical standards and requirements may result in nonpayment. Depending upon time restrictions or other factors it may be unfair to bill our client for time spent adjusting vendor product to published requirements. When possible, product with significant deviations from Firm standards will be returned for correction by the vendor.

As we require client matter number and attorney name on all invoices, make sure you get it up front.

1.02 For Firms

Litigation Support

Litigation Support must be diligent in providing the appropriate requirements to the vendors so that they may have the greatest chance of proper quoting and successful creation of their product for the Firm. Remember, if Litigation

Support doesn't adhere to these standards itself, how can it expect others to do so? As example, the client matter number and attorney name are required on all invoices by vendors. Make sure they have it up front.

The Standard is also how you can prequalify and exclude other vendors. Once vendors can successfully create what the firm needs, why use any who cannot? Prequalify additional vendors with this document. Vendors can take the initiative and deliver a sample production along with the standard to the department.

Attorneys and Paralegals

The lack of a standard when vending discovery means that non-technical people may make technical decisions and the vendor will use their own default technical settings. This will directly impact your ability to litigate. If your firm has not adopted standards, I encourage you to work with litigation support to formalize one for yourself.

If your firm does not have an internal litigation support department then this document can be especially useful for you. Once your standard is established, you do not need the technician to outline them for every subsequent matter. Even firms using a hosted solution, or "ASP", will eventually create a production or use the database elsewhere. The requirements outlined in this document can help your firm make certain the vendor product can be used elsewhere.

1.03 How to Use This Document

This document is a template for a unified approach. As example, the information on every type of label (CDs, tape, hard drives, etc.) is contains consistent information such as client-matter number. If there are load files, the paths always follow the same structure. Change the names, cost codes and contact information, but please leave the names on the cover.

To learn more about implementation, please visit <http://www.Litigation-Support.info> or the pages on [consulting for firms](#) and [consulting for vendors](#).

2.00 Business Standards

This part of the document outlines what information each player (vendor, litigation support, paralegal or attorney) should include from a business standpoint. As example, the firm uses “cost codes”. This is how the firm tracks how much it spent on a service such as computer forensics for a given case or year. At the same time, this document outlines what information the firm should provide to the vendors.

2.01 Outgoing Media Kit

Every project must include this document. If the vendor follows the technical standards included in this manual, returned product should be available to the legal team very quickly. This will also result in a quicker turnaround for paying invoices.

Vendors will do their best to meet the client requirements. It is unfair and difficult to impose these requirements after a project is completed or in process. Therefore, it is the responsibility of the Firm and Litigation Support to present the standards and practices manual with RFQs and outgoing data for processing (e.g. paper for scanning or video for conversion).

2.02 Cost Codes for Litigation Support

In order to facilitate the billing process the accounting department would like to have several pieces of information appear on every vendor invoice.

1. The *client–matter number* must appear on invoices, media labels and business correspondence (including emails) between Firm and vendor.
2. *Attorney’s name*, which will authorize the invoice, must appear on the invoice.
3. *Cost code* must appear on every invoice:

Cost Code	Description
001	<u>Litigation Video and Graphics:</u> Video and Graphics used at depositions, arbitration and trial.
002	<u>Litigation Support Data Services:</u> This covers the range of work such as: database creation, hosting, administration and such. Value added services such as OCR, programming, forensics, conversions, and media creation are all “Litigation Support Data Services”.
003	<u>Scanning and Printing:</u> Photocopying paper, scanning paper or printing a paper set from electronic source (a/k/a “blow backs”).
004	<u>Court Reporting Services – Non-Video</u>

2.03 RFQs

All RFQs should include several key pieces of information. As a vendor, please make certain these fields are completed. Missing or incorrect information will almost inevitably result in problems as the project matures. Note that some of the fields should appear on media labels, others on the invoices.

Field	Description
Attorney	The attorney who will “sign-off” and authorize payment.
Paralegal	The paralegal handling inventory, coordination with legal team.
Litigation Support	The technical contact for this project.
Client-Matter Number	A Firm created number, this uniquely ties the project to the case.
Project Name	Each project must have a unique name.
Collection Name	Usually the same as the project name.
Due Date	When the team must have the completed project.
Description of Materials	A description such as boxes, video tapes. The more specific the better.
Services Required	A list of services such as synchronizing text to video or auto-coding.

For certain types of projects, such a computer forensic work, it may be difficult to impossible for the vendor to provide a qualified estimate. Under these circumstances, the Firm prefers to pay an hourly rate for the vendor to gain a greater understanding of project scope. At the end of this hourly exercise, the Firm would like to then receive a formal RFQ.

2.04 Quotes

On vendor quotes, please reference the fields from the RFQ. We may be required to obtain multiple quotes. Therefore, knowing both the project name and client-matter number are critical. Please note that inclusion of the required fields as outlined in this document will curry favor with Litigation Support.

In order to properly compare one vendor quote to another, line-item pricing is the best format. In this fashion we can compare quotes. To this end, the law firm may include a preferred spreadsheet for quoting goods and services.

Quote File Format

The generally preferred formats for quotes are all electronic. In order of preference, most to least, here are the accepted formats.

1. Quoting Spreadsheet – This document is available from the firm or from <http://Litigation-Support.info>, as part of the book on the Litigation Support department.
2. An electronic format that allows for copy, paste and printing functions.
3. Hardcopy or fax

2.05 Weekly Updates

Overview

The legal team and support teams need weekly status and invoice updates. The update does not need to be too formal. The attorneys should always have a solid idea about the project progress and costs. It never hurts to communicate too much. This information helps everyone to plan and schedule their own efforts. If a project is going to be complete by a new date or the budget is heading away from the quoted target, these reports and forecasts will help all parties involved. These updates are all about zero surprises.

Please keep the firm informed about the following information.

Project Name

Please include this name in your quotes, reports and on your invoices.

Client-Matter Number

This number is required on all communications. This is how the firm tracks all projects and all payments and work relate to this number.

Project Complete %

On a weekly basis, the vendor should provide a report of all open projects and their status. Are there any issues which will cause the project to end more quickly than expected?

Current Bill

Projects, especially electronic discovery projects, can grow to unexpected sizes. This can happen due to various factors. The main concern here is that the attorney understands their current costs and is not surprised by the bill.

Original Bid

What was the original project estimate?

Paid To Date

What has the Firm paid to date for services on a project?

Work-In-Progress (W-I-P)

What the vendor has yet to bill.

Estimated Final Bill

Vendor gives an estimated grand total. This number is critical. It falls under the “told you so” part of the law that helps attorneys sign off on final invoices.

Estimated Completion Date

Please provide an updated completion date in your reports. In this manner, the team can plan their time and efforts accordingly. It is OK for this date to change. It is required to give as much notice as possible. It is better for the attorneys to schedule their time accordingly.

2.06 Quality Control

The firm notes quality in terms of what is written on the label and what is written on the CD.

The firm runs programs, which perform “QC” checks on every delivery. The QC programs evaluate the vendor delivery, looking for missing data or files, gaps and things of this nature.

It is important that the media label and format of the data comply with Firm standards. A CD containing critical data can be lost due to a handwritten label that simply reads “Data”. At the same time, if the format of data contained on said CD is unusable, it does no one any good. It is a safe bet that every law firm has CDs in inventory that do not match what is on the server. In order to make the CD look like the server did (in case of deletion or crash) may take hours.

Content

Media labels, invoices and other important correspondence should always include certain information.

Test	Description
Client-Matter Number	Did vendor provide required information?
Responsible Attorney	Did vendor provide required information?
Responsible Paralegal	Did vendor provide required information?
Project Name	Did vendor provide required information?
Description	Did vendor provide required information?
Services Required	Did vendor provide required information?

Format

These tests look at things such as naming conventions and matching up image keys to the physical file.

