ACC AMERICA
Association of Corporate Counsel

ACC Association of Corporate Counsel

704:Document Retention & e-Discovery in a Post-Enron/Andersen World

Bradley W. Jordan
Principal
Jordan Lawrence Group, LLC

Kathie S. Lee
Vice President, Legal
Cendant Corporation

Carl D. Liggio
General Counsel
EthicsPoint, Inc.

Faculty Biographies

Bradley W. Jordan

Bradley W. Jordan is principal at Jordan Lawrence Group, L.C. and is an industry-noted expert on top breakthroughs that enable large, complex companies to establish and enforce legally defendable corporate records programs covering all records platforms. Using a large dose of objective experience and common sense, and an even larger dose of sophisticated technology, he advises top U.S. executives on how to cut the legal risks and operating costs related to poorly designed and unenforced corporate records programs. His company has worked with more than 800 major corporations including JPMorganChase, Cingular Wireless, Duke Energy, Oracle, Maritz Corporation, Fujitsu, Weight Watchers, Applica, and many others.

Mr. Jordan has more than twenty years experience as an information management executive. His first industry experiences were at Eastman Kodak's business information systems group and later with Anacomp.

Mr. Jordan received his undergraduate degree from Indiana University and a MBA from Maryville University in St. Louis.

Kathie S. Lee

Kathie S. Lee is currently vice president, legal for Cendant Corporation. She provides counsel and advice to Cendant Corporation's real estate franchise group regarding the franchise relationship. This includes advice on a variety of compliance issues, general commercial issues, insurance issues, bankruptcies, and management of litigation matters.

Prior to joining Cendant Corporation, Ms. Lee was assistant counsel to Summit Bancorp where she was involved in negotiating and drafting contracts, settlement, and workout agreements related to commercial real estate transactions. Prior to Summit Bancorp, Ms. Lee was the law clerk to the Honorable John J. Connelly and the Honorable Dennis D. O'Brien, United States Bankruptcy Judges.

Ms. Lee is secretary for ACCA's Litigation Committee. She is also a member of the Women in Federal Practice, National Asian Pacific American Bar Association, and the ABA.
Ms. Lee received her BA from the University of California at Los Angeles and a JD from the Benjamin N. Cardozo School of Law.

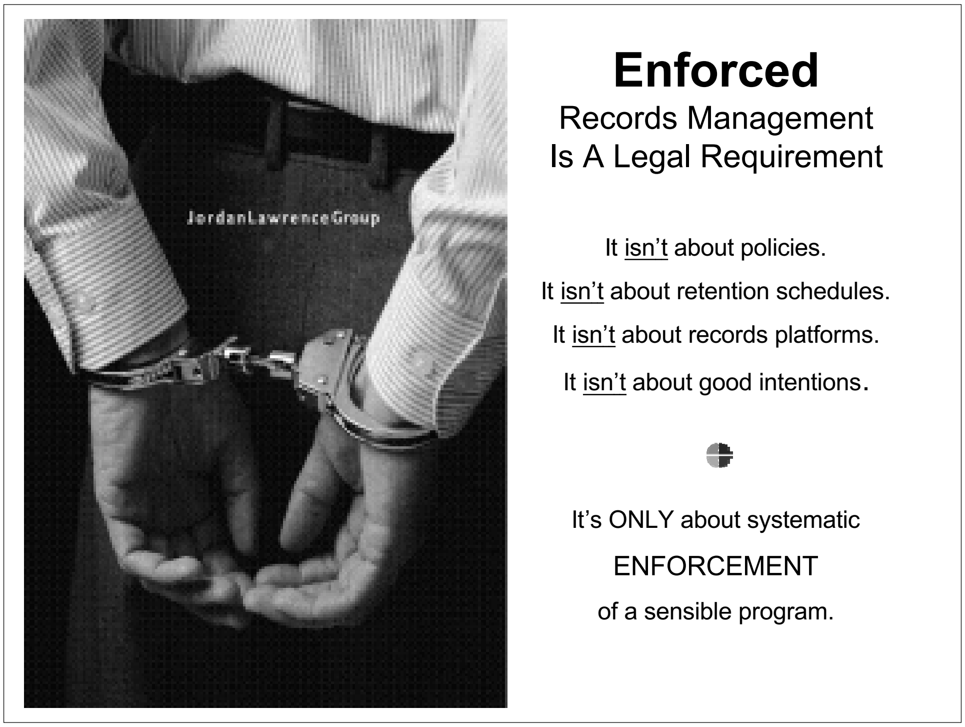
Carl D. Liggio

Carl D. Liggio currently serves as general counsel and director of Ethicspoint, a company which has a web based reporting system for Sarbanes Oxley 301(4) compliance. He is also a director of FIOS, Inc., a company that provides electronic discovery services to litigators and corporate clients.

He spent most of his career as general counsel of Ernst & Young and its predecessor, Arthur Young & Company. He has also served as managing partner of the Chicago office of Dickson, Wright, as of counsel to the law firm of McCullough Campbell & Lane in Chicago, and as a trial attorney with White & Case.

Mr. Liggio is a founding member of ACCA, served on its board of directors for 13 years, and is a past chair. He has written and lectured extensively on the subjects of professional responsibilities, securities law, management of the law, the use of technology in a legal practice, and accounting and legal issues.

Mr. Liggio received his undergraduate degree from Georgetown University and attended New York University School of Law from which he graduated *cum laude.*



Enforced

Records Management
Is A Legal Requirement

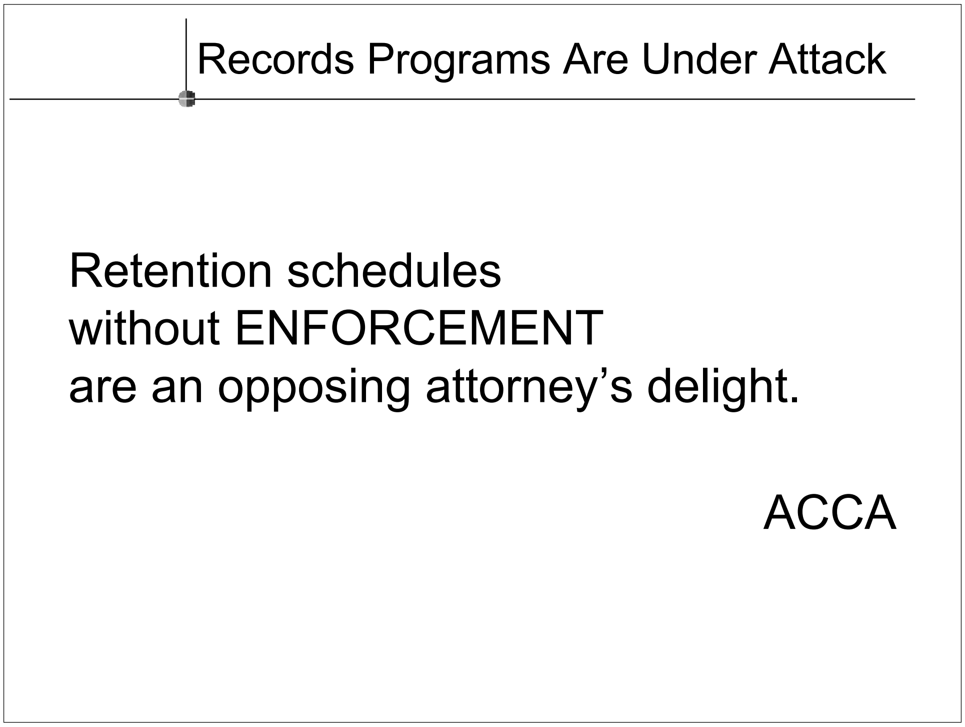
It isn't about policies.

It isn't about retention schedules.

It isn't about records platforms.

It isn't about good intentions.

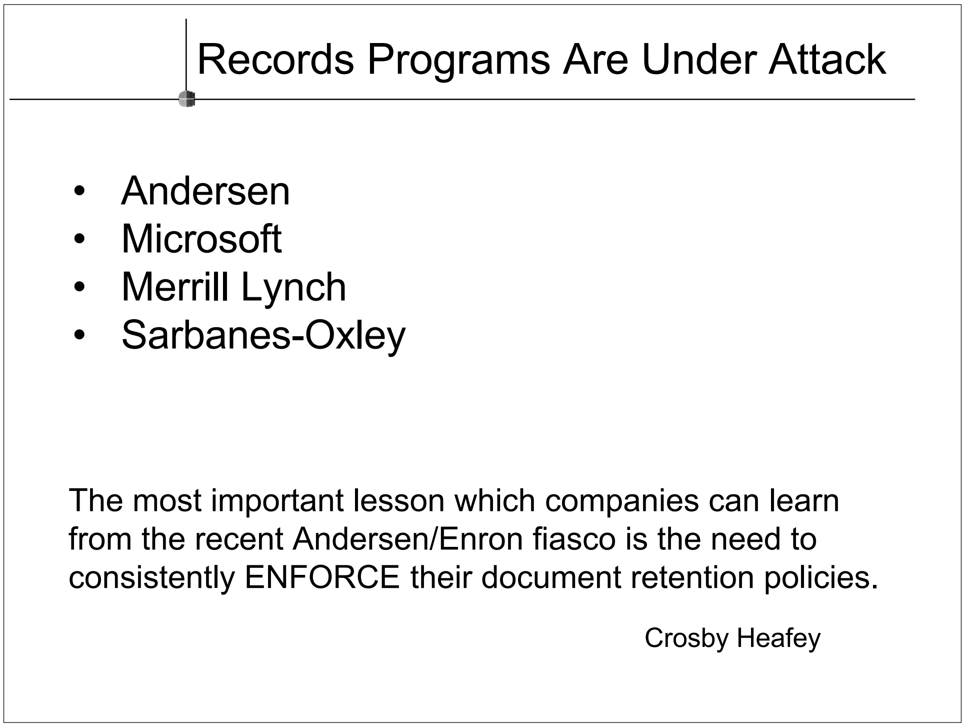
It's ONLY about systematic
ENFORCEMENT
of a sensible program.



Records Programs Are Under Attack

Retention schedules
without ENFORCEMENT
are an opposing attorney's delight.

ACCA

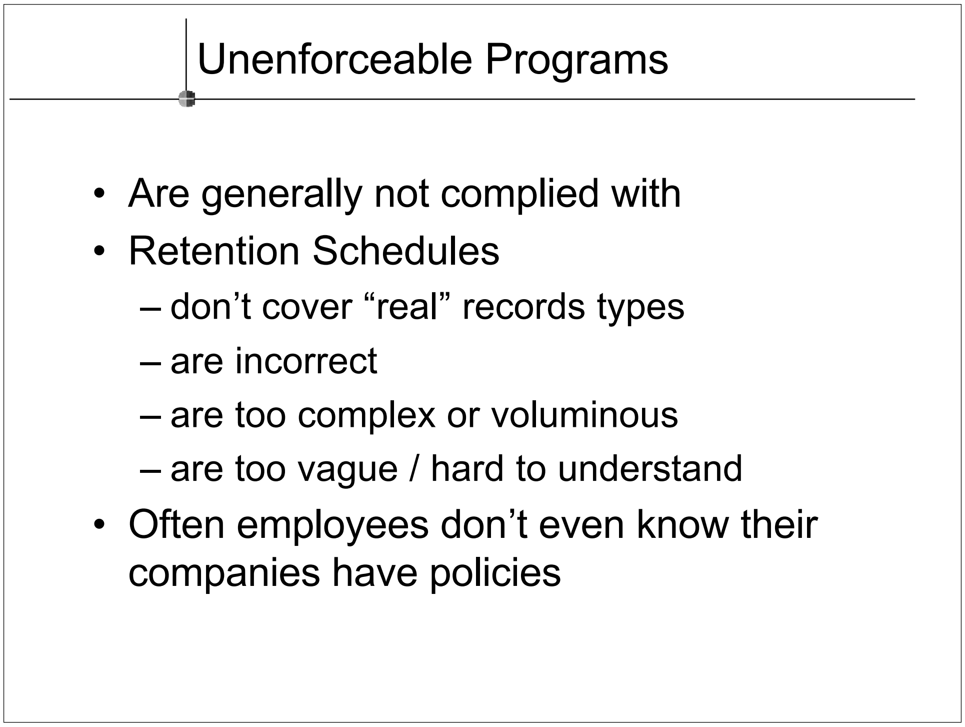


Records Programs Are Under Attack

- Andersen
- Microsoft
- Merrill Lynch
- Sarbanes-Oxley

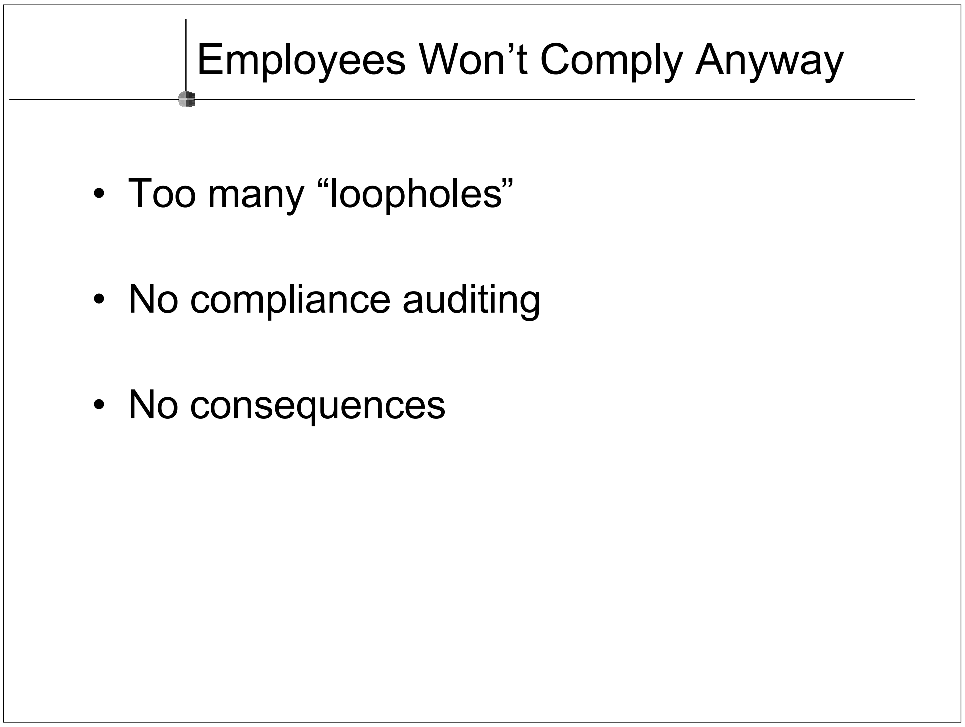
The most important lesson which companies can learn from the recent Andersen/Enron fiasco is the need to consistently ENFORCE their document retention policies.

Crosby Heafey



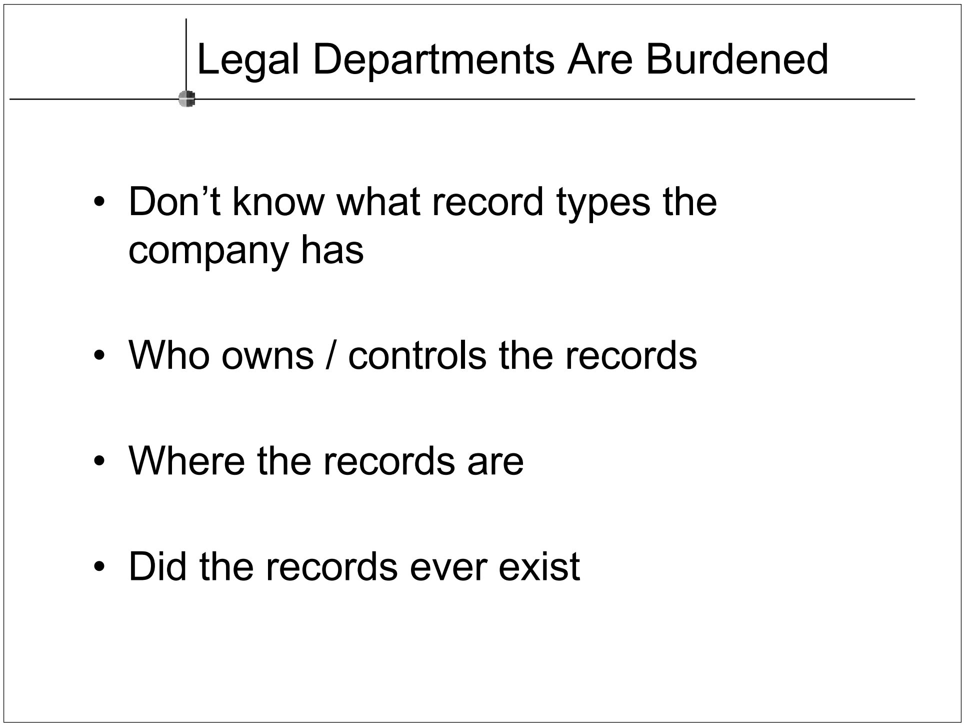
Unenforceable Programs

- Are generally not complied with
- Retention Schedules
  - don't cover "real" records types
  - are incorrect
  - are too complex or voluminous
  - are too vague / hard to understand
- Often employees don't even know their companies have policies



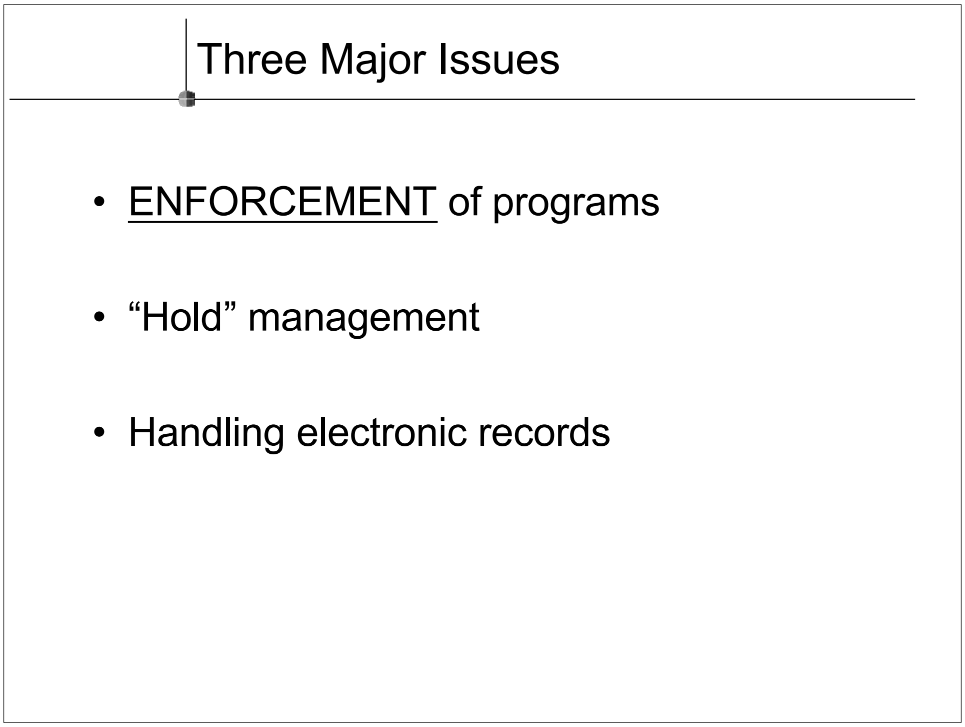
Employees Won't Comply Anyway

- Too many "loopholes"
- No compliance auditing
- No consequences



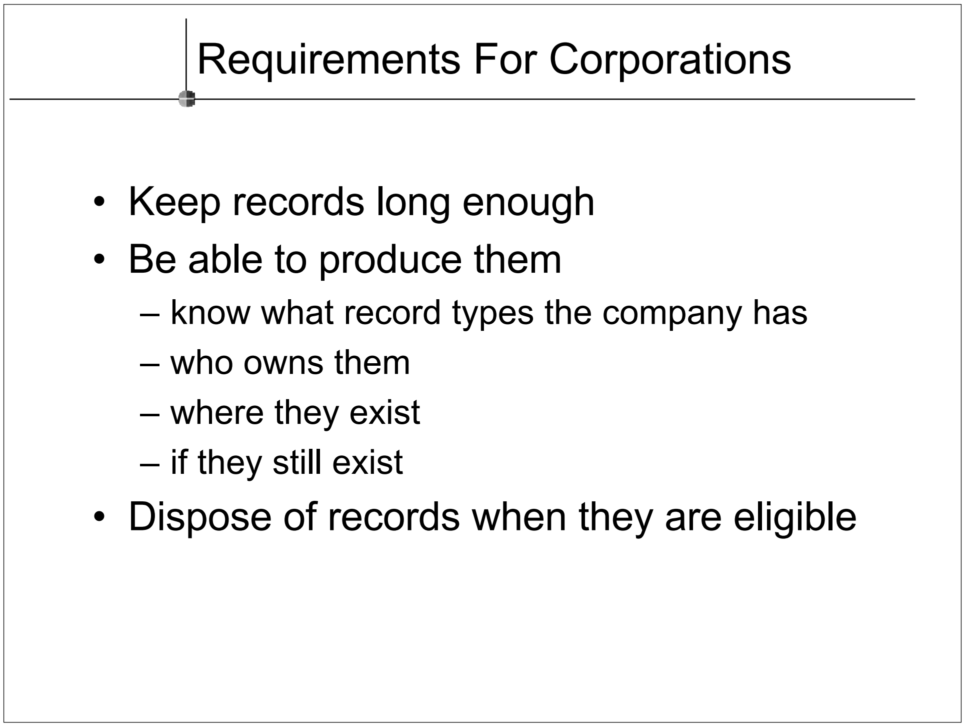
Legal Departments Are Burdened

- Don't know what record types the company has
- Who owns / controls the records
- Where the records are
- Did the records ever exist