Test	Description
Image File Count	Number of image files on delivery.
Document Count	Number of documents in the database load file.
Cross Reference Image Count	Number of images referenced in Opticon load file.
Cross Reference Document Count	Number of document breaks in Opticon load file.

Image File – Cross Reference Match	Does every image listed in the Opticon load file exist? Does every image file on the CD exist in the load file?
Document – Cross Reference Match	Do document start and end number have matching entries in the imagebase load file? Are all Opticon load file ranges in the document database?
Database Load File	We are looking at field delimiters, gaps and a field name header line.
Full Text Formatting	Was the format maintained or were there “odd” characters such as semi-colons instead of spaces.
Folder and File Naming	Does the delivery comply with naming standards and use client matter subfolder structure?

3.00 Technical Standards

This section of the document clearly outlines what information to include and how to format the information. As example, what information should appear on every label, from CD to VHS tape. This section also covers how the Bates, file, folder and volume naming conventions work.

The vendor should immediately alert the firm if it is unable to meet these standards. Litigation Support and the firm will use this document to identify any deviations from the requirements. The law firm is very serious about consistency of product within a case and across all cases.

3.01 Media Labels

The following information should be visible on the CD, DVD, Tape, or a space provided where the information can be written clearly later. Some vendors write the Firm information on the labels of CDs and DVDs. It is more helpful to have the contact information for the media's creator.

Please reference the examples on the following pages.

Required Fields	Sample Values and Examples
Vendor Name	ACME Scanning and Coding
Vendor Address	123 Main Street
Vendor Phone	(555) 555-1212 (voice) / (555) 555-1213 (fax)
Date of Media Creation	12/02/2003
Format Type	CD, DVD
Volume Name	Examples are "FER001, FER002, FER003"
X / Y	Examples are "1 / 3", "2 / 3", "3 / 3"; or just "1 of 1"
Bates Ranges	Examples are "FER000001 – FER001300"
Client-Matter Number	Example, "320123 – 00123"
Image Count	13,000 TIF images

A Note on Handwriting

A hand-written label is not acceptable. Handwritten labels are cryptic and illegible. Further, a pen can damage the media. Five minutes time to create the label can save an hour of research. The purpose of the label is to provide a complete picture of the origin and content. If a label must have handwriting, script is preferred. Cursive handwriting is never acceptable. Bear in mind that other people have to read your writing.

Remember, the firm is making significant investment of the client's money. The media provides the first impression of the vendor's work. Perception is reality. If the delivery looks boiler room, then the Litigation Support persons will view the vendor as boiler room quality. If a vendor is sloppy on their label, chances are they are sloppy in all else they do.



Basic Information – Example with Play Data

3.02 File, Folder and Volume Naming

The following explains how data should be organized on your deliverable to the Firm. Please let us know if you cannot meet this standard, and what standard you are prepared to deliver.

16-Bit vs. 32-Bit

Older computer systems could only use filenames and folder names of very limited length. This is known as the 8.3 naming convention. If a filename is wider than 8 characters, 16-Bit programs truncate the name. So suddenly the filename “AMURPHY0000001.TIF” becomes “AMURPH~1.TIF”. If the vendor is using older software, they may be restricted to 8.3 filenames. This is a serious problem about which the vendor must contact Litigation Support to determine a remedy.

File and Folder Names

1. Only the characters A...Z and the numbers 0...9 are valid
2. Filenames should be unique, matching the image key
3. Image folders should be zero-padded to 3 wide (i.e. 001, 002, 003, 004...)

NOTE: **The filename must match the image key.** The only exception is where the image key contains additional characters that must be echoed in the TIFF file name. A filename such as “000001.TIF” is absolutely unacceptable. If the vendor cannot provide the full image key, please alert the firm immediately so that the firm may find a vendor who can.

While the following is syntax for an Opticon “.LOG” load file, please use the same logic when generating the IPro “.LFP” and Summation “.DII” formats as part of every delivery. For additional examples, please look at the section [4.00 Software Specific Requirements](#).

Database Image Key	Cross Reference File		Actual Filename
	Image Key	Path To TIFF	
A001	A001	D:\A001\IMAGES\001\A001.TIF	A001.TIF

Volume Names

Each CD should conform to the same standard: [PROJECT NAME][999]. So, if our project is named SMITH, the first three CDs delivered should be named: SMITH001, SMITH002, SMITH003. Note the zero-padding.

Unless the project name is “VOL”, the volume name of the first CD should never be “VOL001”. Many applications use the name “VOL” as a default value. This has resulted in many CDs named “VOL001”. This can make identification of the related case and content difficult. Do not use the vendor name as the volume prefix. Use project name as the volume name. The Bates prefix can be an acceptable project name as the volume prefix, but must confirm with the firm as to final decision. Vendors should never use their company name as the prefix. Some firms prefer to use the client-matter number as part of the project name.

3.03 CD Content and Organization

Each CD should contain the same folders each time. This structure is important, as the media is not copied to a single subfolder. Instead, “data” goes under a different folder tree than “images”. If not segregated, Litigation Support will have to perform this separation.

D:\[VOLUME NAME]\	Your CD should have a root folder, named the same as the volume name.
D:\[VOLUME NAME]\IMAGES\	All images and image subfolders reside here.
D:\[VOLUME NAME]\OCR\	This folder contains multi-page ASCII text files. The filename matches the “begBates” key, e.g. A001.TXT, A011.TXT, and A013.TXT.
D:\[VOLUME NAME]\DATA\	All “load”, “database”, “structure” and technical files reside

- | | |
|---------------------------|--|
| D:\[VOLUME NAME]\PROJECT\ | here.
1. Document coding instructions,
2. Project manuals,
3. Vendor contact information,
4. Source information,
5. Ranges information. |
| D:\[VOLUME NAME]\ATTACH\ | All native files reside here, as applicable. |

NOTE: Each CD must be self-contained. This means a CD containing A001...A010 must contain the images, database load file, OCR and cross reference file for A001...A010. A delivery of CD01...CD10 should have the load files for CD01 on CD01. Having load files for CD01...CD10 all reside on CD10 is incorrect. If the firm loses the "load files" CD, the corresponding CDs may not be usable. Further, this means tracking down 2 CDs every time there is a problem.

3.04 Organization of Sub-Folders

We understand that certain applications construct subfolders automatically in different configurations from that listed below. Therefore, this storage convention may not be possible for your organization without unreasonable effort. Do your best.

Standard sub-folders for each delivery: Images, OCR, Data, Project and Attach.

3.04.01 Images Folder

FOLDERS	CONTENT
D:\[VOLUME NAME]\IMAGES\0001\	IMAGE0000001.TIF...IMAGE0001000.TIF
D:\[VOLUME NAME]\IMAGES\0002\	IMAGE0001001.TIF...IMAGE0002000.TIF
D:\[VOLUME NAME]\IMAGES\0003\	IMAGE0002001.TIF...IMAGE0003000.TIF
D:\[VOLUME NAME]\IMAGES\0004\	IMAGE0003001.TIF...IMAGE0004000.TIF
D:\[VOLUME NAME]\IMAGES\0005\	IMAGE0004001.TIF...IMAGE0005000.TIF

NOTE: Zero-Padding is very important. This is especially true if you have 9999 subfolders.

One of the reasons why it is important to have a standard number of images per folder is so that Litigation Support can easily determine where missing images may reside. In this fashion it is also simple to locate the right folder based upon the .TIF filename. Empty or "skipped" subfolders are not acceptable.

3.04.02 OCR Folder

This folder contains multi-page ASCII text files named for the image key or the first page of the document.

Different document review systems load OCR in different fashions. As such, this document includes organization and formatting considerations valid for every format. To learn the actual technical syntax, please refer to the software section for examples. Regardless of software title, there are commonalities to each software, such as names and organization.

While different OCR programs produce different types of output, the firm requires the vendor's product to match the naming conventions and organizational schemas outlined here.

The OCR filename must match the complete image key. As most programs load the OCR by matching the image key to the text file, all OCR for the entire document should reside in the single image key text file.

IMAGE KEY	BEGBATES	ENDBATES	OCR FILENAME	CONTAINS
-----------	----------	----------	--------------	----------

A001	A001	A005	A001.TXT	OCR for A001...A005
A006	A006	A006	A006.TXT	OCR for A006
A007	A007	A070	A007.TXT	OCR for A007...A070

Please refer to the Document Coding Instructions for an explanation of “image key”.

3.04.03 Data Folder

All load files, except OCR, reside in this directory. While many vendors automatically include load file versions formatted for every major application, they all reside in the same folder. For the document review systems, there are basically two main load files: database load file and imagebase cross-reference load file. The first contains the discovery bibliographic coding including Bates. The second file is an index that correlates Bates to .TIF file.

Required files:

The following files should be found in every DATA folder of each delivery:

1. Database Load File
2. Database Structure File
3. Imagebase Load File

Database Load File Format

The database should be an ASCII delimited file. It is preferable to use the delimiters appropriate to the application. If the vendor wishes to include the OCR inside of the database load file, specific step-by-step instructions on how to load this data is required.

The first line of the load file should be the field names. This provides the database administrator a level of confidence upon seeing the field names and values line up perfectly.

Database Structure File

The database structure file is an ordered list of every field name, field type and size. In the case of a date field, the size should show format (e.g. MM/DD/YYYY, DDMMYYYY, YYYYMMDD, etc.).

Imagebase Load File Format(s)

Please provide the image “cross-reference” (a/k/a “xref”) load file in the following formats:

- Opticon
- IPRO
- Summation

All load files, along with paths and technical preferences are contained inside this document.

Naming Convention:

Name the data files to match the CD volume name. In this fashion it is possible to match any load file with the right CD.

3.04.04 Project Folder

These files identify the project, associated information such as attorney name and all the treatments such as stamping or OCR. These files are extremely helpful when identifying an “orphaned” volume. They are also helpful when an old project begins again after a one-year hiatus.

1. Bibliographic Coding Instructions:

This document identifies how the database was codified. This document shows which fields were coded and any rules around the codes themselves, such as valid document types.

2. Project manuals:

As available.

3. **Vendor contact information:**

A simple text file that tells us which vendor made the CD and their contact information.

4. **Source information:**

If there was an **intake form**, include it here. What was the source of the project – boxes or electronic discovery? Sometimes we need to backtrack from the produced CD to the originating data.

5. **Ranges information:**

Show a list of the Bates ranges on the CD.

3.04.05 Attach Folder

This is the home of native files. If there is a movie clip or spreadsheet associated with the collection, this is where it must reside. Of course, once on the folder, the full path will be:

X:\ATTACH\CLIENT#\MATTER#\DATABASE NAME\VOLUME\ATTACH\

FOLDERS	CONTENT
D:\[VOLUME NAME]\ATTACH\0001\	1,000 native files
D:\[VOLUME NAME]\ATTACH\0002\	The next 1,000 native files
D:\[VOLUME NAME]\ATTACH\0003\	And so forth
D:\[VOLUME NAME]\ATTACH\0004\	
D:\[VOLUME NAME]\ATTACH\0005\	

NOTE: Zero-Padding is very important. This is especially true if you have 9999 subfolders.

One of the reasons why it is important to have a standard number of native files per folder is so that Litigation Support can easily determine where missing attachments may reside.

3.05 Bates Schemes

While it is helpful to have a “significant” Bates prefix (“ACME” versus “A”), brevity is not without merit. Use the “KISS rule”: keep it simple, Simon. Just think how many times you will need to write or enter the prefix. Also, there are computer issues that almost mandate certain syntax. Please use the following conventions when constructing the prefix.

We recommend using the project name for the prefix. In this fashion, each collection, project, Bates and files carry the same name. If you have 5 separate discovery collections, you will have 5 projects, 5 unique Bates prefixes and 5 unique filename prefixes. When it is time to produce, a new Bates scheme can be applied.

Guidelines for creating a correct Bates prefix:

1. No more than 5 characters wide
2. Only use uppercase letters from A to Z
3. Bates prefix should not end in “L”, “O”, “I” or “D”
4. No spaces

Explanations / Examples:

1. Good Bates prefixes include: A, AA, AAA, AAAA, AAAAA.
2. Bad Bates prefixes include: 9A, 9A9, A9, A9-, A-A-A, 0-A-0, A 0001.
3. The reason for point 3, above, is that future OCR may mistake these letters for the numbers “1” and “0”.

Rules for creating a correct Bates suffix:

1. Suffix should be numeric
2. Suffix should be zero-padded to four (4) positions, or the ten-thousandths place.

3. Suffix should never contain spaces, hyphens, underlines or characters other than 0 through 9.

Good Bates suffixes include: .0001, .9999, .0100

Bad Bates suffixes include: .A, .0A, .A01, .A-1

Examples:

The following two tables will show the identical ranges using proper and incorrect prefixes and suffixes.

Correct Bates ranges:

BegBates	EndBates	Description
A001	A010	Prefix is short. Easy to see where prefix ends and
A011	A011	the Bates number begins.
A011.0001	A011.0026	Zero-padded and numeric ensures proper sorting and
A011.0027	A015	software friendly format.
A016.0001	A016.0010	

Incorrect Bates ranges:

BegBates	EndBates	Problem
A9001	A9010	Is the document number 1 or 9001?
A9011	A9011	Is the prefix A or A9?
A9011.A	A9011.Z	Suffix is a letter, resulting in sorting issues.
A9011.BA	A9015	Suffix contains letter(s), resulting in sorting issues.
A9016.1	A9016.10	Suffix has no zero-padding resulting in bad sorting. Also, inconsistent suffix naming convention.

3.06 Data Files

The following files must reside in the Data folder on every delivery by the vendor:

1. Database load file,
2. Database structure file, and
3. Imagebase cross-reference load file.

1. Database Load File:

1. Delimiters – Although this document does not truly favor one application over another, the Concordance standard delimiter characters have proven reliable time and again. They are:
Comma (020), Quote (254), Newline (174).
2. The first line of the database load file should be the field names.
3. The name of the database load file should match the volume name.

2. Database Structure File:

The Firm has a standard database structure we use for all databases (electronic and paper). We understand that software restrictions may result in a non-standard product. If the vendor is unable to then post-process their data to match our standard, Litigation Support will have to perform this work as billable time to the firm’s clients. As one can imagine, knowing the database structure as relates to the database load file is critical.

This is a text file showing a sample structure file. The following is just for illustration and does not match Firm standards. Please refer to the Bibliographic Coding Manual for this information.

Field Name	Type	Size
Author	Paragraph	-
Date	Date	MM/DD/YYYY
Title	Text	60
Pages	Number	3.0

3. Imagebase Load File.

The following are the rules governing a good load file:

1. The imagebase load file name should match the volume name
2. All images referenced in the load file must be contained on the same volume
3. Document breaks
4. Page counts
5. Image path:

D:\IMAGES\[CLIENT#]\[MATTER#]\[DATABASE]\[VOLUME]\IMAGES\...

Note: While the path may seem long, it provides everyone with a standard everyone can understand. The database folder may seem redundant at first. That is until there are 12 databases for a given matter number. At that time, one becomes grateful for the database subfolder. The firm uses this structure for many reasons. When the load file does not match this path, the vendor will have to fix this. If Litigation Support has to fix this, then the client may be paying twice for the same work.

Sample Opticon Load File:

Please note the path, image key matching the file name, doc breaks and page counts. Please refer to [Section 4](#), to see examples for other types of load files.