Three Major Issues

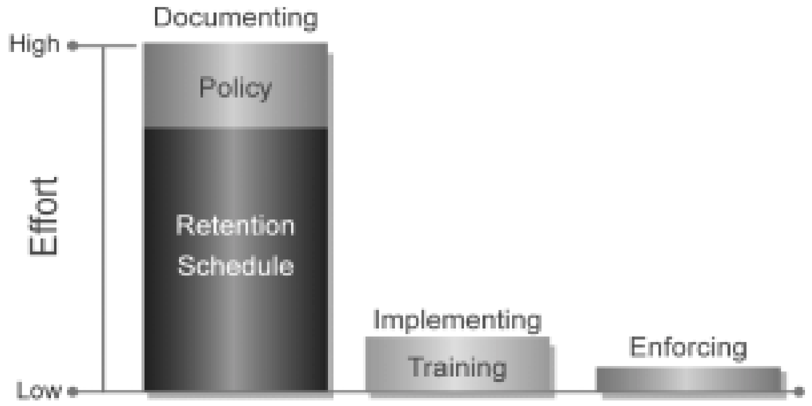
- ENFORCEMENT of programs
- "Hold" management
- Handling electronic records



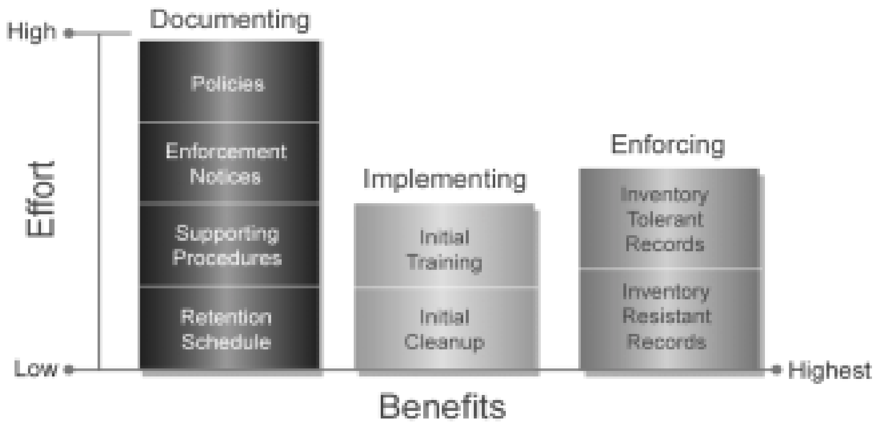
Requirements For Corporations

- Keep records long enough
- Be able to produce them
  - know what record types the company has
  - who owns them
  - where they exist
  - if they still exist
- Dispose of records when they are eligible

Policies Provide No Protection



Enforcement Is Most Important



There Are Two Types Of Records

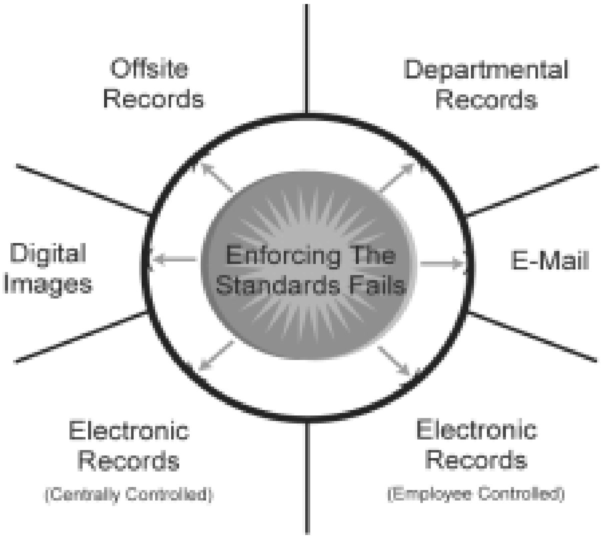
Inventory Tolerant Records

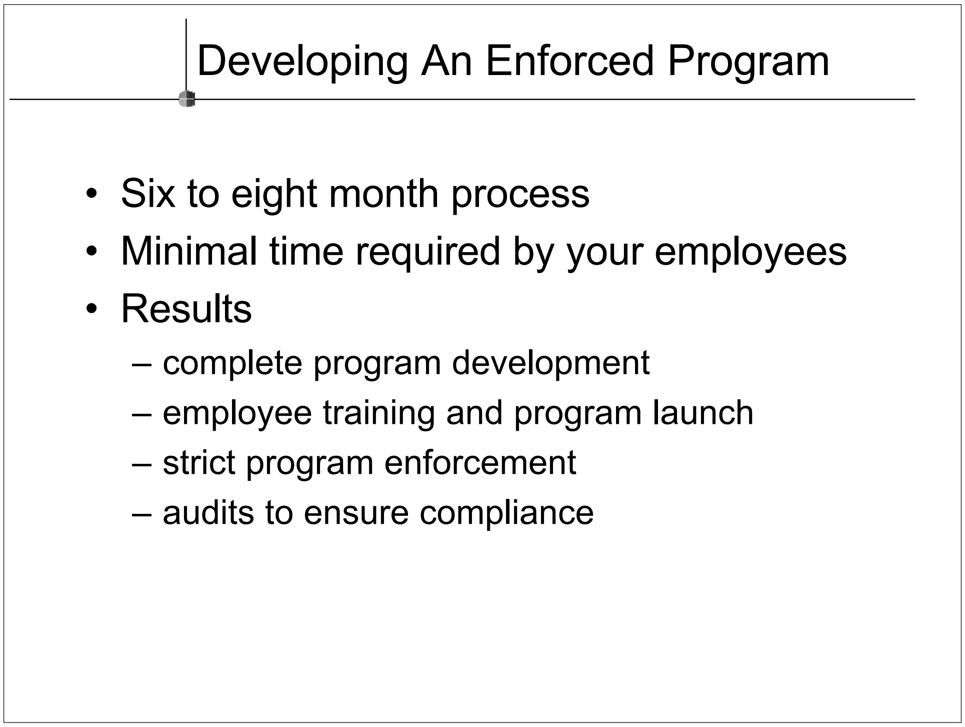
- Trackable records
- Easy to enforce

Inventory Resistant Records

- Employee controlled
- Difficult to enforce
- Emphasize user responsibility

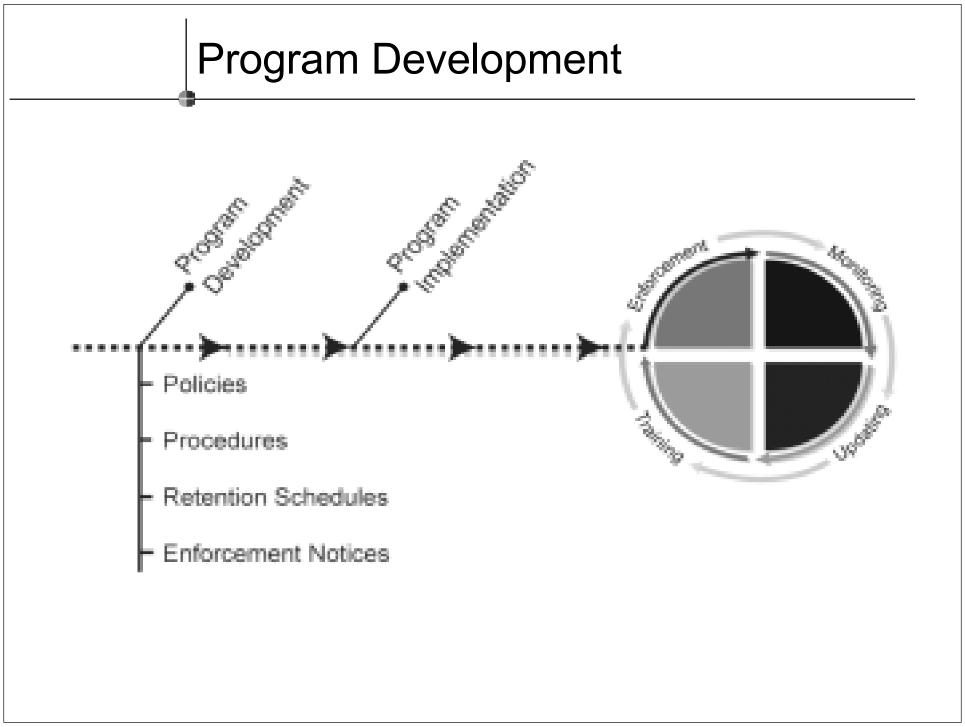
Platforms Fragment Enforcement


Offsite Records
Departmental Records
Digital Images
Enforcing The Standards Fails
E-Mail
Electronic Records (Centrally Controlled)
Electronic Records (Employee Controlled)



Developing An Enforced Program

- Six to eight month process
- Minimal time required by your employees
- Results
  - complete program development
  - employee training and program launch
  - strict program enforcement
  - audits to ensure compliance



Program Development

Program Development

Program Implementation

- Policies
- Procedures
- Retention Schedules
- Enforcement Notices

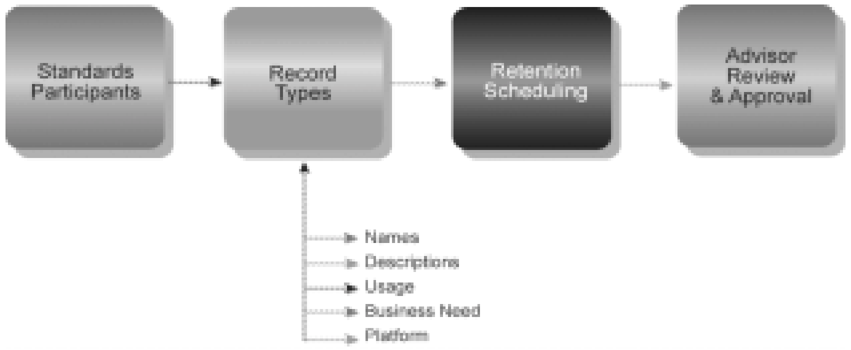
Enforcement

Monitoring

Updating

Training

Development – Standards Setting



Standards Participants

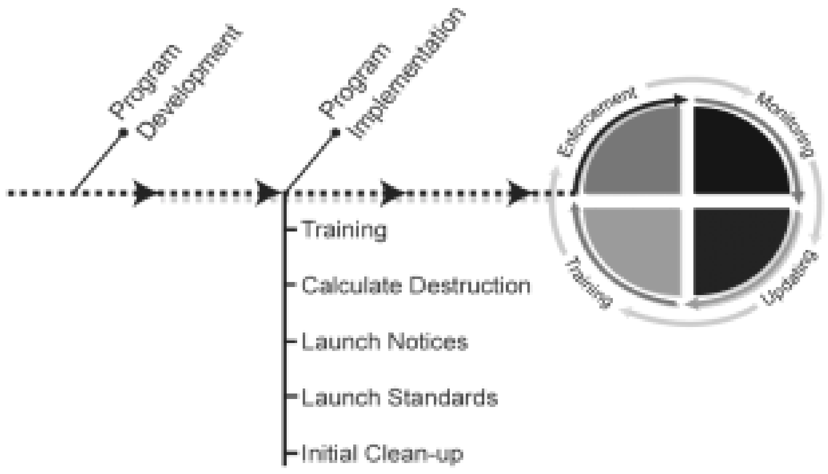
Record Types

Retention Scheduling

Advisor Review & Approval

- Names
- Descriptions
- Usage
- Business Need
- Platform

Program Implementation



Program Development

Program Implementation

- Training
- Calculate Destruction
- Launch Notices
- Launch Standards
- Initial Clean-up

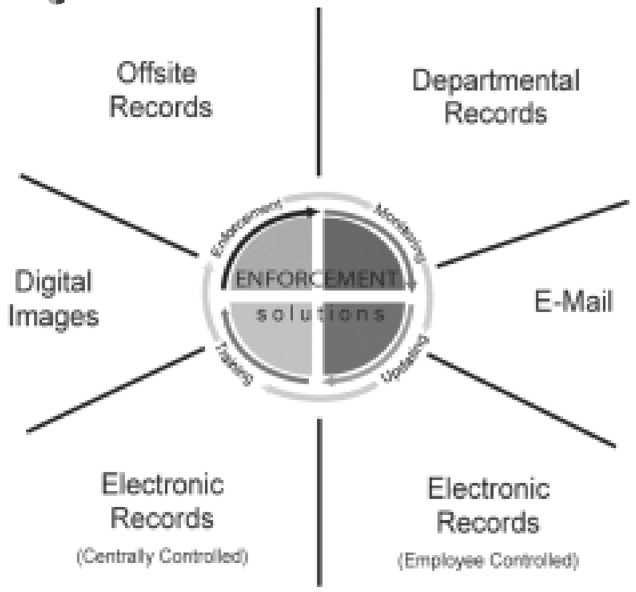
Enforcement

Monitoring

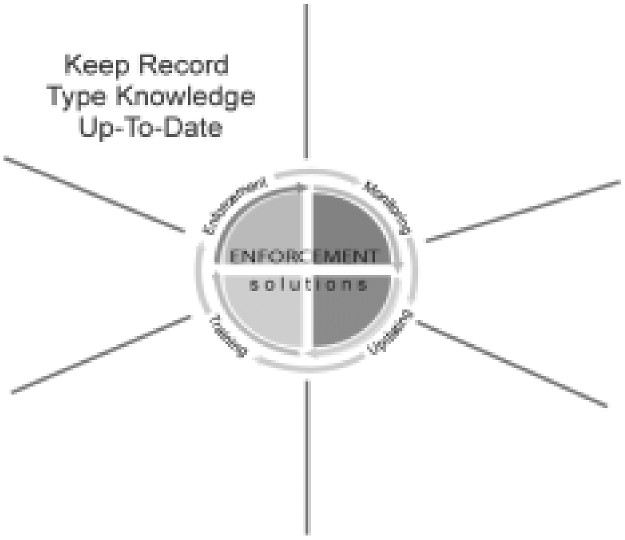
Updating

Training

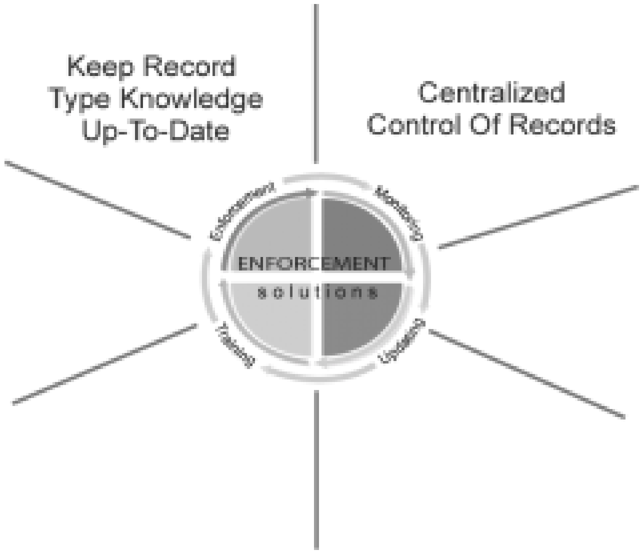
Program Enforcement Hub



Program Enforcement



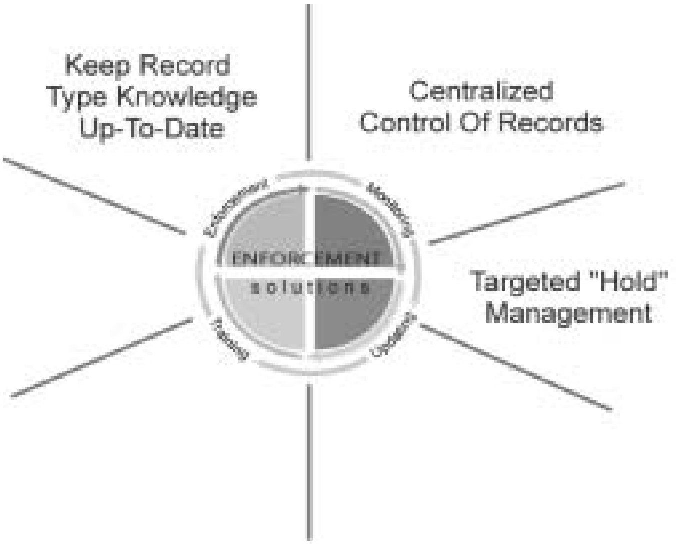
Program Enforcement



Keep Record Type Knowledge Up-To-Date

Centralized Control Of Records

Program Enforcement

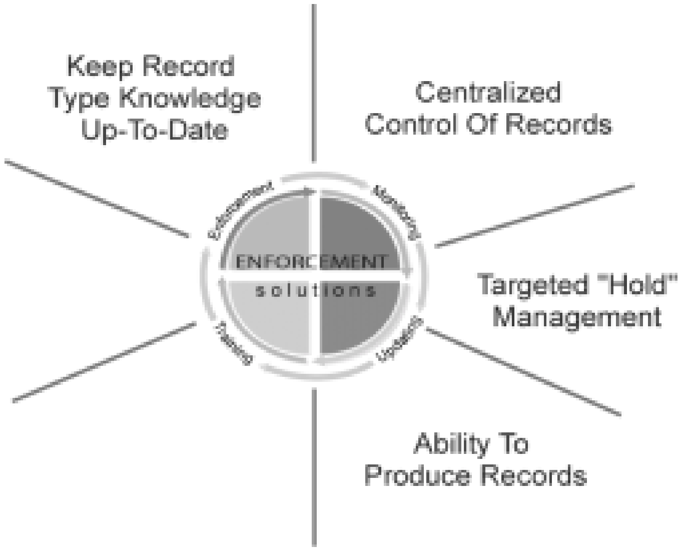


Keep Record Type Knowledge Up-To-Date

Centralized Control Of Records

Targeted "Hold" Management

Program Enforcement



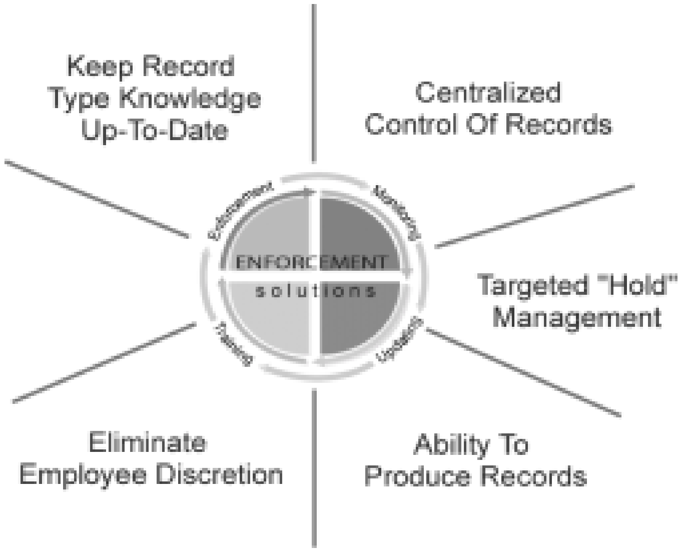
Keep Record Type Knowledge Up-To-Date

Centralized Control Of Records

Targeted "Hold" Management

Ability To Produce Records

Program Enforcement



Keep Record Type Knowledge Up-To-Date

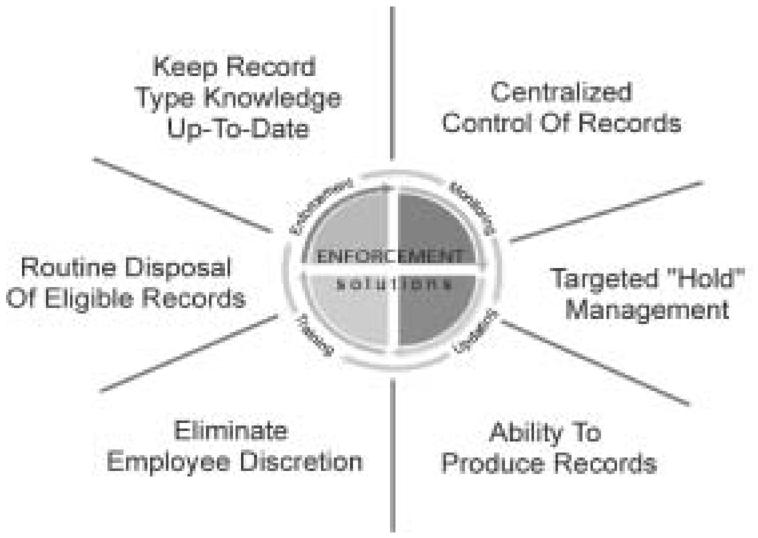
Centralized Control Of Records

Targeted "Hold" Management

Eliminate Employee Discretion

Ability To Produce Records

Program Enforcement



Keep Record Type Knowledge Up-To-Date

Centralized Control Of Records

Routine Disposal Of Eligible Records

Targeted "Hold" Management

Eliminate Employee Discretion

Ability To Produce Records

Benefits Of Enforcement

- Reducing legal risks
- Ease of operations
  - Ability to find all the records you need, when you need them
  - Comfort that you are fully complying with discovery; that availability of records is fully defendable
- Reducing costs
  - Increased efficiency in research
  - Lower storage volume

Document Retention – the Flip Side to e-Discovery or is it the other way around

Carl D. Liggio, Sr.[**]

I. In the past four decades, society has undergone a dramatic change in the amount of information that is available and in information technology that manages, retrieves, and stores that information.

A. The volume of information. A study done in 1995 found that the total amount of information/knowledge that was created in the prior thirty (30) years (1965 to 1995) was equal to the amount of recorded information that had been generated in prior 3 millennia.

1. It has been predicated that this volume will continue to double every five years thereafter.

2. Apparently, if the entire wealth of knowledge (and print media) that existed at the start of the Renaissance were captured in one place today, it would not fill a week day edition of The New York Times.

B. The over whelming majority of this data is now electronically created, captured, managed, used, and stored.

1. Until the advent of the computer, we were a paper oriented society. Our paper orientation was reinforced in the early 60s with the generation of the dry copier by Xerox.