[Field 1]	[Field 2]	[Field 3]	[Field 4]	[Field 5]	[Field 6]	[Field 7]
A001	[VOLUME]	D:\[VOLUME]\IMAGES\001\A001.TIF	Y			2
A002	[VOLUME]	D:\[VOLUME]\IMAGES\001\A002.TIF				
A003	[VOLUME]	D:\[VOLUME]\IMAGES\001\A003.TIF	Y			1

Here is an explanation of the Opticon load file format:

- | | | |
|-----------|-------------------|---|
| [Field 1] | Production Number | This is a text field which contains the "Production" or "Control" or Bates number for that page of the document. It is a unique value and is the load file "key". |
| [Field 2] | Volume ID | This is also a text field. It should contain the Volume ID of the CD on which the images are delivered. |
| [Field 3] | Full DOS Path | This contains both the path to the image and the actual image filename. |
| [Field 4] | Document Break | This is a text field. If this particular image is the first page of a document, this field should contain a "Y" (Yes). |
| [Field 5] | Folder Break | This is a text field. It's fairly rarely used but if used is intended to work just like Document Break, i.e. it would contain a "Y" if this is the first page of a new folder. |
| [Field 6] | Box Break | This is a text field. Also rarely used but intended to work like Doc and Folder Break...would contain a "Y" if this is the first page of a new box. |
| [Field 7] | Pages | This is a text field although it contains numeric data. If this is the first page of a new document, "Document Break" will contain a "Y" and this field will show the number of pages for the document. |

Each of these fields is "separated", or "delimited", from the others, by a **comma**. When a technician imports a load file into Opticon, the content for each field is divided by the commas. Therefore, one can not have a directory named "5,312,591 PATENT" since Opticon will view each comma as the start of the next field, per below:

[Field 1]	[Field 2]	[Field 3]	[Field 4]	[Field 5]	[Field 6]	[Field 7]
[Prod Num]	[VOLUME]	\5	312	591 Patent	,	[Pages]

3.07 Database Conventions

There are two types of discovery: electronic and paper. Electronic discovery software extracts “metadata” from the file. The metadata contains fields and values ranging from author to the last print date. Different file types may yield fields (“metadata”) that do not exist in other file types. What does this mean to us?

Whereas an email should yield the author and date, a Word document or PowerPoint presentation may not. This means the firm needs to pay for bibliographic coding for certain kinds of electronic discovery to achieve a complete database.

All Electronic Discovery yields “full text”. Full text is quite literally all the text inside a word processing or spreadsheet file or any other electronic files. Full text extraction removes the need for OCR. Like OCR, the full text does not provide bibliographic coding such as author and recipient. Full text will however provide 100% accurate content where paper OCR may be 80% accurate or better, depending on the quality of the paper.

Load File Field Order

To help make life a little simpler for our legal teams, we outline the minimal fields that we require for each document, irrespective of origin, format or file type. This requires a certain amount of bibliographic coding for certain types of electronic discovery and all types of paper discovery.

As possible, the firm attempts to keep field order consistent for like types of databases. As such, the firm appreciates the vendor matching their load file to our field order. Our document reviewers expect to see the same fields in the same order for all databases. Please help us make this happen.

The exclusion of certain fields or their incorrect order may require Litigation Support to bill time to the client for correcting these problems. This is one reason why the Database Structure file (see 3.06) is so important.

While the vendor should provide the Firm with every field possible for electronic discovery, the following list from the law firm includes certain fields that we require at a minimum and in the following sequence. Depending upon the production or pre-production status of a collection, certain fields may contain no data.

Please refer to the *Bibliographic Coding Instructions* for bibliographic coding.

Note: These files, “load file field order” and “bibliographic coding instructions” should reside in the “DATA” and “PROJECT folders”, respectively, on the delivery.

3.08 Native Files

There are many types of electronic discovery. This section describes how the firm prefers to handle certain types.

3.08.01 Spreadsheets

Spreadsheets are not printer friendly. Spreadsheets are rarely formatted to accommodate the printed page. As such, when vendors “petrify” (convert to .TIF file) a spreadsheet, the results can be hundreds or thousands of pages of almost useless information and lots of blank pages.

The following is the approach the firm wishes to employ in representing spreadsheets in the database, review and production:

In the Database:

For every spreadsheet, there should be a record. Add the appropriate full metadata to the record. Include a hyperlink to launch the spreadsheet as a native file. For the image, put a placeholder explaining that the document is being reviewed as native file. Therefore there is no image of the spreadsheet.

The File:

Files should retain their original filenames and be read-only. These files go under the traditional X:\ATTACH\[client#]\[matter#]\[DB]\

Production Options –

Native: Spreadsheets should be produced natively. The CD label should have a Bates stamp. Included with the CD is a cover-page, bearing the next Bates stamp number. The cover-page should include a list of every native file, along with filename, file date and all other attributes.

Petrified: Spreadsheets should be converted to .TIF, per our technical standards. Any load files you return should reference the internal control number as well as the new production Bates start and end numbers.

Printed: Spreadsheets should be “blown back” to paper with any associated text and production Bates number.

3.08.02 Relational Databases

There are times when the data in electronic discovery will not make sense if printed. A relational database presents such a dilemma. A “relational” databases stores related information in multiple tables. Each table therefore only presents part of the actual picture. It is only by matching up records across the tables that the combined data presents useful and complete information.

The best way to present this information is to print a report using the relational database software.

Relational databases should be reviewed in native format. The document review system should include a record that references the relational databases.

In the case of an accounting system, the Firm may decide to employ an expert witness who will review in native format. Production may consist of reports the expert printed. Alternately, the other side may be able to send their own expert to use the accounting system and print reports for their review.

Other databases may result in printed pages as a result of review. The pages should be treated per the accounting system example.

3.08.03 Deduplication, Keywords and Culling

The attorney may decide to employ strategies to cull discovery, such as deduplication, keywords and date ranges. We also prefer to receive two databases: a privileged and a non-privileged database. The legal team may provide the vendor with privileged keywords, such as attorney names, in order to help the team identify and review documents which do not contain privileged terms but do meet other criteria.

3.09 Project Specifications Document

Every project should include a Project Specifications Document. This document outlines the scale and scope of the project at hand.

3.10 Bibliographical Coding Manual

This associated document shows the fields, treatment, valid values (such as document types) and other such standards. This information is not included in this document.

3.11 Image Format

The majority of documents imaged only require black and white. On a less frequent basis, we may need color images. The following are our standards:

1. Black and White images should be 300 DPI, Group IV TIFF;
2. Single page TIFF images;
3. Color images should be discussed on a per image or per document type basis;

Please use the following guidelines, borrowed from Adobe's recommendations when scanning images:

- Choose grayscale for variable contrast pages in the scanning software.
- Increase brightness and contrast by 10% for text on colored paper, or filter out the background if your scanner includes that option in the scanning software.
- Adjust manual brightness control on the scanner if your scanner includes that option (for example, a knob).
- Use 16-bit color or less; 24-bit color images take considerably more system resources in the scanning software.
- Don't use any dithering or half toning options in the scanning software.

Oversized Documents

Every oversized document should result in two pages in the database. The first page shows a legend, Bates or any other identifying marks. The second page should be a full-scale image.

Should the document be a five-foot color map, a full-sized .JPG may be required. In this case, an accompanying first page black and white image of just the map legend is still required.

3.12 OCR

Vendor must use auto-rotate when generating OCR. Most OCR software offers an auto-rotate option. When auto-rotate is enabled, the software will OCR each image four times, rotated 90 degrees each time. It determines the best result and publishes the content to the load file. The majority of documents have the same orientation: portrait. Without auto-rotate, these documents can yield good results. The rest of the documents may be designed for a landscape layout, such as an HR chart. Other documents still may have been scanned "upside-down", resulting in garbage OCR.

Quality Check

The OCR text should best approximate and recreate the formatting found on the original image. The OCR field should never be just the words in one long string.

Multi-Page Text Files

There should be a one document to one OCR text file ratio. The OCR filename must match the document image key. So, a 10 page document with the image key of AA001 should have a corresponding file AA001.TXT that contains the OCR for AA001 through AA010.