2. The development of computers, particularly the laptops began a migration of this information from paper to digital: the "electronic media"[1] revolution. This has produced the following:

i) between 70% and 80% of all corporate data is stored or exists solely in an electronic format;

ii) 30% of the data that is generated in computer form never gets *published* in paper format;

iii) the amount of "corporate mail" has been reduced by about 35% from several years ago.

[**] Carl D. Liggio, Sr. cliggio@mcandl.com A founding member of ACCA and its Second Chairman, currently is a member of the Board of Directors of FIOS, Inc. He also serves as General Counsel to EthicsPoint and Tempico, Inc.

[1] For purposes of this presentation, the term "electronic media" includes any media which is stored electronically whether it is in a manipulable file or stored as an image file using one of the standard imaging file formats such as .pdf, .tif, .jpg, etc., which *theoretically* are not manipulable.

C. The implications of this data transformation are significant for businesses and corporate counsel.

1. From a business standpoint, policies must be developed for the management, use, retention, and safe keeping of this electronic data maze.

2. Counsel must be ever mindful of the "legal" implications and responsibilities that exist with respect to this ever increasing mountain of bits and bites and their potential toxic nature including the direct and "indirect" costs generated in litigation by this data mass.

II. The Need for a Document Retention Program:

A. Business needs and Legal Needs.

1. Information is the life blood of a business. It must be available for use and when needed by the corporate entity.

2. It becomes a key component of the litigation process. Counsel must carefully study how a client's document retention policies will play in and interact with the litigation process. e-Discovery has thus becoming one of the hottest fields in litigation.

III. The Development of a Document retention policy:

A. A joint effort involving multiple units of the business.

1. The business people: what are the records that are needed to run the business on a day to day and historical basis.

2. What are the legal requirements with respect to those records and others?

3. This must be a team effort in creating, implementing, managing, enforcing, and continually reviewing the Document Retention policy.

B. Considerations and Criteria in creating a Document Retention policy.

1. Is there a difference between electronic media and all other forms of media (i.e., paper bound)? In theory no, in practice yes.

i) *findability*;

ii) accessibility;

iii) availability

2. Are there options as to what you can keep in electronic media as opposed to paper media?

i) unless a hard or paper copy is statutorily required, almost all data may be stored in electronic media format and a separate hard copy need not be maintained. E.g., until the IRS updated its receipt management guidelines a few years ago, taxpayers were required to maintain an original copy of their T&E receipts. The new guidelines provide that companies can use electronic or imaged receipts in place of paper. *See also* the Uniform Photographic Copies of Business and Public Records as Evidence Act and the Uniform Electronics Transactions Act. Illinois has adopted its own form of the later in the Electronic Commerce Security Act 5 ILSC 175/5 *et seq.*

3. How long must you keep records:

i) Three criteria for determining the retention length:

a. legal mandates [need to address federal, state, and even contractual obligations[2]];

b. time needed to do the job to which the records are related;

c. any other retention requirement such as historical needs[3].

ii) The records should be kept for the longest period covered by their need.

4. Enforcement of the policy: procedures must be put in place for the uniform and regular enforcement of the policy including the ability to stop enforcement in the case of litigation. A policy which is not regularly enforced will be suspect if the only time a company shows concern over the implementation of the policy is

[2] For example, T&E records are only required to be kept for a period of three years after the tax return has been filed under IRS regulations, but certain states require that the records be kept for a longer period. In addition, under many contractual arrangements (e.g., franchise document, lease covenants, etc., a party may have contracted to maintain certain records for the duration and then some of their contractual arrangement.) Typically, contractors and subcontractors on government contracts will have numerous contractual retention requirements built into the contract including the application of certain government regulations which otherwise would be inapplicable to the contractor (or subcontractor).

[3] An example of a multi-purpose need for records is discussed in Records Retention Aspects of the Antitrust Laws, Corporate Counsel's Records Retention Report, Feb. 1998, at 1 [although antitrust laws do not require specific records retention periods, "the antitrust laws inevitably require some type of records retention program if the company later wants to establish something, like the availability of an antitrust defense."]

when it is faced with litigation. Enforcement of the policy must be regularly monitored.

5. What standards will a document retention policy by judged by:

   i) Does it meet all legal requirements for duration?

   ii) Does it meet the statutory requirements as to what must be retained?

   iii) Is it a "reasonable standard?" *See, e.g., Lewy v. Remington Arms Co.,* 836 F.2d 1104 (8th Cir. 1988)

   iv) The issue of whether a document retention policy will typically arise in the context of a specific litigation where the issues would revolve around spoliation of evidence and adverse jury charges. [In *Lewy* the documents had been destroyed pursuant to a document retention program in place at the company. The 8th Circuit directed the trial court to determine whether it was reasonable given all of the facts and circumstances; and, even thought the policy was created when the instant litigation had not been brought, whether the policy could be found to have been created in *bad faith.*][4]

6. How accessible are those records. This is the perfect segway into the presentation by Richard Lazar, CEO of FIOS, which specializes in electronic discovery.

IV. Resources that are available on Document Retention:

A. The ACCA InfoPAKSM at ACCA OnlineSM at www.acca.com/infopaks/recretent.html

B. OSHA Records Retention Requirements, Corporate Counsel's Records Retention Report, Apr. 1998, at 1-7 (discussing the requirements outlined at 29 C.F.R. part 1904 (1994)); Records Retention Requirements Related to Independent Contractors, Corporate Counsel's Records Retention Report, May 1998, at 1, 4 (discussing the requirements of 26 U.S.C. § 6041(a) (1994)); Records Retention Requirements for Importers, Corporate Counsel's Records Retention Report, June 1998, at 1-8 (examining 19 U.S.C. §508 (1994)); Record-Keeping Requirements under the Fair Labor Standards Act, Corporate Counsel's Records Retention

[4] The court held that "if the corporation knew or should have known that the documents would become material at some point in the future then such documents should have been preserved." 836 F.2d at 1112. In reaching this result the court found that "a corporation cannot blindly destroy documents and expect to be shielded by a seemingly innocuous document retention policy." *Id.*

Report, July 1998, at 1-8 (exploring 29 U.S.C. § 211 (c) (1994) and 29 C.F.R. § 516 (1994)); Records Retention Requirements under ERISA, Corporate Counsel's Records Retention Report, Sept. 1998, at 1-6 (discussing 29 U.S.C. § 1027 (1994) and 29 U.S.C. § 1059 (1994)); Record Retention Requirements under OSHA's Hazard Communication Rule, Corporate Counsel's Records Retention Report, Oct. 1998, at 1-8 (exploring 29 C.F.R. § 1910.1200 (1994)); Records Retention and Reporting under the Fair Employment Practices Laws, Corporate Counsel's Records Retention Report, Nov. 1998, at 1-7 (discussing 29 C.F.R. part 1602 (1994)). See generally Donald S. Skupsky, Recordkeeping Requirements 69-110 (1988) (discussing specific requirements for tax records, employment records, general business records, and selected common business records).

C. For other records retention guidelines see: Henry E. Knoblock & Christopher J. MacKrell, Sample Document Retention Guidelines (1995), in American Corporate Counsel Association, Records Retention Manual (1st ed. Supp. 1995)

D. In the September issue of the Docket, in an article[5] I co-authored, we published a basic document retention policy. This should only be considered as a building block for the development of a company specific policy that reflects the needs of your particular company:

> The objective of this guideline is to establish a requirement for corporations and each of their subsidiaries and divisions to develop and implement an appropriate records retention program that meets the following criteria:
>
> 1. All records shall be retained for the period required by applicable state and federal laws and regulations.
>
> 2. Adequate records shall be developed and maintained to document the company's compliance with all relevant laws and regulations.
>
> 3. All records necessary for business reasons shall be retained for a period of time that will reasonably assure the availability of those records when needed.
>
> 4. Records vital to the ongoing operation of the business shall be identified and appropriately safeguarded.
>
> 5. All records not necessary for legal or business reasons and not required to be retained by law or regulation shall be destroyed in

[5] After the Storm: A Post-Enron Look at Document Retention Policies, ACCA Docket, September, 2002, Carl D. Liggio, James G. Derouin, and J. Edwin Dietel

order to reduce the high cost of storing, indexing, and handling the vast amount of documents that would otherwise accumulate.

6. Destruction of records shall take place only in compliance with a standard written policy in order to avoid any inference that any document was destroyed in anticipation of a specific problem.

7. Documents that are not otherwise subject to retention for business reasons may need to be retained because of unusual circumstances, such as litigation or a government investigation. If for any reason it is felt that an unusual circumstance exists or arises, the legal department shall be notified immediately. When litigation or investigations occur, the legal department will notify the appropriate departments and direct that relevant categories of documents be labeled for retention until further notice.

8. The privacy and security of records shall be appropriately assured.

9. This policy shall apply to records maintained on microfilm and microfiche, magnetic tape, or other electronic data processing storage media.

10. Records, such as notes, memoranda, letters, reports, computer disks, tapes, and so forth, maintained in individual offices, at home, or any other offsite location are subject to these guidelines and shall be managed consistent with these guidelines.

Delegation

Responsibility for establishing and implementing an appropriate records retention program in each corporate division is delegated to the division controller.

Final Approval

Each records retention program shall be in writing and must be approved by the corporate controller and general counsel. This policy has already been approved by the corporate controller and general counsel and, if used without change, does not require further approval. Changes must be approved by the corporate controller and general counsel before they are implemented.

Audit

The corporate controller and general counsel shall be responsible for auditing the existence and content of all records retention programs. Each division controller shall be responsible for auditing the actual implementation of such programs at the various operating units.

Exceptions

Requests for exceptions from this policy must be submitted to the corporate controller and general counsel for approval before implementation. In order to obtain an exception from this policy, there must be a program that will assure compliance with the basic objectives stated above at least as effectively as this policy.

Review

The corporate controller and general counsel will review this policy annually. Suggested changes should be submitted to the corporate controller. Changes in this policy necessitated by changes to laws or regulations (or the existence of new litigation or an investigation) will be communicated directly by the corporate controller to each of the division controllers who shall cause appropriate changes to be made in the records retention plans of their respective division.

Interpretation

The corporate controller and general counsel will be responsible for interpreting any portions of this management guideline or the records retention plans as they may apply to specific situations.

V. Now Comes the Document Discovery

A. E-Discovery is one of the fastest growing areas in the law. As noted above, with the overwhelming volume of materials now in electronic format, whether you are a plaintiff or defendant you cannot ignore the benefits and detriments of e-discovery.

B. Rather than repeat materials, I have attached the FIOS guide to Electronic-Discovery and an article on cost compliance issues developed by us at FIOS. The former is a comprehensive guide to e-discovery issues. Notwithstanding the comprehensiveness of that guide, I would like to highlight a few issues that you need to focus on:

1. Preservation of electronic records once a suit is started: This can become a major issue with significant cost issues associated with it. Most companies routinely backup their electronic files on a daily basis. They then routinely rotate and recycle those backup tapes. The backup tapes may contain the only electronic records for certain periods of time as the original files are written over or erased. Be careful: (i) if you are the subject of a request for your electronic files, you may have an obligation to preserve all of the backup tapes absent a dispensation from the court. This can then require your IT department to purchase additional backup tapes at considerable expense; (ii) if you are the party seeking electronic

discovery, you want to ascertain the opposing sides back up tape policy and whether you need a preservation order for those tapes. Which ever side of the issue you are on, you will want to address who pays the costs of this process.

2. <u>Where are the electronic records located</u>: Is everything on a central file or with the advent of laptops, are many corporate electronic records now located only on laptops for which there may be no central corporate access. How do you identify these to make sure you are in compliance with the discovery order or conversely, how do you make sure your opposing counsel is searching these records. A corollary to this is the use of an employee's personal computer (maybe at home) for the storage of company records. You will be surprised to find how often this happens and even more surprised to find out what is there.[6]

3. <u>Understand the "meta data" issues.</u> Every electronic record has considerable "meta data" associated with it. This is not readily seen by the normal user of the document, but a skilled electronic discovery person can quickly extract that data. Some of the things that can be shown by this data include:

   i) the computer on which the record was originally created;

   ii) who first created the record;

   iii) all changes in the record;

   iv) the number of revisions to the document from when it was first written;

   v) the amount of time spent editing the record;

   vi) the identity of the last computer or possibly person who edited the document;

   vii) the date the document was first created (this can be very important because a document may have no date or a "pseudo" date – *i.e.,* the date it was printed is automatically printed on the document if certain settings are used in Word or Excel. Thus, you can be misled by the production date on a document.)

[6] When I was general counsel of Ernst & Young, I had a policy of requiring all personnel who were involved in a matter that was in litigation to provide a written confirmation to the legal department that they had no records at home or at any other site and had in fact turned over all of the records to us. The first time one of the firm's partners encountered this with me, he thought it was overkill until two days later an audit manager walked into his office and asked if that meant the two boxes of documents he had in his garage at home hard to be turned over to the legal department. This is a more acute problem with the advent of easily replicated electronic media and one which cannot be ignored by counsel.

Needless to say this data can be extremely revealing in a case.

4. Understand the difference between document types and their alterability. (This is important when you are reviewing discovery of an opposing party.) For example, a "pdf" is less likely to have been altered from the original creation date than a standard file such as a "Word" or "Excel" file. Also understand how documents can be created with "pseudo" creation dates by altering the startup file of a computer.

5. The cost of electronic discovery. Although the price of it may seem high at first, effective electronic discovery will actually be a material cost saver in terms of attorney and support staff review time. Try to avoid the paper morass and utilize the power of the electronic medium to obtain, review, and analyze documents in their electronic format. Understand that there are significant cost savings over not having to print, store, and find paper copies of documents. In most cases, the use of effective electronic discovery will actually be a significant cost saver.

6. Understand the tools available for analyzing electronic media. In the past two decades the industry has developed more cost effective tools for the handling of electronic information. We have gone from flat data bases to relational data bases to extremely powerful search engines which combine many of the characteristics of the historical data base. One of the most significant electronic discovery tools that is available just this year is the "contextual" search tool. This operates like a relational data base, but it uses sophisticated algorithms to search documents for relationships between words and "concepts". The search result produces groupings of documents which are likely to be related. It is the tool to help find the "needle in the haystack." Today, it is the most effective way to search large masses of unindexed textual material.

the guide to electronic discovery

contents

Published by Fios, Inc., Portland, OR

Fios, Inc.
921 SW Washington Street, Suite 850, Portland, OR 97205
Fios is a registered trademark of Fios, Inc.

ISBN 0-9725542-0-3

Printed in the United States of America

what is electronic discovery?

Electronic discovery is the collection, preparation, review and distribution of electronic documents associated with legal and government proceedings. "Documents," as defined in most requests for production, include e-mail messages and associated file attachments, memos, reports, plain text files, spreadsheets, digital art and photos, presentations and any other data that is created or stored on a computer, computer network or other storage media.

The reach of electronic discovery is significant. Data source targets for electronic discovery include files residing on laptops, office PCs, network servers, floppy discs, PDAs, CD-ROMs, tape backups, other archive media and third-party storage and archival systems.

growth in electronic discovery usage

A number of market drivers are creating a dramatic increase in the international use of electronic discovery services. The most significant is the explosive growth in electronic data. According to recent estimates published in *Law Technology News*, at least 93% of business documents are created electronically, and more than 35% of corporate communications never reach paper. The prevalence of e-mail as a primary form of corporate communication adds to the enormity of electronic documents in use today. According to estimates published in *Wired* magazine, U.S. office workers exchanged approximately 7 trillion e-mail messages in the year 2000.

Another key driver is the widespread need for risk management in today's economy. The heightened scrutiny of corporate activity translates into a corresponding increase in the liability exposure of CEO, president, CFO and other "CXO" positions. This exposure is frequently linked to electronic corporate communications such as e-mail and associated file attachments. The electronic data storage and destruction policies of corporations have an increased impact during litigation, and the need for prompt and thorough access to electronic documents is crucial.

electronic data — not paper —
reveals the complete story

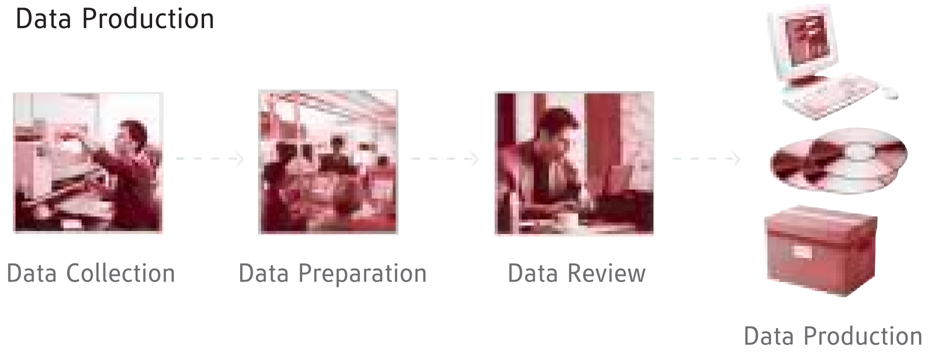
The physical, time and cost limitations of paper-based document review become a harsh reality when viewed against the sheer mass of electronic information created in corporate settings. For those familiar with electronic discovery, it is clear that electronic content and audit trails reveal much more than their paper counterparts. An electronic file contains easily accessible and highly reliable corporate knowledge, leaves a metadata chronology of key dates, comments between collaborators and, in effect, a knowledge map of who knew what, and when it became known.

Given these complexities, the process of electronic discovery may seem overwhelming, leading many to accept the significant legal risks of avoiding the process altogether. Yet, navigating the seemingly complex maze of the electronic discovery process does not have to be difficult. Electronic discovery can be quite manageable when you are familiar with the process and if you select the right partner.

key steps in electronic discovery

The process of electronic discovery is quite straightforward. There are five fundamental phases of the electronic discovery process:

1. Electronic Discovery Strategy
2. Data Collection
3. Data Preparation
4. Data Review
5. Data Production



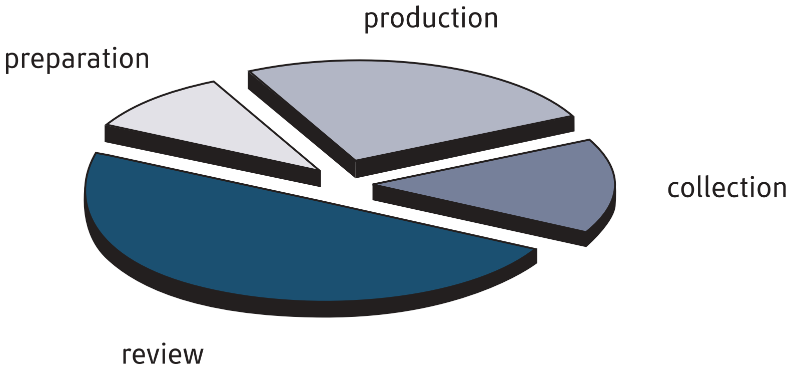
The information in the following chapters of *The Guide to Electronic Discovery* walks you through each of these main categories of activity, providing a practical overview of how to approach each step and how to avoid common pitfalls. As a working tool to help you navigate the process, this guide also describes best practices associated with a successful electronic discovery project.

electronic discovery strategy: the most important step

Early preparation and a well-structured strategy are crucial to simplifying the process of electronic discovery. Knowing up front how to handle the quantity and variety of data you're targeting, the individuals within the company who are most important, and what timelines you'll encounter will save you time and money in the long run.

time expectations

The time required to comply with an electronic discovery request is highly dependent on the scope of subject matter deemed potentially relevant, the amount of data involved, and your desired output for final production. The following time segmentation is a rough estimate of the relative time requirements for data collection, preparation, review and production. Be certain that you allow sufficient time for each phase of activity, especially the document review stage. By anticipating and planning for the resources you'll need, such as staffing, hardware, connectivity and communication methods, you'll be able to significantly streamline the review process.



How to Strategize During Early Stages of Actual or Anticipated Legal Proceedings or Government Agency Investigations:

Step One: Assess the Potential Scope of the Project

1. Frame the legal issues and landscape. Identify and outline what the case is about, what time periods it covers, whether there have been similar suits in the past, characteristics of the opposition, the dollar amount of the exposure, etc.

2. Interview the company's document management team and information technology (IT) managers.

   □ Discuss document retention policies and practices, with specific attention to document destruction policies and procedures. You may have to discuss retention practices with opposing counsel and do not want to encounter any surprises. Retention policies, even if they exist, are not always properly followed.

   □ Find out how the company's document management system, if one exists, is structured. These systems utilize software that labels content based on data source, keywords and other parameters. You'll want to know how much total data is in the system, how old the data is and how much of it is potentially relevant to the specific matter. Additionally, you should identify any parameters that purge data based on content or time periods, as well as the frequency of backups.

   □ Determine the level of readiness to produce documents. How long will it take to identify and collect the live data and restore the relevant backup media, as well as review and produce relevant, nonprivileged documents?

   □ Identify those who have the greatest knowledge of the computer system. These individuals are likely targets for interrogatories and depositions.

   □ Look for public information that could make other IT team members likely targets for the opposition. For example, it is often easy to find the names and contact information of technical managers on your client's Web site who may unintentionally disclose information about data retention or other policies.

3. Quickly address any spoliation issues.

   The term "spoliation" means the destruction or significant alteration of evidence, or the failure to preserve evidence, in pending, imminent or reasonably foreseeable legal proceedings. In such a situation, you may want to advise your client to avoid deleting or erasing information pursuant to their ordinary course of business. This is prudent even if such deletion was normal in the course of their standard business practices. A duty to preserve may arise when your client knows or should have reasonable knowledge that information is:

- The subject of a pending discovery request;
- Reasonably likely to lead to the discovery of admissible evidence;
- Relevant to the anticipated or pending legal action; or
- Reasonably likely to be requested during discovery.

There are a variety of actions you can take to prevent spoliation, including the following:

□ Identify any selective document destruction policies.