Each page of OCR should have a line identifying the page number, or Bates number. In this fashion, people can search for any Bates number and find the correct document. Please include space between the OCR text and page marker.

The following shows sample OCR:

<< AA001 >>

Text for first page

<< AA002 >>

Text for second page

The following chart shows a sample database and corresponding OCR files:

IMAGE KEY	BEGBATES	ENDBATES	PATH	FILENAME
AA001	AA001	AA0010	D:\[VOLUME NAME]\OCR\	AA001.TXT
AA011	AA011	AA0011	D:\[VOLUME NAME]\OCR\	AA011.TXT
AA012	AA012	AA0038	D:\[VOLUME NAME]\OCR\	AA012.TXT
AA039.0001*	AA039.0001	AA0100	D:\[VOLUME NAME]\OCR\	AA039.0001.TXT

* Please refer to Bates prefix and suffix conventions.

3.13 Slip-Sheets or Unitization Rules

If not already done by the client or the firm, the scanning company should place a slip-sheet between each document before scanning. After the documents are scanned, the vendor needs to provide logical document breaks. The Firm requires a 1 document to one database record ratio. Between slip-sheeting during scanning and the logical document breaks service, this ratio should be guaranteed. Slip sheets should not appear in the database or images.

The resulting database must maintain the parent-child document relationships through the “BegAttach” and “EndAttach” fields.

When the firm requires the vendor to print documents to paper, a non-white slip-sheet must separate every document. Blue and green are the

3.13.01 Firm Slip Sheets

On certain projects, the law firm will have already "slip-sheeted" the collection. If the firm

3.13.02 Vendor Slip Sheets

[prescan slip sheet versus vendor slip sheet (scanning)software specific slip sheet -mrl]

[blowback slip sheets may require extra info or simply a blank w/ color]

If legal team puts in plain, colored slip sheets to show vendor where to put in their slip sheets for scanning, they should put Bates/PROJ on there. Not vendor.

The reason a blank appears in the database is due to vendor not replacing with their appropriate slip sheet, or, unless there is a legitimate blank page]

3.13.03 Pre-Production Scanning

3.13.04 Blow-Backs

3.14 Video

While the most frequent purpose of video is for the deposition, there are other purposes where higher quality is necessary, such as showing fine details.

If the firm requires the use of a videographer, that vendor must provide the right format, lighting equipment and experience, meeting or exceeding firm standards. These are the details that can help guarantee a quality level that the attorneys will want to use.

One purpose of these standards is to make certain that our legal teams and clients have the best quality and formats. Another purpose is so that our legal team can concentrate on the law and not worry about deciding which formats to use. The final purpose is to make certain the legal team only works with the true professionals who have invested the time and money into their trade and art.

Format and Video Gear

If at all possible, all Firm commissioned video must be shot in a digital format. With the recent proliferation of consumer, mid-tier and professional cameras on the market, we would like to request the use of only professional level cameras. This means no Sony VX2000 or VX1000, Canon GL1 or GL2 or consumer level cameras.

Lighting Gear

Do not rely on what is in the room. A professional videographer will bring the necessary lighting equipment to make certain the video quality is good. If your videographer does not have such equipment, you may wish to consider using a different company.

Label Information

The following information should appear on every label:

1. Vendor Name
2. Vendor Address
3. Vendor Phone
4. Names of People appearing
5. Dates of appearances

Minimum Video Format Specifications

The Firm requires different encoding based upon the purpose of the video.

Deposition Video

MPEG1 encoded at 2 hour bit rate with quality compression hardware.

Video must be adequate for use at trial using Sanction II or Trial Director, in full-screen mode.

Other than Deposition Video

MPEG2 encoded with quality compression hardware.

For video such as a day-in-the-life video, site tour or outdoors where picture detail and clarity are the overriding factor, on CD-ROM(s) or DVD(s).

Delivery Media

1. Digital Format - Preferred
2. SVHS - Preferred
3. VHS – Backup media provided in conjunction with formats 1 or 2.

3.15 Synchronization

For deposition video, it is best to always get the synchronization done at the beginning with the delivery of the video. Synchronized text should appear black on a solid white background. The text should be clearly legible no matter the quality or color of the background video.

3.16 Transcripts

While there are a wide range of court reporters and companies in the marketplace, every professional should be able to provide their product in one of our accepted formats. Our goal is to minimize the amount of time required by our staff to get your product into our system for use by the legal team.

When printing any documents for the Firm, the vendor should **never** add their company information to the header or footer.

File Format

The Firm uses [Enter Application Name] for transcripts. The following formats are our accepted formats, in order of preferences (most to least):

1. _____;
2. _____;
3. _____;
4. _____;
5. _____;
6. _____

Many court reporters like to deliver an executable (.EXE). These are fine, so long as the “.exe” can export to a preferred format, as outlined above.

In today’s age of email viruses, many email systems automatically reject .EXE and .ZIP attachments as potential viruses. ASCII text files have the best chance of successful delivery.

Media (Delivery Format)

While transcripts are very small in file size, the floppy is slowly disappearing from the PC landscape. Today, a CD burner comes standard on the majority of PCs and laptops. An external CD burner costs ~\$50. A blank CD costs about \$0.10. Further, no one can accidentally modify or delete the court reporter’s work.

The following are our accepted formats, in order of preference (most to least):

1. CD or DVD;
2. 3.5” floppy disk;

Email is acceptable as a stopgap measure for emergencies or for a draft version before delivery of the final. All transcripts require a final CD, DVD or floppy.

Labels

The Firm has a vast library of transcripts. While every transcript is loaded into LiveNote / RealLegal / Summation, it is critical that the delivery media be easy to identify for inventory purposes. The Firm requests that the following information appear on all transcript media deliveries:

1. Vendor Name
2. Vendor Address
3. Vendor Phone
4. Names of Deponents
5. Dates of depositions
6. Times of depositions

Quality and Production Errors

Errors may result in outright rejection of product, reduction in payment due to internal technical time spent to reformat a transcript for import or a simple request for resubmission of product. The decision may be a factor of time and attorney direction.

3.17 Delivery Media

As of August 2004, one can buy a 200GB external USB2.0 hard drive for ~\$200. As most vendors charge an average of \$25 per CD, any delivery of 10 CDs or more should come on a hard drive. Aside from cost savings, loading from a hard drive saves time. It is much easier and expedient for Litigation Support to copy a single hard drive to the server than to copy 10 CDs.

Our firm prefers to receive productions greater than 10 CDs on an external hard drive with a USB2.0 connection.

4.00 Software Specific Requirements

As a result of coincidence, Concordance and Opticon are the two applications used for examples in earlier sections of this document. The following subsections identify how to create other load files which are in keeping with the organizational rules already outlined.

If your company would like to be included in this section, please contact Mark Lieb of Ad Litem Consulting, Inc.

4.01 Casesoft Suite

This section provided by CaseSoft.

CaseMap Load Files

1. A comma or tab-delimited text file with no supporting files
2. The first line of the text file should be the field names
3. One document record per line in the text file
4. The first field should have a value for every row. If necessary, this can be a sequentially numbered column that is skipped during the import.
5. No particular sort order is necessary. CaseMap will sort the records automatically.
6. Beginning and ending document numbers are often included, but not required.
7. The Bates – Begin field is a text field; so, Bates numbers should be padded with zeros in order to sort correctly. Example: 0001, 0010, 0100, 1000
8. Since many document databases use the beginning Bates number as the image ID, it is possible to set up links to images without importing any other information.
 - a. Set up the appropriate File Viewer for the document database in the target case and make it the default File Viewer.
 - b. From the Tools menu, choose Options and click to the Doc. Bates # tab. Check the option for “Use for default Linked File value” under Bates – Begin. (There is also an option here to populate Full Name and Short Name with the beginning number.)
 - c. With these settings, your imported documents will be linked to their respective images at the completion of the import.

To learn more about CaseSoft, please visit <http://www.casesoft.com>.