□ Send out legal hold instructions to your client's information technology team. Schedule informational sessions with the team to explain the risks of spoliation and why the hold instructions are critically important.

□ Suspend normal data destruction until further notice. Be sure, however, that a moratorium on destruction does not unnecessarily or adversely affect unrelated departments or areas of data storage. Also be sensitive to operating considerations, as computers need to be cleared of transaction data and logs in order to maintain peak performance.

□ Send a spoliation letter to the opposition (see Appendix A for a sample spoliation letter).

□ Take "snapshots" (exact copies) of e-mail boxes and other data stores of key individuals to maintain important evidence as well as data integrity and authenticity.

Step Two: Anticipate and Assess the Potential Data Collection Universe

Identifying and assessing the logical and reasonable criteria for the discovery related to your client will enable you to proactively manage the process of electronic discovery. The following four exercises will help you properly assess the potential data collection universe:

1. Determine the criteria for "Privilege" and "Relevancy."

   □ Determine what constitutes work product.

   □ Consider ways to limit inadvertent waivers of privilege.

2. Determine the potential quantity of electronic data involved.

   The amount of data involved affects many aspects of the project, including the method of collection and review, the number of reviewers needed, the time the project will take and how much it will cost. Work with the company's information technology staff to gain an understanding of the potential quantity of data that may need to be collected.

3. Identify potential time, custodian and/or user splices.

   Time, custodian and user splices are the best ways to identify and rapidly cull (reduce) your data set so you won't unnecessarily collect too much data or perform inordinately large searches. Identifying these splices early in the process will also help you stay organized and reduce the likelihood that you'll do redundant searches at a later date.

4. Review and negotiate the scope of discovery requests.

   Review requests for production to help determine what data to target and collect. Negotiation for the scope of discovery is sometimes based on timeframes, custodians, users and search terms and may be a way to help manage the size of the undertaking.

Step Three: Decide Your Form of Review

There are many available methods for reviewing electronic discovery data. Early determination of your preferred method of review can help the project go more smoothly. Review can occur via a Web-based electronic discovery application, enterprise or stand-alone litigation support software, native file application software or paper. Each method has time, cost, efficiency, human resource and reliability implications, so this decision must be made carefully. A capable electronic discovery vendor can help guide you through this decision process.

Step Four: Decide Your Form of Production

While it may seem like a small logistical detail, outlining your production requirements during the strategy phase can result in a significantly more efficient production process. You'll want to evaluate the following production options:

□ Decide whether you want to produce as paper, electronic data or both.

□ Determine if you'll need multiple production sets.

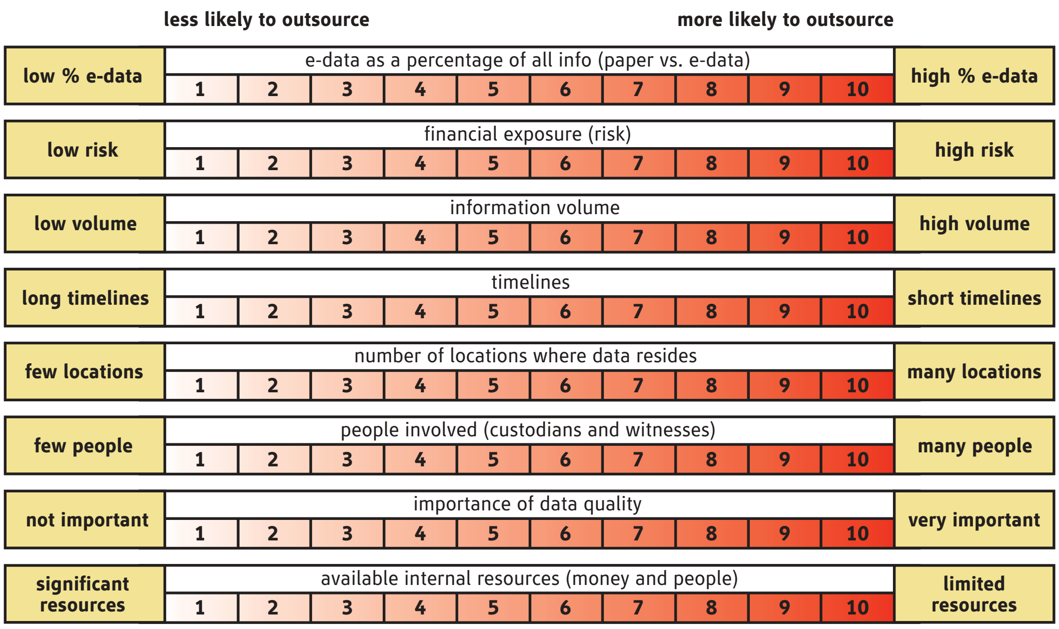
□ Clarify delivery deadlines and locations.

Step Five: Select an Electronic Discovery Service Vendor

For the sake of efficiency, you will likely want to engage an electronic discovery vendor to help you manage the process. A qualified vendor will be able to expedite the completion of the various discovery phases. The benefits of engaging professional help for electronic discovery are significantly greater if you involve the vendor early in the discovery process.

There are a number of evaluation factors to consider before deciding to outsource your electronic discovery project. The following chart illustrates a number of situational criteria you'll want to assess prior to making your decision.

factors influencing e-discovery outsourcing decision



	less likely to outsource									more likely to outsource	
low % e-data	e-data as a percentage of all info (paper vs. e-data)										high % e-data
	1	2	3	4	5	6	7	8	9	10	
low risk	financial exposure (risk)										high risk
	1	2	3	4	5	6	7	8	9	10	
low volume	information volume										high volume
	1	2	3	4	5	6	7	8	9	10	
long timelines	timelines										short timelines
	1	2	3	4	5	6	7	8	9	10	
few locations	number of locations where data resides										many locations
	1	2	3	4	5	6	7	8	9	10	
few people	people involved (custodians and witnesses)										many people
	1	2	3	4	5	6	7	8	9	10	
not important	importance of data quality										very important
	1	2	3	4	5	6	7	8	9	10	
significant resources	available internal resources (money and people)										limited resources
	1	2	3	4	5	6	7	8	9	10	

vendor selection suggestions

The vendor you engage for an electronic discovery project can significantly enhance or hinder your litigation management process. Spending a modest amount of time to gain a detailed understanding of each vendor's capabilities and track record will ensure that you select the most appropriate partner. There are a number of evaluation criteria to consider during the selection process. The following list provides a cross-section of consideration points to assist you in your decision.

- □ Talk with your colleagues and litigation support teams for initial recommendations on the most capable and reliable electronic discovery vendors. The vendor you select must have extensive electronic data collection experience for use in litigation.

- □ Initial due diligence on your vendor should include a dialogue review of their most complex electronic discovery projects and a detailed description of how they manage discovery projects from start to finish.

- □ Particularly for large projects, it is essential that you conduct site visits to the electronic discovery vendors' facilities to gauge the level of sophistication, organization and security they offer for client projects.

- □ The vendor needs in-depth expertise in rapidly identifying and collecting large quantities of data from multiple sources such as laptops, office PCs, network servers, floppy discs, CD-ROMs, tape backups, other archive media and third-party storage and archival systems.

- □ Be certain the vendor has comprehensive procedures to establish and maintain a continuing chain of custody for sensitive electronic material, ensuring data security, accuracy and authentication.

- □ The vendor must be able to digitally fingerprint each file to ensure a legally defensible audit trail.

- □ The vendor must identify documents not by extension but by state-of-the-art signature analysis. Because discovery strategies often require prioritizing the document types to review, knowing unequivocally the true file type distribution is critical.

□ The vendor must ensure that original data is never compromised by virus scanning all data to prevent cross-contamination.

□ The vendor should provide state-of-the-art security for data storage, including secure buildings, processing centers and employees. Firewall intrusion detection and protection, secure digital certificates, and a minimum of 512-bit SSL keys with 128-bit SSL encryption for any data transmissions are also extremely prudent.

□ The vendor's client services team should have in-depth legal experience and data engineers who understand the litigation process. You should request an introduction to the specific project manager who would be assigned to your account.

□ Make sure the vendor isolates the experts you may need to testify for chain of custody purposes from the people helping you with your review strategy.

□ Be certain that the vendor is able to aggregate and convert electronic data into common formats such as HTML, TIFF and PDF, creating a unified and secure database of original content and metadata, fully indexed and optimized for searching.

□ When it comes to the process of reviewing data, flexibility is of paramount importance. You should have the option to review data using the tool of your choice. Whether Web-based review is selected, or any number of stand-alone litigation support review tools, the choice should be yours rather than the vendor's.

□ Because Web-based (online) review is becoming the platform of choice for litigators, your vendor should offer a Web-based tool to enable your review team to work in any location that has Internet access, and to collaborate via secure instant messaging.

□ For Web-based data review tools, you should select a vendor that offers a pricing model that allows unlimited usage and downloads for any number of users and for any length of online time. You should avoid those that charge on a per-seat, per-document (i.e., per click) or per-hour license basis for data review.

□ Regardless of the review tool's document display format, you should select a vendor whose review tool allows individual reviewers to quickly download copies of original documents in their native format. Further, the tool should support controlling who can download copies of original documents, and all downloads should be logged, enabling you to track who downloaded what material on specific dates.

□ Your vendor should enable you to identify the types of content to be included in the unified database via search parameters such as date ranges, custodians or data locations.

□ Your vendor should be able to help strategize and consult on search term filtering as a way to limit the discovery project. Identifying and filtering the right terms is a critically important activity.

□ Your vendor's production standards should be impeccable. Output of selected electronic data must be produced quickly and delivered to your exact specifications. Flexibility is critically important. The vendor should be able to produce selected documents in a wide variety of formats, including new Web-based environments, PDFs, TIFF images with or without associated databases, native files, paper or any combination of these formats.

□ Your vendor should offer content management flexibility and multiple output options related to the review and production phases of your matter.

□ Talk with vendor references, and inquire how the vendor handles the data collection, management and production challenges that are a normal course of litigation.

Step Six: Negotiate the Discovery

Because the quantity of corporate electronic data is immense, it is important to limit electronic discovery to only that portion of the data population that is potentially relevant to the proceeding. Sometimes the process of limiting discovery requires negotiation with opposing parties or court involvement. Your electronic discovery vendor should be able to help you during these negotiations, especially regarding search terms that can affect the information you'd need to produce, as well as what evidentiary information you'll receive. Strategies can include:

□ Meet with your opposition directly, or schedule and conduct a Rule 16 conference with a judge.

□ Propose user and time splices.

□ Propose location splices.

□ Propose search terms.

tip — interrogatories and depositions

A wealth of information can be targeted via electronic discovery interrogatories, requests for production and depositions. Appendix B of this book provides detailed suggested interrogatories and requests for production. Appendix C provides sample FRCP 30(b)(6) deposition questions.

After you have taken the recommended steps to establish your strategy and have selected your vendor, you're ready to begin the data collection phase of electronic discovery. The following chapter will help guide you through this process.

data collection:
maintaining the audit trail

Due to the distributed nature of data in today's business environment, you will find that your client's set of corporate electronic information is physically located in different geographic locations and systems, controlled by many people and departments. The first step toward gaining visibility into corporate content is data identification.

Many businesses are diligent and organized about their electronic data storage and backup procedures, yet, at many other companies, data is not predictably stored or routinely preserved. You'll find that systematically stored and backed up data is easier to access and can flow smoothly into a preservation system structured for legal purposes. Whether a company is diligent or lax about data storage and backup, however, transitory data such as e-mails or documents kept on desktop computers is challenging to secure and gather. Fortunately, the complexities of collecting data can be easily managed by a competent electronic discovery services vendor.

To perform comprehensive data collection, your electronic discovery vendor will need to work with your client's management and information technology teams to identify those areas that are of critical interest to your legal team. Your vendor should have a technically skilled and articulate data collection team that can partner with your client to collect information in a professional manner. The data preservation methods chosen should take into account personnel resources, capital cost and the necessity of keeping business processes running at optimal levels.

key steps in data collection

For large organizations, the process of data acquisition from multiple office locations and scores of employees can be difficult without the right tools and processes. You'll want to be sure that your team or vendor is experienced at rapidly identifying and properly collecting large quantities of data from multiple sources.

Maintaining data integrity is crucial. If you can't authenticate the data, legal risks are created. You'll need to be certain that a complete chain of custody for collected data is maintained to ensure accurate and authenticated information. Appendix D of this guide provides useful data collection forms to help you document the chain of custody. Also, be sure that you employ procedures to minimize the risk of damaging, destroying or otherwise compromising evidence during the collection process.

The following steps are crucial for complete and methodical data collection:

□ Identify stores of data that may fall within the discovery requirements.

□ Maintain chain of custody for any data that is collected.

□ Document the source, as well as the rationale, for what data you decide to include in your data set.

□ Request a list of all relevant data custodians from your client.

□ Match actual name (last, first) with ID used by the client's data management system. Review names from prior collection(s).

□ Determine the location of data, from sources such as:

- E-mail servers
- File and print servers
- Desktops and onsite laptops
- Field laptops
- Shared directories
- Backup tapes

□ Work with an infrastructure specialist to set up collection servers at optimum network points.

□ Use appropriate tools and chain of custody documentation and, when feasible, copy the data over the network to removable hard drives. Be aware that doing a straight copy can result in altered metadata, hence you'll want to make sure that any copying is done by an expert in data collection for electronic discovery. Regardless of whether your electronic discovery vendor or your client accomplishes the collection, your vendor will need some assistance from the client to access the appropriate areas of the network and specific drives for security reasons.

□ If your team performs data copying tasks, you'll need to securely package the data and send or hand-carry it to your electronic discovery vendor with chain of custody forms. To maintain chain of custody documentation, your electronic discovery vendor must properly receive the data.

The integrity and quality of your electronic data set is clearly of paramount importance. To assist you in proper tracking and documentation during this phase of the electronic discovery process, we have included a variety of useful data collection forms in Appendix D.

tip — date range negotiations

If your client's data has already been collected, find out how it was collected prior to negotiating the date ranges for inclusion in the discovery set. Date range splices, if done without special collection tools, should be driven by the "last modified" date, as this date is typically the most reliable. Specifying files dated after a given date introduces the risk of too many files being collected as part of your production, given that normal copy tools can change the "create" and "access" dates of files to that of the date of collection.

data magnitude explained

When facing the prospect of data collection, many litigation teams are not yet familiar with the sheer magnitude of data stored on CD-ROMs, hard drives and other storage media. The terms "megabytes," "gigabytes," and "terabytes" are commonly used, but often it is difficult to gauge what those data volumes mean in terms of page equivalents. The following conversion table provides a data equivalent overview.

The assumptions for this exhibit are that the average banker's box holds 2,500 sheets of paper, and one page of information on average equals 20 kilobytes (.02 megabytes). This page-to-data size conversion factor is conservative. For example, collections consisting largely of e-mails and spreadsheets may have a conversion factor of .01 or even .005, resulting in two to four times as many "pages" or page-equivalents.

Boxes of Paper	Total Pages	Megabytes; Gigabytes; Terabytes
1	2,500	50 Megabytes
10	25,000	500
20	50,000	1 Gigabyte
100	250,000	5
200	500,000	10
300	750,000	15
400	1,000,000	20
500	1,250,000	25
1,000	2,500,000	50
2,000	5,000,000	100
5,000	12,500,000	250
10,000	25,000,000	500
20,000	50,000,000	1 Terabyte
40,000	100,00,000	2
60,000	150,000,000	3

the importance of chain custody

A defensible chain of custody is clearly vital, yet the passage of time can pose some unexpected challenges. Larger legal cases can often last for years, and information technology teams often experience frequent turnover. It is critically important, therefore, to gather the contact information for both the subject of the data collection as well as the person doing the collecting.

You may want to utilize serial numbers and asset tags to tie the data collected to a specific machine, which thus ties it to a person. You'll also want to track what external devices are on a custodian's machine to ensure knowledge regarding the machine's external media capacity. For example, documenting that the CD drive is a CD-Read Only device will reduce the likelihood that the opposition can charge that CD backups were produced from that custodian's computer.

This physical audit trail to the electronic data can be of crucial importance. Appendix B of this guide provides sample data collection checklists that are useful for managing the process with appropriate documentation. The primary purpose for this level of detail is that the data collection staff will not have to rely on memory should they need to be deposed.

Managing chain of custody includes maintaining records of chronological and logistical information such as:

- □ What was the source of the data (custodian and location)?
- □ Where on the hard drive were individual files physically located?
- □ What metadata is available, and what files does it link to?
- □ What was the relationship between e-mails and their attachments?
- □ When did the data arrive?
- □ What was the complete chain of custody up to the production phase of the project?

tip — data collection efficiency

It is best to do as much live data collection as possible up front rather than having it come in small increments. To optimize the data management and culling, it is best to process and review sets of media (such as quarterly backups), rather than one-at-a-time items (such as daily backups). Once your data arrives, your electronic discovery vendor should be able to do rapid production in rolling deliveries. With that said, however, there is often an unavoidable incremental nature to data collection. You may have to refresh the data set when new information becomes available. Try to keep these collection events as consolidated as possible to make the process easier and more productive for your team. Your electronic discovery vendor should assist you in tracking collection events.

common data collection questions

Are electronic discovery and computer forensics the same thing?
No, electronic discovery and forensics are not the same thing, but forensics is an important subset of the electronic discovery process. The first requirement for electronic discovery is to find and record all necessary "live," or easily accessible, files on a company's computer system. Digital information, however, is not always readily accessible. Some files may be deleted, corrupted or otherwise hidden. In addition to collecting currently active files through "live data" collection (common electronic discovery collection), computer forensics allows the collection of deleted, hidden, password-protected and encrypted files and file fragments.

Electronic discovery of "live data" is essential in most cases, but computer forensics is required in only limited circumstances, such as cases relating to key personnel, bad faith, intellectual property and/or wrongful termination. To reduce the risk of exposing your work product to the opposition, be certain that your electronic discovery vendor maintains appropriate confidentiality of information and does not unnecessarily share information with the forensics experts who may be deposed and testify about forensic findings.

What is a program file, and why can it be culled?
Program files, like those associated with software applications and operating systems, provide the mechanical ability for items such as documents and spreadsheets to be created and for computers to function. They do not contain any material that the end user has created. Discovery does not include those program files (with file extensions such as ".EXE," ".HLP," ".DLL," ".LST," etc.), thus they are not required and can be culled.

Your electronic discovery vendor should maintain a library to identify and remove common application program files such as those installed with Microsoft® Office, so that your reviewers will not need to review the standard templates provided with the programs. For example, Microsoft® Word comes with many templates and sample files, which on the surface appear to be user-created files due to their file extensions (.DOT or .DOC). A sophisticated electronic discovery vendor can cull these templates and samples, thus your review team won't waste time reviewing them.

What is metadata, and why is it important?

Metadata is often described as "data about the data." It includes information such as file dates, authors, source locations and e-mail routing information that generally does not appear on the printed page. The electronic discovery process is the best method to access such metadata, which can contain important information about who created and reviewed electronic documents and to whom they were distributed.

Metadata is indicative of, but not definitive of, dates, creators of documents and other properties. Metadata is collected and stored differently by each application program. Even versions of the same program can treat metadata differently. Your electronic discovery vendor should help you understand the nuances of metadata and ensure that your data is collected and produced with impeccable chain of custody procedures. You should also make sure that your electronic discovery vendor has the ability to export all types of metadata to your review database.

There are five key types of metadata your electronic discovery vendor should identify and manage:

- File system metadata – data that can be obtained or extracted about a file from the file system storing the file.

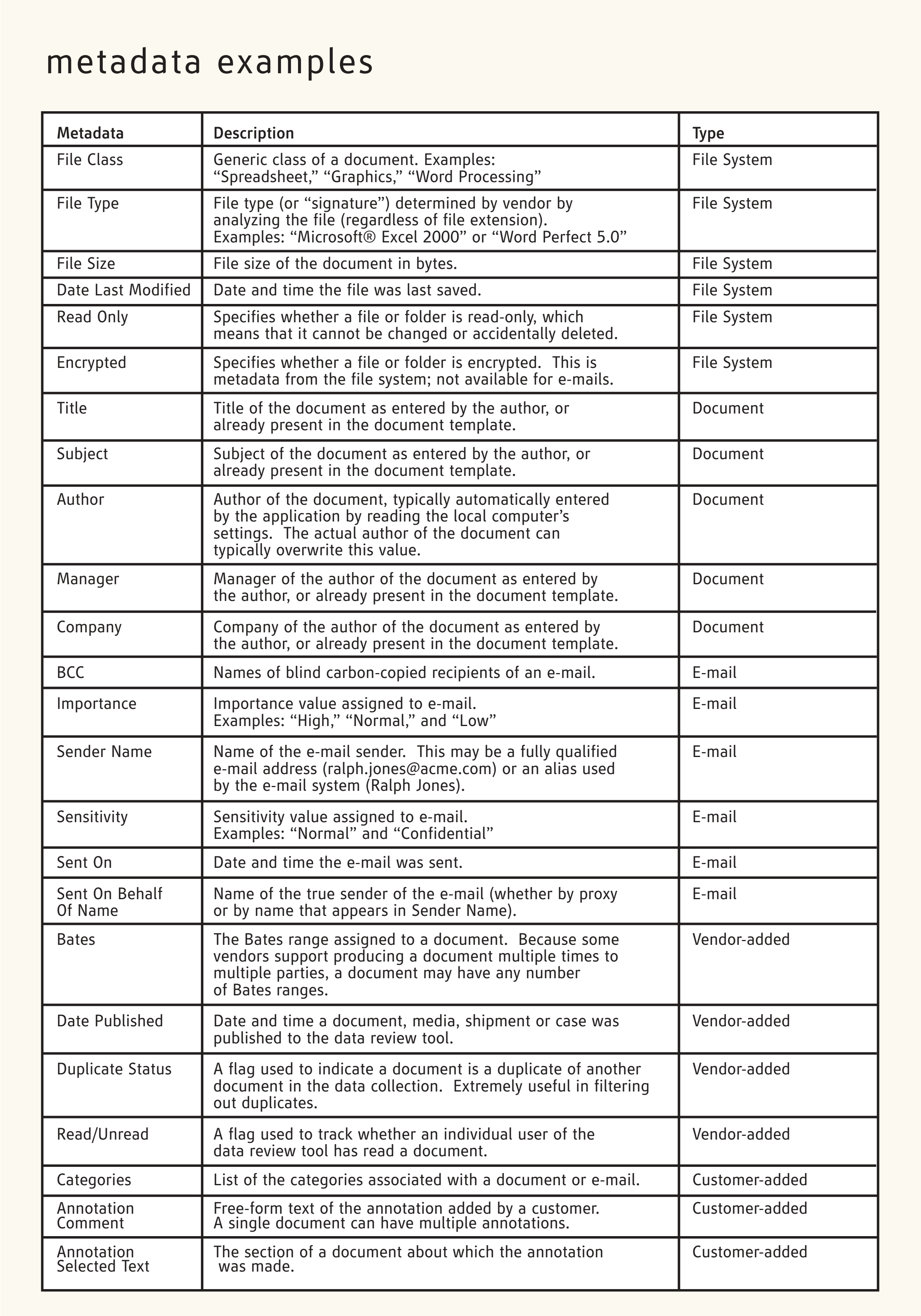
- Document metadata – data stored in the document about the document. Often this data is not immediately viewable in the application used to create/edit the document but often can be accessed via a "Properties" view.

- E-mail metadata – data stored in the e-mail about the e-mail. Often this data is not even viewable in e-mail client applications used to create the e-mail. The amount of e-mail metadata available for a particular e-mail varies greatly depending on the e-mail system.

- Vendor-added metadata – data created and maintained by your electronic discovery vendor as a result of processing the item. While some of the vendor-added metadata has direct value to you, much of it is used for process reporting, chain of custody and data accountability.

- Customer-added metadata – data or "work product" created by a customer while reviewing the document set.

See the table on the following page for examples of each type of metadata.



metadata examples

Metadata	Description	Type
File Class	Generic class of a document. Examples: "Spreadsheet," "Graphics," "Word Processing"	File System
File Type	File type (or "signature") determined by vendor by analyzing the file (regardless of file extension). Examples: "Microsoft® Excel 2000" or "Word Perfect 5.0"	File System
File Size	File size of the document in bytes.	File System
Date Last Modified	Date and time the file was last saved.	File System
Read Only	Specifies whether a file or folder is read-only, which means that it cannot be changed or accidentally deleted.	File System
Encrypted	Specifies whether a file or folder is encrypted. This is metadata from the file system; not available for e-mails.	File System
Title	Title of the document as entered by the author, or already present in the document template.	Document
Subject	Subject of the document as entered by the author, or already present in the document template.	Document
Author	Author of the document, typically automatically entered by the application by reading the local computer's settings. The actual author of the document can typically overwrite this value.	Document
Manager	Manager of the author of the document as entered by the author, or already present in the document template.	Document
Company	Company of the author of the document as entered by the author, or already present in the document template.	Document
BCC	Names of blind carbon-copied recipients of an e-mail.	E-mail
Importance	Importance value assigned to e-mail. Examples: "High," "Normal," and "Low"	E-mail
Sender Name	Name of the e-mail sender. This may be a fully qualified e-mail address (ralph.jones@acme.com) or an alias used by the e-mail system (Ralph Jones).	E-mail
Sensitivity	Sensitivity value assigned to e-mail. Examples: "Normal" and "Confidential"	E-mail
Sent On	Date and time the e-mail was sent.	E-mail
Sent On Behalf Of Name	Name of the true sender of the e-mail (whether by proxy or by name that appears in Sender Name).	E-mail
Bates	The Bates range assigned to a document. Because some vendors support producing a document multiple times to multiple parties, a document may have any number of Bates ranges.	Vendor-added
Date Published	Date and time a document, media, shipment or case was published to the data review tool.	Vendor-added
Duplicate Status	A flag used to indicate a document is a duplicate of another document in the data collection. Extremely useful in filtering out duplicates.	Vendor-added
Read/Unread	A flag used to track whether an individual user of the data review tool has read a document.	Vendor-added
Categories	List of the categories associated with a document or e-mail.	Customer-added
Annotation Comment	Free-form text of the annotation added by a customer. A single document can have multiple annotations.	Customer-added
Annotation Selected Text	The section of a document about which the annotation was made.	Customer-added

risks of data collection by employees

A significant area of risk arises if company employees are allowed to determine which data residing on their computers or storage media is potentially relevant. There are multiple problems inherent in employee data selection and collection:

- There may be an inconsistent understanding or interpretation among employees as to what constitutes relevancy. If a corporate management team tells ten of its employees to gather all of their electronic documents related to a specific topic, there will likely be ten different opinions regarding what is relevant. The attorneys responsible for discovery should clearly define and determine what data is considered relevant, rather than leaving that determination up to the original data custodians.

- The lack of a cohesive collection strategy may make the data unreliable. If the data collection activity is too narrow or has the potential for being inconsistent, any changes in the scope or in the issues of the case may drive the need for future rounds of data collection.

- Having employees copy their information over to a centralized location creates a significant risk of data alteration resulting from automatic updating functions within word processing and spreadsheet programs ("AutoSave"), as well as viral exposure.

- Avoiding risk — Employees who are aware of relevant documents within their data sets may be inclined to avoid potential legal risks created by the documents; hence they may inappropriately self-cull the data. Other employees may have embarrassing material that could cause them to cull excessively or delete documents.

- Employees who are involved in the data collection are immediately made fair targets for being called to testify regarding the completeness and accuracy of their data collection.

the dark side of unprotected desk-side data review

When a company provides data or access to custodian desktops to its law firm, the attorneys are often tempted to do a desk-side, native file review of the material by opening and identifying which files meet the relevance criteria, and then copying the files over to another location. The following risks of doing an unprotected review of data are significant:

- It increases the risk of infecting files with viruses.
- Opening and/or moving files that have auto-date or auto-path features activated can change dates and file path references associated with the file, even though you didn't intentionally modify the content.
- Drag and drop activity with files can also automatically change file reference dates, for example, the Last Accessed date. If your legal proceeding is date sensitive, you can inadvertently cause the data to be included or excluded from the "response set."
- Legal teams can encounter unwanted liability if their desktop reviews result in an alteration, damage or total erasure of an important piece of evidence.

After your client's data has been collected, the next phase of the electronic discovery process is data preparation. The following chapter details the activities and tools involved to prepare data for review.

data preparation: aggregating for true visibility

Once your client's data has been collected, you'll need a way to review all data in a format that allows full visibility of every aspect of the data. You'll also want the ability to do large-scale data searches, rather than by individual data custodians one at a time. For this to be done efficiently, those files must be aggregated into a unified database. To create a comprehensive and secure database of original file content and metadata, your electronic data is best managed by converting it into a common viewable format such as HTML or TIFF. The most useful unified databases are fully indexed and optimized for searching and can be output to a variety of formats for review, including the Web, databases or paper.

It is important that your data preparation include activities such as documenting the chain of custody, virus scanning the files and inventorying every file. A culling process can also be done to separate user-created files from system and program files, and to identify and reduce duplicate files.

tip — data culling

Careful data culling can often reduce the volume of desktop data by as much as 85 percent.

Sophisticated Duplicate File Reduction (De-Duplication)

Just "getting rid of duplicate files" should never be considered as an option. You need to make sure there is a legally defensible audit trail for any duplicate files. The question of "who had access to what when" is very relevant and can be missed if duplicate data files are simply deleted. Experienced litigation support managers recognize the value of identifying duplicates across the whole data set so they can be coded and categorized consistently, which will reduce the time and effort expended by your data review team.

Some litigation support vendors practice de-duplication, meaning there's only one copy of any given item within the data set. The problem inherent in this sort of de-duplication approach is that it doesn't allow you to see the data in, or produce it from, the other locations in which that data resided. The fundamental reason for being wary of de-duplication is that the most important data custodian may not be the source of the first copy that was included in the data set. If it is taken out of the data set, then it cannot be seen in context. You must be certain that de-duplication still allows you to get to and produce a copy, even if it's not the primary copy.

There are three primary methods for de-duplicating: backups, case and custodian. The preference for backup, case or custodian de-duplication is your decision. The differences between the three methods are as follows:

> **Backup de-duplication** will look for and retain single copies of documents in the exact same context. If an identical document name, date and file size is found on ten backup tapes, it can be assumed that it is the same file and only one copy of the file needs to be preserved. Unless there is bad faith alleged, it is unlikely that you will need to report on each file of the backup tape.
>
> **Case de-duplication** will look for and retain single copies of documents per case. So if an identical document resides with Mr. A, Mr. B and Mr. C, only the first occurrence of the file (i.e., Mr. A's) will be saved.
>
> **Custodian de-duplication** will only de-duplicate a document if multiple copies of that document reside within the same custodian's data set. So if Mr. A and Mr. B each has a copy of a specific document, and Mr. C has two copies, the system will maintain one copy each for Mr. A, Mr. B and Mr. C.

When it comes time to produce documents, some vendors provide another very useful method for de-duplicating: production.

> **Production de-duplication,** like custodian de-duplication, will only de-duplicate a document if multiple copies of that document reside within the same production set. So, if two identical documents are both marked responsive, nonprivileged, the system will only produce one of those documents.
>
> One pitfall to de-duplicating production sets is allowing the de-duplication process to result in incomplete e-mail-attachment chains. For example, you want to avoid the situation where the same file is attached to two different e-mails, both marked responsive, and one of the attachments is not produced because it is a duplicate. This would result in a broken or incomplete chain. Ask your vendor if they support a rule that ensures that a chain of e-mails and attachments are always produced together, and that this rule "trumps" (takes priority over) any de-duplication rules.

Some electronic discovery vendors or software packages can only thoroughly de-duplicate e-mail files, while others are able to do precise file comparisons to identify unique or identical files. You'll want to be sure that your electronic discovery vendor can do precise file comparisons for all types of files, not just e-mail.

Thorough Data Preparation

Data preparation for electronic discovery purposes is a complex process requiring extensive technology expertise. Electronic document populations generally contain large volumes of disparate file types. For an electronic discovery project to be successful, this data must be prepared and aggregated quickly and reliably.

The following steps should be performed during proper data aggregation:

□ Fully inventory and uniquely number each file in the data population for 100 percent file accountability.

□ Be sure that you have a log of all passwords required to access files. Otherwise, securing prompt access to important file information can be difficult. Make sure that your vendor can break the passwords for critical data custodians.

□ Uncompress compressed files while maintaining the folder structure of the compressed files. Also, uncompression must be done recursively so that if you have a parent ZIP file that contains a set of children ZIP files, all files will be uncompressed and ready for review, rather than just the parent ZIP file.

□ Digitally identify each file's signature to determine its true file type rather than relying on the file's extension, which is often inaccurate.

□ Remove operating system and software application files.

□ Cull your data based on file type, custodian, date and/or predetermined search terms.

□ Flag duplicate files so they may be filtered out during review. You should be able to produce all original native files, including duplicates you may have filtered.

□ Extract file content and metadata from the files (e.g., e-mail routing data, file property data and other "hidden" data that can be relevant during discovery).

- □ Convert all of the electronic documents to common viewable format (HTML, TIFF or PDF) to create a consistent format for data review.

At this stage, you have the option of merging your electronic documents with paper documents that have been scanned, coded and/or optically character recognized (OCR'd).

tip — bates numbering capabilities

In anticipation of production, you'll want to have a clear understanding of the capabilities, and limitations, of your vendor's Bates numbering capabilities. Be sure that you can assign multiple Bates numbers to an electronic document in case you'll need to produce it in multiple matters or to multiple recipients. This enables the reuse of information that has already been reviewed for privilege and relevancy, and this will save you time in the data review stage. You'll also want to know the legal matter it was reviewed for and what past production sets it was included in.

e-mail preparation

All of a computer user's e-mail messages and attachments are stored in the form of self-contained mailbox files that typically reside on network servers, though may also reside on the data custodian's desktop computer. Difficulties can arise if mailbox files are collected and reviewed incorrectly. For example, if an attorney reviews these files from within an e-mail software program, critical content and metadata may change. Further, if the collected mailbox contained unsent messages in the "Outbox," an attorney reviewing the mailbox may inadvertently send the messages.

Efficiency issues often arise due to the sheer bulk of e-mail communications in today's workplace. If your client has a significant amount of e-mail, it is highly inefficient to open one custodian's e-mail message box at a time. If you're opening messages from within the e-mail software application, you can't do global searches on the attachments associated with the messages. Although you can search the e-mail message content, the attachments provide valuable information that you'll need to access for full data review.

tip — document formatting surprises

Document formatting issues can raise a number of challenges that your electronic discovery vendor should be able to address:

- If a document has multiple print formats within it (such as switching from portrait to landscape and back), you'll need to be sure that all data is going to be captured correctly.
- You also need to be sure your vendor is able to make visible the hidden rows, columns and worksheets that are often used in spreadsheet files.
- Some users create files that have a section that automatically prints or displays "today's date" and/or "path/filename" of the document. This can cause confusion during the data review process as it will print the day of printing as part of the document. Your vendor should offer you the option to redact it, or substitute a phrase of your choosing such as "Today."

litigation support package limitations

Early litigation support software packages were primarily designed to accommodate scans of paper documents. They have subsequently been upgraded to support TIFF images of electronic documents. Some litigation support packages indicate that they manage electronic documents via tagging and searching the information, but you may not be getting what you need. The problem inherent in those litigation support tools is their inability to find information and to Bates stamp and track the TIFF files as stamped documents.

To avoid this logistical hassle, you should consider using technology that will allow you to convert native files into TIFF files so you can Bates-stamp the images and automatically track them.

You should also note that some electronic litigation support packages may have the ability to process e-mail but lack the ability to create TIFF images of the review set. After legal teams complete the review, they may not be able to accurately or efficiently output the results and reliably track production.

tip — taking out the garbage

You may find that your client provides a massive amount of data, yet much of it serves no purpose (such as system and software application files that do not contain relevant content). You can save significant amounts of time by having your client’s files cataloged so a fast cull can be done to filter those unneeded files. Well-structured, efficient preculling and categorization of data via search terms are highly recommended methods prior to creating the data review/control set.

time zone challenges

When preparing data, your electronic discovery vendor must be able to manage date and time complexities. You may often need to combine e-mails from various offices throughout the nation or the world. This introduces time zone issues (even as common as East Coast vs. West Coast). Date/time sequences can be crucial in identifying who knew what when, yet the sequences will be incorrect unless you base it all on a standard such as Greenwich Mean Time (GMT). You’ll also want to be certain that your vendor correctly handles daylight saving time differences. As an added complexity, Windows® 95 stores dates differently than Windows® 2000. Fortunately, all of these date and time issues can be handled by experienced electronic discovery vendors.

Once your client’s data is properly converted into a unified database, you will then be ready to begin the data review process.

data review: fast and thorough methods

During the discovery phase of large or complex cases, there can be significant time and resource burdens associated with paper document review. Reviewing printed versions of electronic materials is often an inefficient and time-consuming process. It is also subject to risks of omission due to "reviewer's fatigue," resulting in potential misidentification and misclassification of relevant and privileged evidence.

To meet tight electronic discovery response deadlines imposed by courts and government agencies, legal teams should use reliable electronic data review tools that are fast and efficient. Review tools must scale to meet the demands of a project of any size and provide means for tracking progress and ensuring that all materials are reviewed thoroughly and accurately.

ways to reduce search and review time

There is a misperception about electronic discovery that you'll have to read thousands of documents online. In reality, the sophisticated search capabilities available in today's electronic discovery data review tools reduce the document count dramatically, allowing you to decide whether you want to review the final documents on screen or on paper.

To facilitate the most accurate and efficient data review process, you should consider utilizing a Web-based (online) review tool that enables your legal team to search, organize, categorize, annotate, cull and produce information. The goal is to reduce the time and resources you need to allocate to the search and review process while protecting the integrity of your client's data. Online discovery enables remote, secure access to your data for concurrent review by members of your team.

Sophisticated online review tools also enable instant messaging capability. This helps save time by bridging the geographical gap between reviewers. Your team can immediately clarify document review criteria or other questions by sending a secure instant message to a peer or to all reviewers on the team.

potential privilege review

To start the review process, you'll want to create a list of all counsel names and law firm names so that potentially privileged items can be separated from the data set at the beginning of the review cycle. The greater the number of potentially privileged documents you identify, the greater protection you provide for your client and the faster you can complete the review process.

review for relevance and privilege

You'll then want to review the data set for relevance and privilege by performing the following activities:

Step One: Define Criteria for Relevance and Privilege:
While this may seem like an obvious step, establishing the criteria for categorizing documents as potentially relevant or privileged can have a profound impact on the effectiveness of the data review. The larger the number of reviewers on your project, the more important it becomes to clearly document criteria for categorizations. Any ambiguities in the criteria will prolong the review and may delay the project.

Step Two: Review for Relevance

- □ Do a first-pass review of the documents for relevance, and flag the questionable documents.
- □ Do a second-pass review of the subset of documents that have been marked as potentially relevant.

Step Three: Review for Privilege

- □ Create a privilege log. Sophisticated review tools automatically track privilege information and allow you to generate a privilege log at the end of your review.
- □ Do a first-pass review of documents for privilege.
  - Flag privileged documents and record the privilege claim.
  - Flag partially privileged documents to be redacted later.
- □ Do a second-pass review of documents for privilege.
- □ Redact documents that are partially privileged, and maintain a privilege log.

proper indexing is crucial

A common problem in data review, as well as during data preparation, is a lack of completely indexed data. The quality of the indexing is very important, as it is the key to finding the data you need. A variety of indexing methods are possible, the use of which depends on the nature of your data.

For e-mail, it is best to extract all the e-mail metadata (including the body of the e-mail) and attachments and then build an index for searching. In their original form, e-mail messages and their attachments are contained in a single large “mailbox” file. For example, Microsoft® Exchange mail systems typically store e-mails in a “.PST” file while Lotus Notes® mail systems typically store e-mails in an “.NSF” file. If you were to do a keyword search on a mailbox file, you would not be able to search the individual e-mail and attachments within that file. Once all the content is extracted from a mailbox, however, each individual e-mail message, all its metadata and any attachments are available for searching. For example, you can then pull up only those documents that were sent or received by a particular person.

For all other files, it is best to do native file indexing, because those files do not have the same limitations that affect e-mail files.

the right search terms

One pitfall in the review stage is the use of overly broad search terms (such as “office,” “company name,” “contract,” etc.), which can result in a massive number of irrelevant hits. You need to ensure that your search parameters and mechanism appropriately narrows the resulting data set. You’ll want to be able to do things like proximity searches (such as looking for the word “contract” within a number of words of another word, such as “contract w/5 of employment”). You’ll also want your review tool’s search engine to do specific subsearches within metadata fields such as “to,” “from” and “subject” for e-mails, and Microsoft® Office “properties” metadata such as “Author” or “Title.”

Using electronic discovery data review tools, data searches can be done on an impressive number of terms. Advanced tools can do simultaneous searches in excess of 2,500 characters. Depending on word count, that can average between 300 and 500 words during a single search. Make sure that your search tool allows you to exclude certain sets of data from your search. This is critical when your data is flowing into the system at different times.

data file extensions — there's more than meets the eye

Unless you're using appropriate search technology, data file extensions can be problematic during the review stage. As an example, the file extension for a Microsoft® Word document, which is traditionally ".DOC," can be changed to other file extensions such as ".EXE," ".PSD," ".XLS," etc. It is very simple for a data custodian who does not want information to be easily discovered to alter file extensions within his or her data set. Some programs even change file extensions on their own. For instance, when "auto-saving" documents, some word processing programs save temporary, at times hidden, ".TMP" files.

Consider the scenario where Mr. Smith alters the file extensions on a set of confidential documents. Later, the data reviewer who gathers Mr. Smith's computer files and searches the entire data set for "*.DOC" (which is supposed to turn up all word processing document file types) will not find all document files. The benefit of using sophisticated data review tools is that they will search through all content, thus eliminating the risk of incorrect file extensions.

tip — the importance of e-mail chains

To get full visibility into your data, you'll need access to complete e-mail chains, including the original e-mail and all of its attachments. This is critically important because you often need to produce the entire chain and therefore need to review a chain in its entirety to determine if any single item is confidential or privileged.

the time crunch

Despite the best intentions, litigation teams don't always allow enough time during the planning process for data review, which can cause a problematic time crunch. If you are in a situation where you need to quickly assign additional document reviewers to your project, you'll want to be sure your vendor doesn't charge you on a per-seat license basis, otherwise your costs will go up. Potentially more detrimental than the extra seat fees could be the delays in purchasing the licenses, setting up the new accounts and/or installing new software.

Some law firms use data review tools that are provided on a per-seat basis and share individual passwords among multiple data reviewers. Aside from potentially violating software license agreements, such activity results in reduced security and data reliability. You should be able to authenticate exactly who had access to confidential client data and who reviewed each document. Every system user should be accessing the data set via his or her unique account and password.

tip — what you don't see on paper

When electronic information is directly converted to paper for review, much information can be lost and therefore never seen. Things such as metadata, linkages between e-mail messages and attachments, hidden or changed text in word processing documents, formulas and hidden rows and columns in spreadsheets are not printed to paper. Therefore, you'll want to utilize the appropriate electronic discovery tools and processes.