4.02 IPRO

This section provided by IPRO.

IPRO Tech, Inc. - LFP File Format

LFP files (also called load files) are used to build the image database and to instruct IPRO View how to display a project's images. The image's file location; file type, boundary, and Bates number make up the LFP file.

Here are three methods for creating an .LFP file.

1. The discovery vendor provides an LFP file upon completion of the imaging phase for a project / case.

2. Use IPRO's free utility, IConvert.

This is a free download at <http://www.lproCorp.com>.

The IConvert tool will convert many load file formats to either LFP or other software formats, such as Summation and Opticon. It is a handy tool for people who don't even use any other IPRO products internally.

3. Manually edit or create an LFP file using a text editor.

Each record (or line) in the LFP file begins with a 2-letter code that determines the action the viewer will perform. To load an image, we will start the line with IM.

Commas separate some parts of the record along with semi-colons for the rest of the line. The command in the LFP file takes effect after you load or import the LFP file. (Import > Import LFP File from the IPRO Tech Utility menu).

Here are two examples of an LFP record. In our examples, we use the MSC collection. It contains two documents, two pages each. The volume name is MSC001.

Example 1: Single Page .TIF files

```
IM,MSC00014,D,0,@MSC001;IMAGES\ 00\ 00;MSC00014.TIF;2
IM,MSC00015,,0,@MSC001;IMAGES\ 00\ 00;MSC00015.TIF;2
IM,MSC00016,D,0,@MSC001;IMAGES\ 00\ 00;MSC00016.TIF;2
IM,MSC00017,,0,@MSC001;IMAGES\ 00\ 00;MSC00017.TIF;2
```

Example 2: Multi-Page .TIF files

```
IM,MSC00014,D,1,@MSC001;IMAGES\ 00\ 00;MSC00014.TIF;2
IM,MSC00015,,2,@MSC001;IMAGES\ 00\ 00;MSC00014.TIF;2
IM,MSC00016,D,1,@MSC001;IMAGES\ 00\ 00;MSC00016.TIF;2
IM,MSC00017,,2,@MSC001;IMAGES\ 00\ 00;MSC00016.TIF;2
```

Note: Because the files are multi-page, the entire bates range (or image key range) must point to the same .TIF file. As example, MSC00014 contains both "14" and "15". Therefore, to view page 15, the computer must display MSC00014.TIF.

The following provides a breakdown of the fields:

Value	Description of Purpose
IM	Import code identifier (Importing New Page/Image database record)
MSC00014	The image key/document id number

D	Document designation; only designate the first page of each document.
0	Offset to the Tiff file. Always 0 for single page tiff files. When creating Multi-Page Tiff files, this number will increment for the pages within the file. (If there is an 11 page document, the offset would start at 1 and end at 11 and the next tiff file would start over at 1.
@MDEMO	CD volume name
IMAGES\00\00	Directory path on the CD for the image
MSC00014.TIF	Filename for the image.
;2	Tells IPRO the Types* of image file, e.g. tiff, PDF

*Supported Image Types and their specification in the LFP file are:

4. Type 1 is for IPRO Tech image from DOS-Based version, still supported (.IMG)
5. Type 2 is for Standard single and multiple page black & white or color TIFF (.TIF)
6. Type 3 is for IPRO Tech stacked TIFF (.STF)
7. Type 4 is for Color image (.BMP, .PCX, .JPEG or .PNG)
8. Type 5 is for black & white .PDF
9. Type 6 is for Color .PDF
10. Type 7 is to Auto-detect the .PDF type, e.g. Color or Black & White

To learn more about IPRO, please visit <http://www.iprocorp.com>.

4.03 Dataflight's Concordance and Opticon

This section was provided by Dataflight.

Concordance Database Load Files

The most reliable format for Concordance data delivery is a Concordance Database. This will ensure that the vendor has the correct fields, and that the data will load without a hitch. Rolling productions, delivered as Concordance databases, can be merged into the working set utilizing the standard "Import Concordance Database" option. Additional fields of data can be imported into existing records in the same fashion, allowing for initial base level coding to be done, and then more detailed coding for a subset of "Key" documents, identified through initial review. In instances where this is not an option, data can be delivered utilizing standard delimited files for coded data, and TXT or RTF files for OCR data. Refer to the load files section of this document to see the firm's preference.

Delimited Load Files

The first line of the delimited text database load file should be the field names.

Concordance allows users to specify delimiters; however, the best practice is to use the “Concordance Standard Delimiter” characters, which are:

- Comma (020),
- Quote (254),
- Newline (174)

OCR Load Files

OCR is loaded into Concordance through the READOCR CPL (Concordance Programming Language) script, which is designed to import document level OCR (one database record represents one document). Your text files should be on the document level to import properly with this CPL.

The choice of multi-page OCR files, or "Document level" files, means that the full document, including all pages, resides within a single file. If the database has five records, then there are five documents and five OCR text files, each containing however many pages. Most vendors will delineate between OCR pages by adding text such as, << ABC0000001 >>.

The OCR text filename must be unique. Otherwise the READOCR program may import that text into multiple records. The filename, therefore, should match the image key field for the associated document in the database (IMAGEKEY.TXT). The script will scan selected volume directories for the filename that matches the value of the “IMAGEKEY” field.

Example:

Two documents have been OCR’d for import into a Concordance database, with Bates ranges corresponding MSC000001 and contains 3 pages. The second begins at MSC000004 and contains 2 pages. The corresponding OCR text files are named MSC000001.TXT, and MSC000004.TXT.

BEGBATES*	ENDBATES	PATH	FILENAME
MSC000001	MSC000003	D:\[VOLUME_NAME]\OCR\	MSC000001.TXT
MSC000004	MSC000005	D:\[VOLUME_NAME]\OCR\	MSC000004.TXT

* *Image key* - unique value.

Opticon OPT (Load) Files

The Opticon load file details the link between documents in Concordance and their corresponding images. Each line reference defines the image key (the reference from the database), its volume label (for identification purposes), and the associated image (with its full file path). The load file entries also define the document breaks and, optionally, page counts.

The Opticon load file format is a text-delimited file containing all information necessary to link the imagebase with the database. There is one line entry per image file, whether it is a single-page or multi-page image file. The load file consists of seven delimited entries as follows:

ALIAS,VOLUME,PATH,DOC_BREAK,FOLDER_BREAK,BOX_BREAK,PAGES

Example:

The following is a 5-image load file example. It details 2 documents; the first relates to the image key MSC000001 and contains 3 pages. The second begins at MSC000004 and contains 2 pages.

```
MSC000001,MSC001,D:\IMAGES\001\MSC000001.TIF,Y,,,3
MSC000002,MSC001,D:\IMAGES\001\MSC000002.TIF,,,,
MSC000003,MSC001,D:\IMAGES\001\MSC000003.TIF,,,,
MSC000004,MSC001,D:\IMAGES\001\MSC000004.TIF,Y,,,2
MSC000005,MSC001,D:\IMAGES\001\MSC000005.TIF,,,,
```

Value	Description
ALIAS	Should match your image key from the Concordance database. Concordance stores this key in order to reference the image.
VOLUME	This entry is the name of the volume where the image resides. This is typically the volume name of a CD or server. (Optional)
PATH	This is the full path and file name (and extension) of the image.
DOC_PATH	Enter a 'Y' to denote whether this image marks the beginning of a document.
FOLDER_BREAK	Enter a 'Y' to denote whether this image marks the beginning of a folder. (Optional)
BOX_BREAK	Enter a 'Y' to denote whether this image marks the beginning of a box. (Not Currently Supported)
PAGES	This entry is the number of pages associated with the image. (Optional)

Opticon currently supports the following image types:

- TIFF files: (single and multi-page): (.TIF)
- JPEG files (.JPG)
- GIF files (.GIF)
- Bitmap files (.BMP)
- PCX files (.PCX)
- CALS files (.CAL, .MIL)

To learn more about Concordance and Opticon, please visit <http://www.dataflight.com>.