After your litigation team has performed a detailed review of your client's data, you'll be ready to produce a set of documents to other parties.

data production: speed, flexibility and accuracy

Once your team has reviewed all documents associated with the discovery process, the relevant nonprivileged data set must be delivered to parties such as opposing counsel, partner firms, outside counsel or the requesting government agency. You'll want to be sure that your electronic discovery vendor can provide a variety of delivery options, including Web repositories, exports to other database environments, native files and paper.

gap-free bates production

It is common for files to require some level of manual intervention for successful processing. Files that are encrypted, password-protected and/or those containing macros need special attention to make them print-ready. The challenge during production is that you'll want all of the documents produced in a specific order to maintain an appropriate Bates numbering sequence. Therefore, rather than printing directly from a file, it is best to convert the file into TIFF format and then print it. Any problematic files will be discovered during the TIFF creation stage and can be addressed prior to printing. This prevents confusion, eliminates labor-intensive assembly and assures consistent, gap-free Bates numbering.

readability and fonts

Having the appropriate fonts loaded into the production system is vital to properly rendering and producing the wide variety of files generally found in electronic data sets. Documents are often created using a surprising number of standard and nonstandard fonts. If a font is not available, font substitution occurs that can result in visual distortion such as overlays and other alterations. You'll want to be sure that all fonts are available so documents can be produced in the original format used by the data custodian.

tip — data production complexities

Your electronic discovery service vendor can help you identify which sorts of files are likely to "fail" during the production step. Complications such as password protection, macros or presentation files that are too large are common examples that can cause production limitations. You'll want to make sure that your vendor can handle file complexities and provide detailed exception reporting for file output and, if necessary, export native files for review and production.

multiple production options

A crucial aspect of production is accurately and reliably tracking document productions so you know what documents went to which recipients on specific dates. All production information should be tracked within the vendor's database. Your vendor should also provide you with the option of producing the same documents to multiple recipients, with separate tracking numbers for each, and, if desired, with different Bates numbers. Make sure your vendor allows you to search for a document based on a Bates number, so if you are presented with an individual page at deposition time, you can retrieve the entire document to view the page in context and make sure it is what you produced.

problems with do-it-yourself production

What you print does not always capture everything in a computer file. Often when producing spreadsheets, the selected print area within a document does not contain all the content of the file. You'll need to be sure that any hidden rows, columns and worksheets are unhidden. Spreadsheets also pose the problem of cell contents that are too large for the column or row, resulting in truncated text or placeholder content for numerical data that appears as "######." To prevent this, you'll need the cells to be resized so all the content is visible. Your electronic discovery vendor should offer the option of performing automatic electronic formatting to reveal such content in spreadsheets. After all the content has been made visible, it is then appropriate to convert the file to HTML,TIFF or PDF.

producing the right documents

You don't want to inadvertently produce privileged or irrelevant documents to the opposition. Mishaps at this stage can be devastating to your case. Be certain that your vendor has strict quality control methods in place for document production. Production project managers should clearly understand the data review plan and the methods for data production. You'll want to be confident that your data production is fast, clean and accurate.

glossary of electronic discovery terms

active data
Data currently displayed on a *computer* screen, and/or files on a *computer* that can be accessed without having to use a restoration process.

application
See *software application*.

archive
A copy of *data* on a *computer* drive, or on a portion of a drive, maintained for historical reference.

attachment
A memorandum, letter, spreadsheet or any other electronic document appended to another *document* or *e-mail*.

backup
A copy of inactive *data*, intended for use in the restoration of data.

Boolean search
The term "Boolean" refers to a system of logic developed by an early *computer* pioneer, George Boole. In Boolean searching, an "and" operator between two words results in a search for documents containing both of the words. An "or" operator between two words creates a search for documents containing either of the target words. A "not" operator between two words creates a search result containing the first word but excluding the second.

cache
A form of high-speed *memory* used to temporarily store frequently accessed information; once the information is stored, it can be retrieved quickly from *memory* rather than from the *hard drive*.

case de-duplication
Retains only single copies of *documents* per case. For example, if an identical *document* resides with Mr. A, Mr. B and Mr. C, only the first occurrence of the file will be saved (Mr. A's). Contrast with *custodian de-duplication* and *production de-duplication*.

cluster
In *operating systems* that use a file allocation table (FAT) architecture, the smallest unit of *storage* space required for data written to a drive. Also called an allocation unit.

computer
Includes but is not limited to network servers, desktops, laptops, notebook computers, mainframes, PDAs (personal digital assistants, such as PalmPilot™), and other digital communication devices.

compression
A technology for storing *data* in fewer bits, it makes data smaller so less disk space is needed to represent the same information. Compression programs such as WinZip and UNIX Compress are valuable to network users because they save both time and bandwidth. Data compression is also widely used in backup utilities, spreadsheet applications and database management systems.

custodian
See *data custodian*.

custodian de-duplication
Culls a *document* if multiple copies of that *document* reside within the same custodian’s data set. For example, if Mr. A and Mr. B each has a copy of a specific *document*, and Mr. C has two copies, the system will maintain one copy each for Mr. A, Mr. B and Mr. C. Contrast with *case de-duplication* and *production de-duplication*.

customer-added metadata
Data or work product created by a user while reviewing a *document*. For example, annotation text of a document or subjective coding information. Contrast with *vendor-added metadata*.

data
Any information stored on a *computer*.

data custodian
Person having administrative control of a *document*; for example, the data custodian of an e-mail is the owner of the mailbox that contains the e-mail. Often referred to as “user” by IT personnel, sometimes referred to as “source” by legal review teams.

data formats
The organization of information for display, storage or printing. *Data* is maintained in certain common formats so that it can be used by various programs, which may only work with data in a particular format. This term is commonly used in the industry when asking another person about the state in which particular information exists. For example, "What format is it in, *PDF* or *HTML*?"

deleted file
A *file* with disk space that has been designated as available for reuse. Although a user may "erase" or "delete" a file, all that is really erased is a reference to that file in a table on the hard disk. Unless overwritten with new data, a "deleted" file can be as intact on the disk as any "active" file you would see in a directory listing.

de-duplication
The process of identifying (or for some vendors includes actually removing) additional copies of identical *documents* in a *document* collection. There are three types of de-duplication: *case*, *custodian* and *production*.

digital certificate
A means of providing heightened security for the access of a Web site or a specific *document*. Digital certificates are electronic records that contain keys used to decrypt information, especially information sent over a public network like the Internet. Digital certificates must be applied for and granted by a certificate authority (CA).

document
Any *file* produced by a *software application*.

document metadata
Data stored in the *document* about the *document*. Often this data is not immediately viewable in software application used to create/edit the *document* but often can be accessed via a "Properties" view. Contrast with *file system metadata* and *e-mail metadata*.

e-mail
Electronic mail, or computer-based mail.

e-mail address
An electronic mail address. E-mail addresses follow the formula: user-ID@domain-name. In some e-mail systems, a user's e-mail address is "aliased" or represented by his or her natural name rather than a fully qualified e-mail address. For example, john.doe@abc.com might appear simply as John Doe.

e-mail attachment
See *Attachment*.

e-mail metadata
Data stored in the *e-mail* about the e-mail. Often this *data* is not even viewable in the e-mail client application used to create the e-mail. The amount of e-mail

metadata available for a particular e-mail varies greatly depending on the e-mail system. Contrast with *file system metadata* and document metadata.

encryption
A technology that renders the contents of a file unintelligible to anyone not authorized to read it. Encryption is used to protect information as it moves from one *computer* to another and is an increasingly common way of sending credit card numbers over the Internet when conducting e-commerce transactions.

FENS® number
Fios Electronic Numbering System — a unique *document*-level identifier (as opposed to a page-level identifier such as a Bates number) for each *document* in a matter.

file
An element of *data* storage in a *file system*. A collection of *data* or information that has a name, called the *filename*. Almost all information stored in a *computer* must be in a file. There are many different types of files: data files, text files, program files, directory files and so on.

file system
The system that an *operating system* or program uses to organize and keep track of *files*. For example, a hierarchical *file* system is one that uses directories to organize *files* into a tree structure. Types of *file* systems include *file* allocation table (FAT) and Windows® NT *file* system (NTFS).

file system metadata
Data that can be obtained or extracted about a *file* from the *file system* storing the *file*. Contrast with *document metadata* and *e-mail metadata*.

filename
The name of a *file*. All files have names. Different *operating systems* impose different restrictions on filenames. Most operating systems, for example, prohibit the use of certain characters in a filename and impose a limit on the length of a filename. In addition, many systems, including DOS and UNIX, allow a *filename extension* that consists of one or more characters following the proper filename. The *filename extension* usually indicates what type of file it is.

filename extension
In DOS and some other *operating systems*, one or several letters at the end of a *filename*. Filename extensions usually follow a period (dot) and indicate the type of information stored in the *file*. For example, in the *filename* LETTER.DOC, the extension is DOC, which indicates that the file is a word processing file.

hard drive

The primary hardware that a *computer* uses to store information, typically magnetized media on rotating discs.

HTML

HyperText Markup Language, a language that uses tags to structure text into headings, paragraphs, lists and links. It tells a Web browser how to display text and images.

imaged copy

A "mirror image" or bit-by-bit copy of a *hard drive*, i.e., a complete replication of the physical drive. From an imaged copy of a *hard drive* it is possible to reconstruct the entire contents and organization of the source drive from which it was taken.

input device

Any object that allows a user to communicate with a *computer* by entering information or issuing commands (e.g., keyboard, mouse or joystick).

magnetic or optical storage media

Includes, but is not limited to, *hard drives* (also known as "hard disks"), backup tapes, optical disks, CD-ROMs, DVD-ROMs, Jaz and *Zip drives* and floppy discs.

mailbox

An area in *memory* or on a storage device where *e-mail* is placed. In e-mail systems, each user has a private mailbox. When the user receives e-mail, the mail system automatically puts it in the mailbox. The mail system allows you to scan mail that is in your mailbox, copy it to a file, delete it, print it or forward it to another user. The mailbox format used by Microsoft® Exchange e-mail systems is PST, while Lotus Notes® uses NSF files.

memory

Internal storage areas in the *computer*. The term memory identifies *data* storage that comes in the form of chips, and the word storage is used for *memory* that exists on tapes or disks. Moreover, the term *memory* is usually used as shorthand for *physical memory*, which refers to the actual chips capable of holding *data*. Some *computers* also use virtual memory, which expands physical memory onto a hard disk. See the definitions for two types of physical memory: *RAM* and *ROM*.

metadata

Data about *data*. In *data* processing, metadata provides information about a *document* or other *data* managed within an application or environment. There are five types of metadata: *file system*, *document*, *e-mail*, *vendor-added* and *customer-added*.

native file/format
The source *document*, as collected from the source *computer* or server, before any conversion or processing of the *document*.

network
A group of connected *computers* that allow people to share information and equipment (e.g., local area network [LAN], wide area network [WAN], metropolitan area network [MAN], storage area network [SAN], peer-to-peer network and client-server network).

OCR
Optical character recognition, a method of translating printed text and images into a form that a *computer* can manipulate (into ASCII codes, for example). An OCR system enables you to scan a printed *document* directly into a *computer* file.

operating system
Software that directs the overall activity of a *computer* (e.g., MS-DOS®, Windows®, Linux®, etc.).

network operating system
Software that directs the overall activity of networked *computers* (e.g., Novell™, Microsoft® Windows NT®, UNIX™, etc.).

PDF
Portable document format – a *file* format developed by Adobe Systems.
PDF captures formatting information from a variety of desktop publishing applications, making it possible to send formatted *documents* and have them appear on the recipient's monitor or printer as they were intended. To view a *file* in PDF format, you need Adobe® Acrobat® Reader, a free application distributed by Adobe Systems.

production de-duplication
Culling of a *document* if multiple copies of that *document* reside within the same production set. For example, if two identical *documents* are both marked responsive, nonprivileged, production de-duplication ensures that only one of those *documents* is produced. Contrast with *case de-duplication* and *custodian de-duplication*.

RAM

Random access memory – the hardware inside a *computer* that retains memory on a short-term basis and stores information while the user utilizes the *computer*. RAM is erased when a *computer* is turned off.

ROM

Read-only memory – the hardware in a *computer* that can be read but not written to. ROM can contain the programming that allows a *computer* to boot up each time the user turns it on, and essential system programs that neither the user nor the *computer* can erase.

slack

The difference in empty bytes of the space that is allocated in clusters minus the actual size of the *files*. Also described as the *data* fragments stored randomly on a *hard drive* during the normal operation of a *computer*, or the residual *data* left on the *hard drive* after new *data* has overwritten some of the previously stored *data*.

software

Any set of instructions stored on *computer*-readable media that tells a *computer* what to do. Includes *operating systems* and *software applications*.

software application

A program that instructs a *computer* to perform a specific set of instructions or execute a process. Some software applications are user-driven, such as Microsoft® Word or Notepad, while others are system-driven, such as the Windows® system clock or automatic virus-scanning programs.

storage device

Any device that a *computer* uses to store information.

storage media

Any removable devices that store *data*. See *magnetic* or *optical storage media*.

tape drive

A hardware device used to store *data* on a magnetic tape. Tape drives are usually used to back up large quantities of *data* due to their large capacity and cheap cost relative to other *data* storage options.

TIFF
Tagged image file format – a graphic *file* format used for storing still-image bitmaps. TIFFs are stored in tagged fields, and programs use the tags to accept or ignore fields, depending on the application.

vendor-added metadata
Data created and maintained by the electronic discovery vendor as a result of processing the *document*. While some vendor-added metadata has direct value to customers, much of it is used for process reporting, chain of custody and *data* accountability. Contrast with *customer-added metadata*.

Zip® Drive
A brand-name magnetic *storage device* that can hold between 100 and 250 megabytes of *data*.

appendix a: form spoliation letter to opposing counsel

[date]

[address]

re: [matter (, case number)]

Dear: ___________,

By this letter, you and your client{s} are hereby given notice not to destroy, conceal or alter any paper or electronic files and other data generated by and/or stored on your client's {clients'} computers and storage media (e.g., hard disks, floppy disks, backup tapes), or any other electronic data, such as voice mail. As you know, your client's {clients'} failure to comply with this notice can result in severe sanctions being imposed by the Court {and liability in tort} for spoliation of evidence or potential evidence.

Through discovery we expect to obtain from you a number of documents and things, including files stored on your client's {clients'} computers and your client's {clients'} computer storage media. {As part of our initial discovery efforts, you [are hereby served with/will soon receive] [initial/supplemental] interrogatories and requests for documents and things.}

In order to avoid spoliation, you will need to provide the data requested on the original media. Do not reuse any media to provide this data.

{Although [we may bring/have brought] a motion for an order preserving documents and things from destruction or alteration, your client's {clients'} obligation to preserve documents and things for discovery in this case arises in law and equity independently from any order on such motion.}

Electronic documents and the storage media on which they reside contain relevant, discoverable information beyond that which may be found in printed documents. Therefore, even where a paper copy exists, we [seek/will seek] all documents in their electronic form along with information about those documents contained on the media. We also [seek/will seek] paper printouts of only those documents that contain unique information after they were printed out (such as paper documents containing handwriting, signatures, marginalia, drawings, annotations, highlighting and redactions) along with any paper documents for which no corresponding electronic files exist.

Our discovery requests [ask/will ask] for certain data on the hard disks, floppy disks and backup media used in your client's {clients'} computers, some of which data are not readily available to an ordinary computer user, such as "deleted" files and "file fragments." As you may know, although a user may "erase" or "delete" a file, all that is really erased is a reference to that file in a table on the hard disk; unless overwritten with new data, a "deleted" file can be as intact on the disk as any "active" file you would see in a directory listing.

{Courts have made it clear that all information available on electronic storage media is discoverable, whether readily readable ("active") or "deleted" but recoverable. See, e.g., Easley, McCaleb & Assocs., Inc. v. Perry, No. E-2663 (Ga. Super. Ct. July 13, 1994; "deleted" files on a party's computer hard drive held to be discoverable, and plaintiff's expert was allowed to retrieve all recoverable files); Santiago v. Miles, 121 F.R.D. 636, 640 (W.D.N.Y. 1988; a request for "raw information in computer banks" was proper and obtainable under the discovery rules); Gates Rubber Co. v. Bando Chemical Indus., Ltd., 167 F.R.D. 90, 112 (D. Colo. 1996; mirror-image copy of everything on a hard drive "the method which would yield the most complete and accurate results," chastising a party's expert for failing to do so); and Northwest Airlines, Inc. v. Teamsters Local 2000, et al., 163 L.R.R.M. (BNA) 2460, (USDC Minn. 1999); court ordered image-copying by Northwest's expert of home computer hard drives of employees suspected of orchestrating an illegal "sick-out" on the Internet).}

Accordingly, electronic data and storage media that may be subject to our discovery requests and that your client{s} are obligated to maintain and not alter or destroy, include but are not limited to the following:

> Introduction: description of files and file types sought
>
> All digital or analog electronic files, including "deleted" files and file fragments, stored in machine-readable format on magnetic, optical or other storage media, including the hard drives or floppy disks used by your client's {clients'} computers and their backup media (e.g., other hard drives, backup tapes, floppies, Jaz cartridges, CD-ROMs) or otherwise, whether such files have been reduced to paper printouts or not. More specifically, your client{s} is {are} to preserve all of your e-mails, both sent and received, whether internally or externally; all word-processed files, including drafts and revisions; all spreadsheets, including drafts and revisions; all databases; all CAD (computer-aided design) files, including drafts and revisions; all presentation data or slide shows produced by presentation software (such as Microsoft PowerPoint); all graphs, charts and other data produced by project management software (such as Microsoft Project); all data generated by calendaring, task management and personal information management (PIM)

software (such as Microsoft Outlook or Lotus Notes); all data created with the use of personal data assistants (PDAs), such as PalmPilot, HP Jornada, Cassiopeia or other Windows CE-based or Pocket PC devices; all data created with the use of document management software; all data created with the use of paper and electronic mail logging and routing software; all Internet and Web-browser-generated history files, caches and "cookies" files generated at the workstation of each employee and/or agent in your client's {clients'} employ and on any and all backup storage media; and any and all other files generated by users through the use of computers and/or telecommunications, including but not limited to voice mail. Further, you are to preserve any log or logs of network use by employees or otherwise, whether kept in paper or electronic form, and to preserve all copies of your backup tapes and the software necessary to reconstruct the data on those tapes, so that there can be made a complete, bit-by-bit "mirror" evidentiary image copy of the storage media of each and every personal computer (and/or workstation) and network server in your control and custody, as well as image copies of all hard drives retained by you and no longer in service, but in use at any time from _______________ to the present.

Your client{s} is {are} also not to pack, compress, purge or otherwise dispose of files and parts of files unless a true and correct copy of such files is made.

Your client{s} is {are} also to preserve and not destroy all passwords, decryption procedures (including, if necessary, the software to decrypt the files); network access codes, ID names, manuals, tutorials, written instructions, decompression or reconstruction software, and any and all other information and things necessary to access, view and (if necessary) reconstruct the electronic data we [are requesting/will request] through discovery.

1. Business Records: [All documents and information about documents containing backup and/or archive policy and/or procedure, document retention policy, names of backup and/or archive software, names and addresses of any offsite storage provider.]

a. All e-mail and information about e-mail (including message contents, header information and logs of e-mail system usage) {sent or received} by the following persons:

[list names, job titles]

b. All other e-mail and information about e-mail (including message contents, header information and logs of e-mail system usage) containing information about or related to:

[insert detail]

c. All databases (including all records and fields and structural information in such databases), containing any reference to and/or information about or related to:

[insert detail]

d. All logs of activity (both in paper and electronic formats) on computer systems and networks that have or may have been used to process or store electronic data containing information about or related to:

[insert detail]

e. All word processing files, including prior drafts, “deleted” files and file fragments, containing information about or related to:

[insert detail]

f. With regard to electronic data created by application programs which process financial, accounting and billing information, all electronic data files, including prior drafts, “deleted” files and file fragments, containing information about or related to:

[insert detail]

g. All files, including prior drafts, “deleted” files and file fragments, containing information from electronic calendars and scheduling programs regarding or related to:

[insert detail]

h. All electronic data files, including prior drafts, “deleted” files and file fragments about or related to:

[insert detail]

2. Online Data Storage on Mainframes and Minicomputers: With regard to online storage and/or direct access storage devices attached to your client's {clients'} mainframe computers and/or minicomputers: they are not to modify or delete any electronic data files, "deleted" files and file fragments existing at the time of this letter's delivery, which meet the definitions set forth in this letter, unless a true and correct copy of each such electronic data file has been made and steps have been taken to assure that such a copy will be preserved and accessible for purposes of this litigation.

3. Offline Data Storage, Backups and Archives, Floppy Diskettes, Tapes and Other Removable Electronic Media: With regard to all electronic media used for offline storage, including magnetic tapes and cartridges and other media that, at the time of this letter's delivery, contained any electronic data meeting the criteria listed in paragraph 1 above: Your client {clients} is {are} to stop any activity that may result in the loss of such electronic data, including rotation, destruction, overwriting and/or erasure of such media in whole or in part. This request is intended to cover all removable electronic media used for data storage in connection with their computer systems, including magnetic tapes and cartridges, magneto-optical disks, floppy diskettes and all other media, whether used with personal computers, minicomputers or mainframes or other computers, and whether containing backup and/or archive data sets and other electronic data, for all of their computer systems.

4. Replacement of Data Storage Devices: Your client {clients} is {are} not to dispose of any electronic data storage devices and/or media that may be replaced due to failure and/or upgrade and/or other reasons that may contain electronic data meeting the criteria listed in paragraph 1 above.

5. Fixed Drives on Stand-Alone Personal Computers and Network Workstations: With regard to electronic data meeting the criteria listed in paragraph 1 above, which existed on fixed drives attached to stand-alone microcomputers and/or network workstations at the time of this letter's delivery: Your client {clients} is {are} not to alter or erase such electronic data, and not to perform other procedures (such as data compression and disk de-fragmentation or optimization routines) that may impact such data, unless a true and correct copy has been made of such active files and of completely restored versions of such deleted electronic files and file fragments, copies have been made of all directory listings (including hidden files) for all directories and subdirectories containing such files, and arrangements have been made to preserve copies during the pendency of this litigation.

6. Programs and Utilities: Your client {clients} is {are} to preserve copies of all application programs and utilities, which may be used to process electronic data covered by this letter.