4.04 Image Capture Engineering

This document does not yet provide any specifics for this software. If you would like to submit content, please contact Ad Litem.

To learn more about ICE, please visit <http://www.imagecap.com>.

4.05 RealLegal

This document does not yet provide any specifics for this software. If you would like to submit content, please contact Ad Litem.

To learn more about RealLegal, please visit <http://www.reallegal.com>.

4.06 LiveNote

This document does not yet provide any specifics for this software. If you would like to submit content, please contact Ad Litem.

To learn more about LiveNote, please visit <http://www.livenote.com>.

4.07 Summation

This document does not yet provide any specifics for this software. If you would like to submit content, please contact Ad Litem.

To learn more about Summation, please visit <http://www.summation.com>.

4.08 iCONNECT

This document does not yet provide any specifics for this software. If you would like to submit content, please contact Ad Litem.

To learn more about iCONNECT, please visit <http://www.iCONNECT.com>.

4.09 Trial Director

This document does not yet provide any specifics for this software. If you would like to submit content, please contact Ad Litem.

To learn more about Trial Director, please visit <http://www.indatacorp.com>.

4.10 Sanction2

This document does not yet provide any specifics for this software. If you would like to submit content, please contact Ad Litem.

To learn more about Sanction2, please visit <http://www.verdictsystems.com>.

4.11 JFS Litigator's Notebook

This document does not yet provide any specifics for this software. If you would like to submit content, please contact Ad Litem.

To learn more about JFS Litigation's Notebook, please visit <http://www.bowne.com>.

4.12 This Space Can Be Yours

Simply contact Ad Litem Consulting and submit your load file or other specifications.

5.00 Examples of What Not To Do

The following are examples of bad product. Any of these types of “fouls” will require correction before Firm legal staff can begin their work. Corrections may be performed by Litigation Support or the vendor, as need and time restrictions dictate. All examples are real and taken from deliveries we have received.

5.01 Media Labels

1. Misspellings: Client names, matter names or client-matter numbers are misspelled. This can be especially frustrating when the Firm’s client wishes to see or get a copy of the vendor product. If our client’s name is SMITH, a CD labeled SITH or SMATH or SMYTHE instantly calls into question the quality of the content of the media.
2. Handwritten: Not only is handwriting hard to read, but it also lends itself to missing information.

5.02 File / Folder / Volume Name Conventions

1. Tilde or otherwise truncated file or folder names. As example, AAA0000001.TIF versus AAA000~1.TIF and D:\PROGRAM FILES\ versus D:\PROGRA~1\. Whenever possible, volume, file and folder names should not be wider than eight (8) characters with a suffix not wider than three (3) characters. In technical circles, this is known as the “16-bit” or “MS-DOS 8.3” naming convention.
2. Use of spaces or any characters in a load file that Windows does not allow in a file or folder name. This seems obvious, but we have received deliveries from vendors who used characters in the database that were not valid in the filename. This resulted in files that would either not copy to the server or would copy with strange naming results. We don’t know what kind of software or operating system this vendor used to create their product, but they certainly never tried to load it themselves.
3. I will not name the vendor, but on several separate projects they used VENDOR001 as the volume name. If used, my server could have a dozen VENDOR001 CDs. If the filenames do not have the full image key, e.g. “0000001.TIF”, there is no quick way to determine the associated database, client or matter.
4. Missing or empty folders are a big red flag. If your image folder contains 3 subfolders named 012, 014 and 015, your first inclination is to ask what happened to 001-011 and 013. This just isn’t right on any level. It leads me to assume there will be problems with the data, specifically missing images.

5.03 Database

1. The date field should only include the date. An example of a valid date is “01/01/2004”. An example of an invalid date is “01/01/2004 12:01:01PM”.

2. Dates should have 4-digit years. "01/01/2004" is valid where "01/01/04" is not.
3. OCR and full text from electronic discovery should maintain original formatting. Some EDD and OCR applications replace spaces, soft returns and hard returns with characters other than spaces, soft returns and hard returns. If the original text is "Best Practices", then the database OCR field should never contain: "BestPractices", "Best/Practices" or "Best@Practices".

5.04 Media Content

1. Each CD should be "self-contained". If 5 CDs arrive and the load files for all 5 CDs reside on CD #5, then that is wrong. The idea here is to be able to reload any CD as quickly as possible. Sometimes collections become separated over time. It is conceivable that CD #5, with all the database and image load files, could be lost. This means CD #1 through #4 are now incomplete. Each CD should be self-contained.
2. For a given project, all load files (Concordance .DAT and Opticon .LOG) should use the same field names, ordering and structure as the first delivery.

5.05 Load Files

Concordance Load Files

1. A .DAT load file without a supporting file showing: field structure, field size and field sequence.
2. The first line of the .DAT file should be the field names. When loading a .DAT file, this is the simplest way to see if the data loaded correctly.
3. Badly formatted body Meta-Data. The spaces and returns must match the original text. No odd characters, such as a semi-colon, should appear in lieu of a soft-return or a space. These kinds of problems not only make the text hard to read, but they also interfere with searching.
4. More than one document per database record. This kind of error can cost the Firm hours and days or a case. When the review team identifies all the documents to produce, a ratio other than 1:1 will result in the wrong documents getting produced along with the right documents.
5. Databases and load files should open sorted by "Bates" or "docno". Concordance displays records in the same order that they were loaded. Therefore a disordered load file results in a disordered database.
6. Duplicate, overlapping or gaps in "Bates" or "docno" fields.
7. Bates / Docno prefix contains characters other than A...Z.
8. Bates / Docno suffix contains letters and is not zero-added to four places (.0001).
9. Bates / Docno contains a space, such as "AA 00001".

5.06 OCR

1. When there is bad OCR, an appropriate error code and warning to the firm is required. Things such as handwriting and graphics will not provide good OCR results. As such the vendor must warn the firm and Litigation Support about these issues and the associated "<<OCR ERROR>>" text. In this fashion, the law firm knows a legitimate error from a missed problem. This can result in a "false positive" in terms of QC looking for errors.

2. Vendor must use **Auto-Rotate** on every image. This ensures the 5 – 10% of images facing sideways or upside-down get quality OCR. Documents such as hierarchical employee charts are almost always designed landscape instead of portrait. All of these names and titles should be easy to OCR, unless auto-rotate is off.

5.07 Opticon Load Files

1. The following is based on an actual subfolder name we received on a production by a vendor:

\\BOX 3 - JOHN DOE & OTHERS DOCS REC'D FROM BOB SMITH; AL SMITH'S NOTEBOOKS, PAPERS\

2. There are several big fouls here. In addition to the “&”, “;” and “'” (apostrophe) characters, there really is no purpose in a “significant folder name”. This was part of a series of subfolders that literally went eight (8) folders deep. Not only will the end user never see the folder name when reviewing their discovery, but one cannot load the cross-reference file into Opticon. Opticon requires a comma delimited load file.
3. The example below shows problems we have seen historically:
 - A. Image key, “A001” and filename “001.TIF” do not match
A001,[VOLUME],D:\[VOLUME]\IMAGES\001\001.TIF,Y,,,
 - B. This first page of a document is missing a page count
A001,[VOLUME],D:\[VOLUME]\IMAGES\001\001.TIF,Y,,,
 - C. This page is missing the “;”s and possibly the begin document “Y” and page count
A002,[VOLUME],D:\[VOLUME]\IMAGES\001\A002.TIF
4. Opticon load file extensions should be .LOG, .TXT or .RXF. Some software vendors used to create the log file output with an extension of .OPT. Opticon does not look for .OPT when displaying potential load files.
5. Image Cross-Reference File – Filename Mismatch. The filename inside of the cross-reference file does not match the actual filename. Again, this could be a hiccup in processing. This is caught when we run our QC tests to make sure every file listed is actually on the server.
6. Only images belong in the Opticon load file. Sometimes vendors will put the OCR files into the same folder as the images. This has, on occasion, resulted in a load file that references both the images and the OCR files. In the following example, lines 2 and 4 should not be included:

1. SMI0001,SMI001,D:\IMAGES\SMI0001.tif,Y,,1
2. SMI0001.TXT,SMI001,D:\IMAGES\SMI0001.TXT,,,,
3. SMI0002,SMI001,D:\IMAGES\SMI0002.tif,Y,,1
4. SMI0002.TXT,SMI001,D:\IMAGES\SMI0002.TXT,,,,

Every import line for every delivery should be formatted the same, irrespective of the technician who generated the load file. Right or wrong, at least let the whole delivery be consistently from CD to CD. If the path information isn't “plug and play”, Litigation Support

has to modify the associated load files. Did the vendor not know or not care that their CDs contained inconsistent information?