7. Log of System Modifications: Your client {clients} is {are} to maintain an activity log to document modifications made to any electronic data processing system that may affect the system's capability to process any electronic data meeting the criteria listed in paragraph 1 above, regardless of whether such modifications were made by employees, contractors, vendors and/or any other third parties.

8. Personal Computers Used by Your Employees and/or Their Secretaries and Assistants: The following steps should immediately be taken in regard to all personal computers used by your client's {clients'} employees and/or their secretaries and assistants.

a. As to fixed drives attached to such computers: (i) a true and correct copy is to be made of all electronic data on such fixed drives relating to this matter, including all active files and completely restored versions of all deleted electronic files and file fragments; (ii) full directory listings (including hidden files) for all directories and subdirectories (including hidden directories) on such fixed drives should be written; and (iii) such copies and listings are to be preserved until this matter reaches its final resolution.

b. All floppy diskettes, magnetic tapes and cartridges, and other media used in connection with such computers prior to the date of delivery of this letter containing any electronic data relating to this matter are to be collected and put into storage for the duration of this lawsuit.

9. Evidence Created Subsequent to This Letter: With regard to electronic data created subsequent to the date of delivery of this letter, relevant evidence is not be destroyed and your client {clients} is {are} to take whatever steps are appropriate to avoid destruction of evidence.

In order to assure that your and your client's {clients'} obligation to preserve documents and things will be met, please forward a copy of this letter to all persons and entities with custodial responsibility for the items referred to in this letter.

Sincerely, etc.

appendix b: sample electronic discovery interrogatories and requests for production

Below are suggested interrogatories and requests for production that are meant to be complementary (i.e., any devices or electronic files that are identified in answer to an interrogatory or interrogatories are usually immediately requested in the follow-up request[s] for production).

For more detailed questions that you might want to include in interrogatories rather than in a deposition, see the sample deposition questions.

Sample Interrogatories and Requests for Production

[Note: The precise format for the following suggested interrogatories and requests for production of documents and things should be in accordance with the applicable civil and local rules of the court where the matter is filed.]

[suggested language for inclusion in preamble:]

I. Definitions

For the purposes of the following interrogatories and requests for production of documents and things, the following definitions apply:

Application Software: A set of electronic instructions, also known as a program, which instructs a computer to perform a specific set of processes.

Archive: A copy of data on a computer drive, or on a portion of a drive, maintained for historical reference.

Backup: A copy of active data, intended for use in restoration of data.

Computer: Includes but is not limited to network servers, desktops, laptops, notebook computers, employees' home computers, mainframes, the PDAs of [party name] and its employees (personal digital assistants, such as PalmPilot, Cassiopeia, HP Jornada and other such handheld computing devices), digital cell phones and pagers.

Data: Any and all information stored on media that may be accessed by a computer.

Digital Camera: A camera that stores still or moving pictures in a digital format (TIFF, GIF, etc.).

Document: Includes but is not limited to any electronically stored data on magnetic or optical storage media as an "active" file or files (readily readable by one or more computer applications or forensics software); any "deleted" but recoverable electronic files on said media; any electronic file fragments (files that have been deleted and partially overwritten with new data); and slack (data fragments stored randomly from random access memory on a hard drive during the normal operation of a computer [RAM slack] or residual data left on the hard drive after new data has overwritten some but not all of previously stored data).

Hard Drive: The primary hardware that a computer uses to store information, typically magnetized media on rotating disks.

Help Features/Documentation: Instructions that assist a user on how to set up and use a product including but not limited to software, manuals and instruction files.

Imaged Copy: A "mirror image" bit-by-bit copy of a hard drive (i.e., a complete replication of the physical drive).

Input Device: Any object that allows a user to communicate with a computer by entering information or issuing commands (e.g., keyboard, mouse or joystick).

Magnetic or Optical Storage Media: Include but are not limited to hard drives (also known as "hard disks"), backup tapes, CD-ROMs, DVD-ROMs, JAZ and Zip drives, and floppy disks.

Network: A group of connected computers that allow people to share information and equipment (e.g., local area network [LAN], wide area network [WAN], metropolitan area network [MAN], storage area network [SAN], peer-to-peer network, client-server network).

Operating System: Software that directs the overall activity of a computer (e.g., MS-DOS, Windows, Linux).

Network Operating System: Software that directs the overall activity of networked computers.

Software: Any set of instructions stored on computer-readable media that tells a computer what to do. Includes operating systems and applications.

Storage Devices: Any device that a computer uses to store information.

Storage Media: Storage media are any removable devices that store data.

II. Spoliation: getting information on preservation of information.

S1. Written policies on preservation of records

Interrogatory No.______:

Do you have a written policy for the retention of documents, including but not limited to business records?

Request for Production No.______:

Please produce copies of any and all written policies for the retention of documents, for the time period of ________ to __________ inclusive.

S2. Destruction of documents

Interrogatory No.______:

Do you have a written policy for the destruction of documents, including but not limited to business records?

Request for Production No.______:

Please produce copies of any and all written policies for the destruction of documents, for the time period of ________ to __________ inclusive.

Interrogatory No._____:

Has destruction or overwriting of documents been suspended? If so, on what date did suspension begin?

S3. Persons in charge of maintaining document retention and destruction policies

Interrogatory No.______:

Identified by job title, job description and business address and telephone number, who are all persons in charge of implementing the policies identified in your answer to Interrogatories 1 and 2 above?

Interrogatory No.______ :

If not the same person(s) as identified in your answer to the immediately preceding interrogatory, identify by job title, job description, and business address and telephone number, the person at [party name] who is the most knowledgeable about the retention and destruction of documents at [party name]?

Interrogatory No.______:

With respect to preventing the spoliation of documents and things that may potentially become evidence in litigation, please identify with particularity and in detail:

a. Whether the minutes of the meetings of the Board of Directors, from [date] to [date] contain any references to considerations or discussions of preventing such spoliation of potential evidence.

b. If so, state the dates of the meetings for which minutes were taken.

c. If so, state the name, title, job description, business address and telephone number of the person or persons with custody of those minutes.

Request for Production No.______:

Please produce all documents referenced in the immediately preceding interrogatory.

S4. Preservation of evidence

Interrogatory No.______:

Since [date of opposing party's awareness of client's claim or counterclaim, if not date of complaint, cross-claim or counterclaim], have any documents at [party name] been destroyed? If so, please state which electronic files have been deleted from the magnetic or optical storage media of [party name] or overwritten from that date to the present, and dates of destruction or overwriting.

S5. Storage of documents

Interrogatory No.______:

As to the storage of data generated by the users of your computers (such as word-processed files and e-mail), please state whether:

A. The data are backed up on tape or other media?

1. If so:
   - i. How many such media currently exist with backup data on them?
   - ii. What is the maximum storage size in megabytes for each such media?
   - iii. What is the brand name for each such media?
   - iv. When was the last time each such media was backed up with data?
   - v. What software, including brand name and version number, was used to back up each such tape?
   - vi. What was the computer or other hardware (e.g., individual workstation, server) for each such backup?
   - vii. With respect to the immediately foregoing question, state the physical location and current user of each computer or other hardware listed.

Request for Production No.______:

Please produce all backup and/or archive media, for the time period of ________ to __________ inclusive.

III. Data Universe – identifying it

Interrogatory No.______:

Does or did [party name] maintain, or contract with another party to maintain, an overall inventory of data resources such as a Year 2000 Plan or Disaster Recovery Plan? If so, please provide the name, address, phone number and other contact information for the individuals primarily responsible for maintenance of the inventory and/or plan.

Request No.______:

Produce any and all company organizational and policy information in its entirety, including but not limited to organizational charts, corporate policy and procedure manuals, policy memoranda, system schematic, network topology, system restart procedures, e-mail retention policies, Year 2000 Plan, Disaster Recovery Plan, and other related items.

IV. Information personnel

Interrogatory No.______:

Provide a list of all personnel responsible for maintaining computer hardware, data or information systems on computers for [party name]. Include name, position title, contact information, and official job description and list of duties.

Request No.______:

Produce all formal and informal contact lists and duty rosters for personnel in Information Technology (IT) and Information Services (IS), or equivalent divisions within [party name]. Specifically include rosters for groups such as Incident Response Teams, Data Recovery Units, Audit/Investigation Teams, etc.

Request No.______:

Produce all formal job descriptions, assignments and personnel lists for IT and IS personnel, including revisions, for the period _____ to _____.

V. Loose media (including Backup and Archive)

Interrogatory No.______:

Does [party name] maintain a policy regarding use of loose or removable media in its workstations, computers or networks? If so, state the name of the person(s) responsible for creating and enforcing that policy.

Request No.______:

Provide a copy of the policy mentioned in the preceding interrogatory, as well as any revisions, records or logs related to formulation or enforcement of that policy for the period _____ to _____.

Request No.______:

Produce any and all devices used to place information on loose or removable storage media, including but not limited to hard drives, floppy drives, CD-ROM drives, tape drives, recordable DVD-ROM drives, and removable drives (e.g., Jaz, Syjet, Zip, SuperDisk). Include all instructions for use and maintenance of those devices.

Request No.______:

Produce any and all loose or removable media used to store data, including but not limited to floppy disks, CD-ROM discs and tape drive cartridges, that have been used by personnel or contractors of [party name] to perform work for [party name].

Request No.______:

Produce any and all backup and/or archived data [describe scope of data].

Request No.______:

All slack, wherever located, even if media contains nonproduced data.

VI. Computer hardware

Interrogatory No.______:

List all computer equipment provided by [party name] or used by employees of [party name] to perform work for [party name], including but not limited to hardware and/or peripherals attached to a computer such as computer cases [desktop, tower, portable/batteries, all-in-one], monitors, modems [internal, external], printers, keyboards, printers, scanners, mice [cord and cordless], pointing devices [joystick, touchpad, trackball] and speakers. Include description of equipment, serial number, all users for the period _____ to _____ and dates used, and all locations where the equipment was located for the period _____ to _____.

Interrogatory No.______:

Will [party name] permit, without an order therefore, inspection of the equipment mentioned in the preceding interrogatory?

Request No. [follow-up, if response to preceding interrogatory is negative] ______:

Please produce the following computers, including their magnetic or optical storage media, for inspection and copying, on or before [date], at the offices of [law firm] at [address]:

[list of computers you want image-copied, previously identified in discovery; alternatively, if you know the computer population is relatively small]:

Please produce your computers, including their magnetic or optical storage media, for inspection and copying, on or before [date], at the offices of [law firm] at [address]:

Interrogatory No.______:

List all hardware components (e.g., motherboard, modem, NIC, etc.) installed internally or externally to the PC(s) used by _____ during the period _____ to _____.

Request No.______:

Provide any and all documentation of software and hardware modifications to the PC(s) used by _____ during the period _____ to _____, including but not limited to modification dates, software/hardware titles and version numbers, names of persons performing modifications, location of any backup of the data on the computer performed prior to modification, and disposition of replaced software and hardware.

Request No.______:

Produce any and all documentation instructing in setup and use of the PC(s) used by _____ during the period _____ to _____, and hardware and software installed on that/those PC(s). Include any and all documentation reflecting communication with a computer professional or help desk for help in setting up and using the PC(s).

Interrogatory No.______:

List discarded or replaced hardware and software for the PC(s) (including entire PCs) used by _____ during the period _____ to _____. If the hardware or software is no longer in your control, then include the name and contact information of last known custodian.

VII. Computer Software

Request No.______:

Produce any and all software installed or used on the PC(s) used by _____ during the period ____ to _____. Include all titles and version numbers. Include authors and contact information for authors of custom or customized software. Include operating system(s) software.

VIII. Operating Systems

Interrogatory No.______:

List all operating systems (including but not limited to UNIX, Windows, DOS, Linux and PDA operating systems) installed on all computers used by [party name], the specific equipment the OS was installed on and the period during which it was installed on the specific equipment.

Request No.______:

Provide copies of all operating system software listed in the preceding interrogatory, and all supporting documentation provided with the software, and any manuals and tutorials acquired by [party name] to support use of the software.

IX. Telephony

Interrogatory No.______:

Do you have any graphic representation of the components of your telephone and voice messaging system, and the relationship of those components to each other, including but not limited to flow charts, videos or photos, and diagrams?

Interrogatory No.______:

If so, where are the documents located? Include logical paths for electronic documents.

Request No.______:

Produce copies of any and all graphic representations of your telephone and voice messaging network, and the relationship of those components to each other, including any revisions, for the period of _____ to _____ inclusive. If the documents are electronic, please produce them in their native form, as they existed at the time they were drafted, based on archive or back-up data.

Interrogatory No.______:

List all telephone equipment provided by [party name] or used by employees of [party name] to perform work for [party name], including but not limited to desktop telephones, cell phones, pagers, PDA and laptop modems, calling cards, telephony software and contact management software. Include description of equipment and software, serial number, all users for the period of _____ to _____ inclusive and dates used, and all locations where the equipment was located for the period of _____ to _____ inclusive.

Interrogatory No.______:

Will [party name] permit, without an order therefore, inspection of the equipment mentioned in the preceding interrogatory?

Request No.______:

Produce any and all voice messaging records including but not limited to caller message recordings, digital voice recordings, interactive voice response unit (IVR/VRV) recordings, unified messaging files, and computer-based voice mail files to or from [specified parties] for the period _____ to _____.

Request No.______:

Produce all phone use records for [party name] including but not limited to logs of incoming and outgoing calls, invoices and contact management records, manually or automatically created or generated for the period from _____ to _____ inclusive.

X. Other sources of electronic evidence

Interrogatory No.______:

List all log files (files with suffixes) found on computers in [party name]'s network, and the equipment and logical path where the log files may be found.

Request No.______:

Provide copies of the following log files: [this is a follow-up request to the preceding interrogatory, issued after the list of log files has been reviewed]

Request No.______:

Produce any and all manual and automatic records of equipment use, including but not limited to fax, access, audit, security, e-mail, printing, error and transmission records.

Interrogatory No.______:

Do any employees of [party name] subscribe to or participate in Internet newsgroups or chat groups in the course of their employment? If so, list all users and the services that they subscribe to or participate in.

Request No.______:

Produce any and all information related to newsgroups or chat groups, including but not limited to names and passwords for each and every service, newsgroup messages, text files and programs used to access messages.

Interrogatory No.______:

Do any employees of [party name] use portable devices in the course of their employment that are not connected to [party name]'s network, and that are not backed up or archived? If so, list all users and the devices they use.

Request No.______:

Produce any and all portable devices not backed up or archived, including but not limited to handheld devices, set-top boxes, notebook devices, CE devices, digital recorders, digital cameras and external storage devices.

Interrogatory No.______:

Does [party name] provide Internet access for any of its employees or has [party name] done so at any time during the period from ____ to _____ inclusive? If so, list the employees who had Internet access, the Internet service provider (ISP) used, and describe the method(s) used to connect to the Internet.

Request No.______:

Produce any and all documentation describing installation and use of hardware and software used by [party name] to provide Internet access for its employees during the period from _____ to _____ inclusive.

Request No.______:

Produce copies of all manuals, policies and other guidelines for employee access and use of Internet resources.

Interrogatory No.______:

Describe any restrictions on, controls over or monitoring of employee use of Internet resources.

Request No.______:

Provide any records generated as a result of restrictions on, controls over and monitoring of employee use of Internet resources.

Interrogatory No.______:

Provide a list of any and all Internet-related data on the PCs used by [specific employees or classes of employees], including but not limited to saved Web pages, lists of Web sites, URL addresses, Web browser software and settings, bookmarks, favorites, history lists, caches, cookies.

XI. Data security measures

Interrogatory No.______:

List any and all user identification numbers and passwords necessary to access computers or programs addressed in interrogatories and requests. Your response to this interrogatory must be updated with responses to future sets of interrogatories and requests and updated responses to any set of interrogatories and requests.

Interrogatory No.______:

Please provide copies of your computer security policies and procedures and the name and contact information for the person responsible for security.

Interrogatory No.______:

Please provide information about the security settings for the [program]. For example, please provide the security settings for the Exchange Server, noting who has administrative rights.

XII. Network questions

Request No.______:

Produce any and all documents and things related to networks or groups of connected computers that allow people to share information and equipment, including but not limited to local area networks (LANs), wide area networks (WANs), metropolitan area networks (MANs), storage area networks (SANs), peer-to-peer networks, client-server networks, integrated services digital networks and VPNs.

Request No.______:

Produce any and all components related to networks, including but not limited to information exchange components (e.g., Ethernet, token-ring, ATM), network work file servers, traffic, hubs, network interface cards, cables, firewalls, user names, passwords and intranet.

N1. System overview

Interrogatory No.______:

Do you have any graphic representation of the components of your computer network, and the relationship of those components to each other, including but not limited to flow charts, videos or photos, and drawings? Include network topology documents and network schemas in your response.

Interrogatory No.______:

If so, where are the documents located? Include logical paths and physical locations for electronic representations.

Request No.______:

Produce copies of any and all graphic representations of your computer network, and the relationship of those components to each other, including any revisions, for the period of _____ to _____ inclusive. If the documents are electronic, produce them in their native form, as they existed at the time they were drafted, based on version or backup data.

XIII. Electronic mail (e-mail)

Request No.______:

Produce any and all information related to e-mail, including but not limited to current, backed-up and archived programs, accounts, unified messaging, server-based e-mail, Web-based e-mail, dial-up e-mail, user names and addresses, domain names and addresses, e-mail messages, attachments, manual and automated mailing lists and mailing list addresses.

appendix c: sample FRCP 30(b)(6) deposition questions

Questions to Ask at the FRCP 30(b)(6) Deposition of the Designated IT Person

The following questions may be useful to ask of the deponent in order to track down the sources of electronic data relevant to your case. The broad nontechnical questions are meant to be foundational to the technical ones.

Depending on your jurisdiction's limitation on interrogatories, most of the following questions can be used first in interrogatories, with follow-ups wherever necessary in depositions and requests for production.

These questions are crafted to assist you in getting focused, specific answers. Whether you ask these questions or others, of course, is up to you. You are far more familiar with the specific issues of your case, and our charge is not to provide you with legal but rather technical advice.

After you have had the deposition transcribed, we can help you assess the responses you were given for completeness and accuracy.

Personnel:

1. To whom at your company do you report?
2. Who works directly under you?
3. Does your company have an organizational chart? Who has custody of such a chart?
4. How many people at your company have a direct responsibility for computers and/or networks?
5. What are their job titles and duties?
6. What outsourced services, if any, do you use in the care and maintenance of your company's computer hardware, software or network(s)?
7. Who is the person at your company ultimately responsible for responding to discovery requests made of your attorney(s) in this lawsuit? If you are not that person, what role, if any, did you play in responding to discovery requests made of your attorney(s) in this lawsuit?

8. What have you done to prepare for your deposition here today?
9. What documents did you review prior to your deposition here today?

Systems information:

1. Do you use a computer or computers at work? If so, how many?
   a. Is the computer on your desk?
   b. Do you know the brand name?
   c. Does it run on Windows?
      If so, Windows 3.1, 95, 98, 2000, XP or NT?
      If not, is it a Unix-based system? Linux? Macintosh? Apple?
   d. Are you the only person with access to this computer?
      If not, who else uses it?
      If not, who else has access to it?
   e. Has this computer ever been used in the past by anyone else?
      If so, by whom?
   f. Who keeps the records of purchases of equipment, such as your computer?
   g. When and where was your computer bought?
   h. Does it have a hard drive? More than one? Do you know what its/their storage capacity is in gigabytes?
   i. Do you need to use a password to access the computer? If so, what is the password? How often have you changed that password?
   j. Does the computer have a floppy disk drive? A CD-ROM drive?
   k. Is your computer capable of saving files on a JAZ drive?
   l. Is your computer equipped with software known as Laplink?
   m. Is your computer equipped with software known as PC-Anywhere?
   n. Is it connected in such a way as to be able to access the Internet and send and receive e-mail?
   o. Do you ever back up anything from your computer? If so, how often, what are the times of the last two backups, and

on what kind of media (floppy disk, other disk drive, onto the server, Jaz drive, CD-ROM, streaming backup tape media, [Other? If so, what?])

p. Do you ever make "personal" floppy disk or disks for copies of files or other data from your computer (workstation)? If so, when? Why? How often? Where are those floppies? How many are at work? How many at home or elsewhere?

q. Do you use a laptop, notebook computer or personal data assistant such as a PalmPilot, Pocket PC or other portable computer? How are data moved back and forth between these devices and your computer (workstation)?

r. Do you use a computer at home? [If so, repeat relevant questions above as to type, kind, features, backups, who else has access to it, etc.]

s. Have you received or sent e-mail from your home computer?

t. Have you worked on your home computer to do work for your company?

u. Has your employer paid for all or part of your home computer?

Network Information:

1. Is the computer you use at work connected to a network? What kind of network is it (LAN, WAN, intranet)? If so, do you know what your access rights are? (Translation: level of security to access files, the top access right being "administrator" in most systems)

   a. What is your password to get onto the network?

   b. Who else knows your password?

   c. Do you know what a server is? [If not, have him/her assume it is the main computer to which all the work station (client) desktop computers are connected] Is your computer connected to a server?

   d. Is there more than one server in the network?

   e. Under what operating system does the server work (e.g., Windows NT, Novell, Unix, Linux)? What version?

f. Do you have a network administrator whose job it is to make sure the network keeps running properly or who can fix it when it does not? If not, whose job is that?

g. Is there an Information Services or Information Technology department in your company? Who are those people and who is the person in charge?

h. Is there somebody in charge of doing backups of data from the server? If so, who?

i. Do you know how often the network server is backed up? If not, are you aware that there is a backup system in place? If not, what is your understanding as to how your company's computer data will be preserved in the event of a disaster, such as a fire, flood, theft or vandalism?

j. Do you know what kinds of backup hardware and software are used? If so, what specific hardware and software, and version of that software? If not, who would most likely know the answers to this question?