5.08 Image Format

1. Multi-Page TIFFs. There are two major problems with multi-page TIFFs. The main issue is the inability to easily divide one document into two. Selecting the “logical bindings” option in scanning along with use of slip-sheets is a great way to ensure the required one document to one record division in the database.
2. Unless otherwise specified, we do not want Bates stamps or any other type of stamp applied to our images.
3. TIFF images of Excel spreadsheets where columns are too narrow causing cell content to appear as “#####” instead of the actual value.
4. TIFF images of Excel spreadsheets where the cells show the formula instead of the resulting value. An example of this would be a summing cell that should show the grand total for a column but instead shows something such as “=sum(A1..A10)”.

5.09 Transcripts

1. Transcript is in WordPerfect format or some legacy word processing format such as Wang or Wordstar.
2. Transcript requires manual editing due to extremely irregular formatting.
3. Gaps in text or pages.
4. Control characters in transcript text file.
5. Each line of text has a “line wrap” instead of “hard return”. (Note: UltraEdit, text editor can fix this.)
6. Delivery transcripts on floppy instead of CD. (It is safer and the media cheaper.)

5.10 General Errors / Issues

1. Databases and Opticon load files where every document is one page. While possible and quite likely to have 1 single page document, a database comprised entirely of 13,000 one-page documents is highly unlikely.
2. While a document containing 13,000 pages is possible, is it unlikely. A database with several 13,000 page documents is extremely unlikely. This could be a physical versus logical document breaks issue.
3. Do not create a new image sub-folder for each document. A CD with 300 1-page documents should result in 1 folder of 300 images. 300 folders each containing 1 page is incorrect.
4. When generating electronic or paper documents, the vendor should never add their company information to the header or footer.
5. Add your own real life war stories to the growing list. Visit <http://www.eDiscovery.org> today.

5.11 Real Experiences

All of the examples in the "[things not to do](#)" section are, unfortunately, real. These types of problems eat up a lot of billable time. As the document matures, the list will probably get longer. There are a lot of creative people out there working the controls. What they create is what your firm will use for review, productions and exhibits.

People are welcome to submit their experiences at <http://www.eDiscovery.org>.

These examples are not meant to criticize any product or person. These examples explain the types of problems one may encounter, by product, regardless of product, firm, vendor or persons involved. One goal of this document is to give both good and bad examples that anyone technical person can use to improve their products.

If you use this document, please let me know.

Glossary

The following is a partial glossary of terms. For an expanded industry dictionary, please visit <http://www.DiscoveryDictionary.com>.

About Mark Lieb

MARK R. LIEB | Ad Litem Consulting, Inc.

mark@adlitem.com

PO 124873
San Diego, CA 92112-4873

Ph 619-546-7750



Mr. Lieb has provided litigation support to legal teams for cases ranging from small collections to multi-national, multi-firm, litigation involving millions of pages of discovery. Mr. Lieb currently consults with firms and companies on litigation technology and marketing.

"Mr. Lieb's expertise focuses on the full range of modern electronic discovery. He has been instrumental in creating a robust infrastructure for providing electronic discovery services within the firm. In addition, he was the principal author of Gray Cary's Litigation Support Technical Standards. This document has done much to enhance communications between vendors and end users".

- Browning Marean, III, Partner, [DLA Piper Rudnick Gray Cary](#)

In March of 2004, Gray Cary generously made the document public in an attempt to help the industry. Mr. Lieb continues to update it. As of November, 2004, the document has been downloaded over 700 times by a wide range of vendors and law firms.

Mr. Lieb is the co-founder and technical hand behind the Litigation Support Vendors Association. The LSVA, a not-for-profit organization, provides free forums moderated by industry experts and representatives from major software companies covering topics from computer forensics and electronic discovery to best practices and standards. The website, [LSVA.com](#), is the public home of the Litigation Support Standards document.

Mr. Lieb's most recent efforts center on his book, Litigation Support Department. The documentation outlines how to run the litigation support as a business. The litigation support department can provide a consistent product and experience for the legal team and associated parties. The manual includes case lifecycle, work flow, budget, strategies, check lists, roles, responsibilities and more. Firms that outsource all work will also benefit. The document is available as a free download from <http://litigation-support.info>.

It is easier to explain litigation case lifecycle to a technician than to explain technology to a "layperson". Litigation support now has dozens of pages on technology, lifecycle, strategies and tactics. Put this knowledge to use for you.

Consulting for Firms

Whether your firm has no litigation support department or a series of decentralized departments, outsources all discovery or performs work in-house, the firm needs a technology plan for each case and an overall strategy for the firm. After calling many major firms it became apparent that creating documentation and best-practices is a high priority. Only, no one has the time to write it.

Litigation Support can help attorneys expedite review and keep costs down. At the same time the firm uses the accounting department to track the financial impact. Ad Litem Consulting can work with your firm to put best practices ideas into action.

A good technology plan works for cases of any size. Because all attorneys encounter the same issues in the same phases of litigation, the technological and strategic issues are the same. Further, execution of that plan can benefit the firm in addition to each case. This is true whether your firm outsources litigation goods and services or performs them in-house.

Mr. Lieb has worked closely with dozens of attorneys and paralegals on hundreds of databases. Drawing upon his graduate work in computer science and over a dozen years field experience in information systems, Lieb created an efficient model for handling cases as well as the department. Mr. Lieb shows the law firm how to use this standardized approach for all litigation.

Consulting for Vendors

Train your sales and technical staff on the law firm mentality. Learn how to make litigation support your internal cheerleader. In today's competitive market, any big problems can put your shop "on the bench" for years. There are too many other vendors vying for a chance to provide their worth.

What are the best ways to get that first project at a firm? How can you prove your worth while impressing the right people? Mr. Lieb's grasp of guerilla marketing is well known in the litigation support community. Put this knowledge to work for your company.

If litigation support will administrate, support or use your product, should you cater functionality, workflow and interface design for this role? Improve your product by designing for actual business needs with input from a litigation support expert.

Litigation Support Department - The Book

Read about the litigation support department from the perspective of the litigation support technician and manager. Learn how to use technology and strategy to benefit cases, improve review speed and avoid expensive dead-ends.

This is your chance to get everyone on the same page. If your firm has a department of one or more people, litigation support can not only pay for its own hardware, software and personnel needs, it should. Learn how to make the business case the partners will embrace. If your firm outsources all litigation technology goods and services, this book will help you choose a good vendor, keep costs down and understand how technology helps you to litigate your case.

If you are a vendor, this book can provide valuable insight into the business workings of your clients. Learn how your goods and services compliment the case lifecycle from the firm perspective. Learn what happens after you deliver your product.

The litigation support department and litigation technology inside of the firm have changed since the explosion of electronic discovery. Those vendors approaching electronic discovery from a photocopy shop mindset are headed for trouble.

This book provides a tour of the new mindset and case strategies inside of the firm. It also represents several hundred hours of writing, years of hands-on litigation support experience and lots of formal training.

Bulk and educational discounts are available.