E-mail and Software Usage:

1. Do you send and receive e-mail from your computer? If so, what e-mail program(s) do you use? (If he/she does not know, suggest Microsoft Exchange/Outlook, Lotus CC:Mail, Groupwise or Eudora as possibilities.)
2. Do you know how received e-mail is routed in your company? Does your company have document management software to log incoming and/or outgoing e-mail? If not, is there a policy on how incoming e-mail is logged and routed? How often do you delete your incoming e-mail from your work station? Outgoing e-mail? When is the last time you did that? And the time before that?
3. How often do you delete your incoming e-mail from the server? Outgoing e-mail? When is the last time you did that? And the time before that?
4. Do you know if your e-mail is backed up from time to time, either from your desktop computer or from the network server?
5. Have you ever printed out an e-mail message? If so, how often, when and for what purpose?

6. What word processing software do you use? Is this the same software used by everyone else in the company, to your knowledge? Was there at any time a different word processing program used by the company? If so, when was/were the conversion(s) made?
7. Repeat question 6 as to spreadsheet program (such as Excel); database management program (such as Microsoft Access); presentation software (such as PowerPoint); personal information management software (such as Outlook); contact management software (such as Act!).
8. What other software do you use in your work?
9. Do you do your own word processing? If not, who does?
10. Do you maintain your own calendar? If not, who does?
11. If you have a PalmPilot, Pocket PC or other handheld data device, do you keep your calendar on it? How far back does it go? What other data do you keep on such a device?

Preservation of evidence:

1. Once this lawsuit was brought, did anyone tell you to preserve all your electronic data and not erase any of it? Since [commencement of lawsuit], have you deleted any files from your computer?
2. Have you been asked by anyone since [commencement of lawsuit] to delete any files on any computer to which you have access, whether at work or at home or elsewhere?
3. Do you know what it is to "defragment" your computer? If so, have you ever run such a program to defragment it?
4. Do you know what it means to "wipe" a hard drive? If so, have you ever run software that will "wipe" a drive? Where? When? Why?
5. Does your company have a written e-mail policy concerning such matters as prohibited content, statement of who owns the e-mail, expectations of privacy, or any other policy relating to e-mail? If so, how long has that policy been in place? How many times has it been amended? Who is the author of the current policy? Do you have a copy of the policy?
6. Does your company have a written policy concerning the retention and destruction of paper records? If so, how long has that policy been

in place? How many times has it been amended? Who is the author of the current policy? Do you have a copy of the policy?

7. Does your company have a written policy concerning the retention and destruction of electronic records (such as files generated by computers, voice mail, e-mail)? If so, how long has that policy been in place? How many times has it been amended? Who is the author of the current policy? Do you have a copy of the policy?
8. Does your company have a disaster recovery plan? A Y2K plan?

Third-party sources of information:

1. Are there places outside the physical location of your company where you regularly send e-mail or other electronic files, such as a parent company, subsidiary, outside consultants, government entities, investors, other parties to this lawsuit [hereinafter = "outside entity"]? If so, to whom, when and why?
2. Who are the persons at [outside entity] who regularly send you e-mail? What are their functions and what purposes does each person's e-mail serve (i.e., in terms of his or her job function)?
3. Have you seen what kind of computers the [outside entity] uses?
4. Can you estimate how many computers there are there?
5. Are they networked?
6. Does the [outside entity] direct you or otherwise expect you to use compatible software?
7. [If foreign entity] Do you know how much of what they generate internally is in English and how much is in [foreign language]?
8. Do you know how received e-mail is routed in the [outside entity]? Does [outside entity] have document management software to log incoming and/or outgoing e-mail? Do you have such software? If not, is there a policy on how incoming e-mail is logged and routed? Does this policy come from [outside entity]?
9. Have you ever been told by anyone in [outside entity] that they did not receive e-mail you had sent them? To your knowledge, has lost e-mail been a problem expressed to you by [outside entity]?

10. Who else besides you receives e-mail from the [outside entity]?
11. Besides e-mail, have you ever exchanged documents in electronic form between your company and the [outside entity]? If so, have you ever done so by attaching documents to e-mail? Any other way?
12. Are faxes routinely sent back and forth between you and [outside entity]?
13. Is there a fax log kept of these incoming and outgoing faxes?
14. Do you have such a log kept on a computer? If so, how is that log kept and by what software?
15. In sending or receiving a fax between you and [outside entity], has there ever been a fax sent electronically, i.e., from a computer without use of paper? Is that the typical way faxes are sent? (If appropriate: If not, what if anything explains why sometimes a fax would be sent electronically and why at other times by paper?)
16. Have you received any e-mails from [outside entity] concerning this lawsuit?
17. Have you been told by anyone at [outside entity] to do anything with information in your possession or to which you have access, such as to preserve it, destroy it or anything else?
18. Have you received from or sent e-mails to anyone in [outside entity] in the last ___ years?
19. Have you deleted from your office or home computer(s) any of that e-mail? If so, what and when?
20. Do you have a secretary or assistant who prepares documents for you? Does he or she have a computer? [If so, you may want to repeat questions above about type of computer, whether networked, types of software used, configuration and features, whether he/she keeps duplicate or backup files from his/her computer.]
21. Does your company have voice mail? Are voice mails stored and retained? Do you receive voice mail from time to time from [outside entity]?
22. What is the most often-used medium of document exchange between [outside entity]and you: e-mail, fax, regular mail or something else?
23. How frequent are your communications with [outside entity] (i.e., daily, once a week, monthly)? Has there ever been a significant period

of time (i.e., more than ___ weeks) when there was no communication between you and [outside entity]? If so, was that, if you know, due to technical problems of any kind, such as computer system failures?

24. Do you have any communications with [outside entity] through use of collaborative software, such as "whiteboarding" over the Internet with something like Microsoft NetMeeting, or through use of a Web site where people can meet via the Internet and share ideas and thoughts?
25. Is there a site on the Internet where you and [outside entity] post and share information?
26. Is there a site not on the Internet (such as an intranet or extranet) where you and the [outside entity] post and share information?
27. Does [outside entity] have a Web site? If so, what is its Internet address?
28. In conjunction with this lawsuit, has anyone at [outside entity] asked you to look for records or otherwise supply information, either in paper or electronic form? If so, what documents, when and for what stated purpose?

Storage:

1. Where are backup tapes stored in your company (physical location)?
2. What security measures are in place to protect unauthorized access to those tapes?
3. Are backups kept on anything other than tapes?
4. What person or persons have custody of those tapes?
5. Who makes the backups? How long has that person been doing that? Who did the backups before that person?
6. What is the brand name and type of backup tape?
7. Are the backup tapes labeled? If so, what information is kept on the labels?
8. Have any backup tapes ever been destroyed, erased or altered, to your knowledge? If so, when, where and why?

Other lawsuits:

1. Has your company ever produced electronic data to another party in a lawsuit? If so, in what matter? What data was produced? In what format was the electronic data produced (e.g., on CD-ROMs, floppies, printouts?)
2. To what party or parties was that data produced? Was any of the electronic data used at trial? In support of any motions?
3. What was/were the caption of that case/those cases, and filed with what court(s)?

Retired hardware:

1. What is the usual life span of a computer at your company?
2. Is there a practice of computers used at one echelon of the company (for example, top management) migrating to other employees in the company as newer equipment is purchased?
3. What happens to computers and/or their hard drives when they are retired from service?
4. How many computers from _____ to ______ were retired from service, given away or sold? Who would have records that might answer that question? What would you call such records for the purpose of being able to identify them?

Encryption and legacy data:

1. Is encryption (encoding of data that prevents accessing it without a proper decoding or decryption software or hardware key) used in any of the following?
    a. e-mails
    b. data stored on servers
    c. data stored on backups
    d. data generated by any software application
    e. other

    If so, state the type of encryption used, the name(s) of any software used, the level of encryption, and how, for each instance, the process for decryption of the data works.

2. Do you maintain electronic data on backup tapes, archive tapes, hard drives, or otherwise, for which you no longer actively use the software to read the data on those storage media? [This question is meant to elicit information about “legacy data,” electronic data on such old software programs as WordStar, VisiCalc or WANG.] If so, do you still have the old software and user manuals? Who in your company would know how to recover that data?

appendix d: data collection forms

receipt of media form

Prepared at the request of counsel — privileged and confidential

Received from:
Name ______________________________ Title ______________________

Address ____________________ City ___________ State ____ Zip _____

Phone (w) (____) ___________ Ext ____ Phone (h) (___) ___________ Ext _____

Cell (____) _________ Pager (___) _________ PIN _____ E-mail ___________

Media Type	Serial Number	Notes

Released to:

Name ______________________________ Title ______________________

Signature ______________________________________ Date ____________

Phone (w) (____) ___________ Ext ____ Phone (h) (___) ___________ Ext _____

E-mail ____

desktop collection information form

Prepared at the request of counsel — privileged and confidential

The information below will assist in describing the chain of custody for the data which is collected.

Computer user: Name ___________________________ E-mail address _________________
Address ___________________________ City __________________ State _____ Zip ________
Phone (w) (____) _________________ Ext ______ Phone (h) (____) _________________ Ext ______
Cell (_____) ________________ Pager (_____) ________________ PIN ______ Title ___________
ISID _________________________

Brought to you? Y/N If yes, by whom? Name ____
Phone (w) (______) ________________ Ext ______ Phone (h) (_____) ________________ Ext ____
If not brought to you:
Building ___________________________ Floor __________ Office/Cube/Mailstop __________
Was this the only computer in this person's area? Y/N
Location of computer (e.g., on desk, under desk, in leather bag) _________________________

Machine/drive identifiers:
Type (laptop, desktop, notebook, server) _____________________ Manufacturer _____________
Model ______________ Serial Number _______________________ Asset/inventory tag ________

Number of peripherals attached to the computer with the computer of interest (write 0 if none)
Also, removable media in area:

Hard drives _______________________
CD-ROM Read/write _______________
3.5 floppy _______________________
Zip/Jaz drive _____________________
Docking station ___________________
Floppies __________________________
CD-ROMs _________________________
CD-ROM Read-only _______________________
Fax/modem _____________________________
Printer ports ___________________________
USB ports _____________________________
Network connection _____________________
Tapes _________________________________
Zip, Jaz _______________________________
Other not listed ____
Notes: ____

Bios date/time _____________________ **Actual date/time** __________________________

Name of ghost image ________________ Copied to network drive ____________________

Type of data	Name	Date collected	Collected by:
Home directory			
User computer			
E-mail			
Share 1			
Share 2			

Continue on back of sheet

Gatherer's information and signature:
Signature _____________________________________ Name ______________________________
Gathered at the request of ____
Case identifier ____

server information form (for Microsoft® Exchange)

Prepared at the request of counsel — privileged and confidential

Location:________________________ Server:___________________________

On-site contact:

Name _______________________________ Title ________________________

Address_____________________ City ____________ State____ Zip ________

Phone (w) (___) ________________ Ext ___ Phone (h) (___) _______________

Cell (___) ________________ Pager (___) ______________ PIN____________

E-mail (w) _________________________ E-mail (h) ______________________

Physical Location:
Secure location? Y/N If yes, type of access (e.g., keycard, code) _________________
Phone (in room) (___) ______________________________ Ext _____________
Building ____________________________ Floor _________ Room _________
Location of computer (e.g., rack #, shelf) _____________________________

Hardware:
Machine name ______________________ Manufacturer _________________
Domain name ____
Model _________________ Serial # ________________Asset tag # _________
Tape drive manufacturer ___________ Model # ______Serial # ___________
Tape media size ____

Time:
Bios date/time ____________________ Actual date/time _________________
Time zone ________________________ Daylight savings selected? _________

Software:
Operating system version and service pack ____________________________
Exchange version and service pack ___________________________________
Site name __________________ Organization name ____________________
Exmerge version ______________
Backup software/version _________________ Open file agent? ___________

Gather's information and signature:
Name ______________________________ Title __________________________
Signature _________________________ Company ______________________
Phone (w) (___) ______________ Ext _____ Phone (h) (___) ______________
E-mail (w) _________________________ E-mail (h) ______________________

key personnel list

Prepared at the request of counsel — privileged and confidential

Location :____

Role	Name/Department	Phone/E-mail
Top Manager		
Top IT Manager		
E-mail		
Network		
Desktop		
Security		
Help Desk		
Telecom		
Human Resources		

legal notice

The content provided in *The Guide to Electronic Discovery* ("guide") is intended to provide you with an overview of the activities and issues related to the process of electronic discovery. While our goal is to provide you with useful practical and technical information, this guide should not be considered, and is not intended as, legal advice. You use the Guide at your own risk.

FIOS and the FIOS logo are registered trademarks, and PREVAIL is a trademark, of Fios, Inc. All other trademarks and copyrights are the property of their respective owners.

Please forward all permission requests to:

Fios, Inc.
Attn: Copyright Inquiries
921 SW Washington Street
Suite 850
Portland, OR 97205 USA

info@fiosinc.com

Total Cost of Compliance (TCC)?

They ruined everyone else's fun

By Fios, Inc.

Conversations about the missteps of highly compensated fast-moving corporate executives have become cliché. Late night talk show hosts and "Gen Y" constituents have made references to corporate swindling part of their parlance. As we remain astonished by the audacity of these miscreants, we are hit with a harsher reality: The dastardly deeds committed by these personalities have resulted in legislation and subsequent regulation to make our corporate lives—a little less fun.

We are, generally, a responsive society. In order to alleviate anxiety caused by upheaval, we respond rapidly and deliberately. To ease investors' concern over fraudulent accounting practices and general mismanagement, our governmental representatives have introduced compliance oriented measures for many of our nation's companies. These measures have become looming risk management clouds for your CFOs, CIOs and CEOs. They fear the punitive measures associated with non-compliance. Additionally, undue burden has been placed on your operating managers to institute processes, technology and employee training to ensure compliance. There is no better way to put the "un" in "fun."

The topic of this paper is **not** "How to Develop Corporate Strategies for Sarbanes-Oxley." Strategies to best apply today's knowledge- or records-management solutions toward a "proactive" corporate regulatory compliance initiative will be left to other experts.

Instead, this paper will focus on a concept called Total Cost of Compliance (TCC). We will explore the operational impact of compliance measures on your organization. The impact of proactive as well as reactive compliance measures and the cost of non-compliance in both scenarios will also be discussed.

Proactive Compliance = Predictable Compliance

Much noise has been made about the impact of compliance mandates such as Sarbanes-Oxley, even more about how to interpret and integrate them into business practice. In spite of their vigor, you have undoubtedly questioned the clarity of these mandates' "terms and conditions." However, Sarbanes-Oxley does specifically outline the expectations government regulators have for corporate behavior. With this, your compliance officer can now put a plan in place that enables adherence to the standard. The processes and costs associated with implementing and maintaining such a (proactive) compliance plan are relatively predictable.

Recent reinterpretations of the SEC Act of 1934 targeting data retention policies clearly articulate the processes and procedures that your company must employ to align with these regulations. Again, the cost associated with these types of initiatives is relatively transparent. The penalties associated with non-compliance are also well documented. With this in place, you can now develop an ROI and begin to implement a corporate wide, "proactive" compliance solution to address Sarbanes-Oxley, or any other regulatory mandate.

The cost impact of proactive compliance (or non-compliance) has become relatively predictable.

Reactive Compliance

The reality is that compliance goes beyond these "relatively well-defined and predictable" mandates. This is not meant to trivialize the effort associated with these corporate initiatives. They can be wicked complicated and in most cases require substantial business re-engineering. The requirements are often foreign to those tasked with implementation. The overhead can be burdensome, and the costs significant. Even if you follow the letter of the law, your company's position as it relates to compliance is in the "eye of the beholder", since anyone can call it into question. However, ultimately, compliance with these "known" mandates is a (relatively) predictable process with predictable costs.

A large portion of the costs associated with compliance (or non-compliance for some) has to do with components of your business that are not predictable or well defined. These components have been around long before Sarbanes-Oxley, and will remain even after numerous reinterpretations have been authored.

If your business operates in a highly regulated sector, litigation or government investigation is a constant presence. These events, or rather your response to them, quickly become the largest cost component associated with compliance. These events create a "reactionary" environment that disrupt your business and cause your organization to focus on resolving issues otherwise unrelated to normal business operations. This is the world of "reactive compliance."

Reactive Compliance = Unpredictable Compliance

In the midst of a quarterly strategic planning session your general counsel enters the room. Seems an intellectual property infringement claim has been brought against the company by a competitor. As demonstration of your competitors' intent to pursue the claim, computer forensic experts will arrive in 48 hours with a court order to begin examination of many computers, including yours. All electronic correspondence related to the technology in question, from all related employees, will be collected, reviewed and delivered to the lawyers filing the suit. Moreover, a preservation order has also been delivered that requires your IT department to immediately cease the OS upgrade critical to your company's next product release—so that all systems will remain "as is."

> ***"A large portion of costs associated with compliance has to do with components that are not as predictable as Sarbanes-Oxley."***

$$TCC = \sum (pTechnology + pPeople) + (rTechnology + rPeople + rOpportunityCost)$$

$$TCnC = \sum (pnCCivilPenalties + pnCCriminalPenalties) + (rnCCivilPenalties + rnCLitigationExposure)$$

Figure 1.

This scenario, although hypothetical, could be an actual account from many Fortune 1000 companies. In this scenario, the introduction of the lawsuit was not predictable (or at least the timing was uncertain) and the scope of activities involved with complying with "requests" was even less predictable. As a result, the process burden and therefore, cost burden on your organization is extremely unpredictable.

Total Cost of Compliance (TCC)

So, what is the point of the examples above? Aside from demonstrating the difference between compliance activities that are predictable and activities that are unpredictable, it is important to understand the entire organizational impact of compliance. Since profitability is "top of mind" for corporations, the impact of compliance is best measured by its impact on organizational costs—or the expense line of the income statement. The formulas in Figure 1 above outline the elements contained in a Total Cost of Compliance (TCC) and a Total Cost of Non-Compliance (TCnC) calculation:

Let's break down each formula into its "colloquial" components. The TCC formula says that the Total Cost of compliance is the sum of:

- the proactive compliance (p) programs' technology (software/hardware/IT enabled services) and people (employees and contractors) cost; plus
- the reactive compliance (r) technology (software/hardware/IT enabled services), people (employees and contractors) cost; and
- the opportunity cost associated with re-focusing mission-critical staff on short turnaround time, unpredictable initiatives at the expense of already planned initiatives.

The TCnC formula is equally straightforward. This is the sum of:

- the civil and/or criminal penalties that can be imposed if a company does not comply with proactive (p) mandates; plus
- the reactive (r) civil penalties associated with not complying with a discovery request[1] and the larger cost/risk associated with potentially compromising your company's position to defend itself in a lawsuit.

Predictability sometimes leads to disobedience

In the case of proactive compliance, some organizations choose to disobey the compliance mandates. Why? It is not because they are habitually disobedient. Rather, the costs associated with non-compliance are predictable. There are organizations that weigh the risks of being non-compliant with the (known) costs of compliance and choose to incur the costs of non-compliance. Also, there is a sense that the compliance-happy environment will disappear once the equity markets and investor confidence rises to higher levels. Other organizations are resource constrained, unaware, or believe that the worst will never happen to them. Why not just wait it out and incur known costs without organizational upheaval?

A reactive environment is usually created by a lawsuit or claim against your company, or as a result of an investigation from a regulatory body, such as the federal government. In cases like this, the cost of non-compliance could be a civil penalty or worse, the loss of a lawsuit which could potentially be a "bet the company" scenario. Increasingly, there are impacts to stock prices when correct information is not forthcoming, causing a publicity crisis. Non-compliance is usually not an option in a reactive environment. As a result, the operational costs associated with reactive compliance are unpredictable, but inevitable.

Inevitable costs—Data Collection

In the event of a discovery request, your IT organization is immediately impacted. The discovery request will contain demands for information pertaining to timeframes, people, and topics. Your general counsel, with the assistance from your external law firm, will launch portions of this request into the various organizations affected. In the case of a discovery request associated with electronic information, your IT organization is tasked with locating and aggregating the relevant information. This phase is what is referred to as *data collection*.

For those who are not well versed in the dynamics of a large IT infrastructure, this is no small feat. Even if you have deployed a knowledge management and data retention system, you will learn that only a subset of the relevant data resides there. Much more resides "live" on the network and on individual computers, while still more is stored on backup tapes. The process of aggregating electronic information in response to a discovery request resembles nothing that your IT team deploys on a typical day. The process requires a scalpel type technique, as data preservation (don't throw the stuff away) and non-spoliation (don't tamper with the stuff) rules mandate specific procedures to be followed. Lack of adherence to these processes can translate to immediate penalties, loss of good faith with the decision maker in the matter, or worse.

For this reason, you may choose to outsource data collection. Alternatively, you may train and deploy your own IT department. In the former scenario, the costs of external resources are significant. In the latter, the opportunity cost of re-focusing your in-house team is significant.

TCC and TCnC are about perspective

Total cost of compliance and total cost of non-compliance are concepts that are meant to provide perspective. It is not the intent of this paper to suggest that you embark on a quantitative exercise to calculate your precise TCC and TCnC and make decisions to comply (or not) on the basis of the results.

Rather, the purpose of this paper is to highlight the fact that organizational costs incurred in a reactive compliance situation are unpredictable yet inevitable. In addition, organizations often ignore these costs because the "cost of a lawsuit" is viewed primarily as the legal expenses related to the defense of the suit. Furthermore, courts have not made a habit of asking the requesting party to fund the costs of producing active data, because, by and large, it is assumed that the data requested is easily accessible. Perhaps, but access does not translate to ease of extraction. As a result, for the foreseeable future, you will incur significant costs to comply with a reactive compliance request. The key is to know what your costs are so that you can proactively manage them.

[1] a request for information from the opposition to enable "discovery" of "what happened"

Fios provides IT enabled services to simplify electronic discovery and reduce the operational cost of reactive compliance.