



A **TERIS** White Paper

1001 Fourth Ave, Suite 2110  
Seattle, WA 98154  
206-521-8717  
[www.DiscoverTERIS.com](http://www.DiscoverTERIS.com)

## How to Avoid Sanctions when Sailing the Sea of Data

*By Mark Prewitt, Director of Information Technologies*

February 9, 2009

■ [Seattle, San Francisco, Silicon Valley, San Diego, Phoenix, Austin](#) ■ (888) 99-TERIS ■

## Contents

---

Introduction	3
Problem Statement	4
Evolution of the Industry	5
Solutions Developed	6
Practical Tools Help	7
Efficient Searching	8
Litigation Support	8
Don't Drown!	9
Conclusion	10
Bibliography	11
Appendix B	12
Graphics & Charts	12

## Introduction

In today's society, the storage of documents is increasingly in electronic form. To search these documents requires some knowledge of technology and the best application of this technology to find the important facts in the tsunami of electronic information. In this paper, we will discuss the problems, pitfalls and potential solutions for managing such a deluge of data.

Searching through this ocean of data for cases is usually referred to by one of several common terms, Electronic Discovery, e-discovery, Electronic Data Discovery (EDD), or Electronically Stored Information (ESI). Data is collected from sources known in the industry as custodians, via forensically sound methods that allow attorneys for both sides (Plaintiff and Defense) to review the information for relevance to their case. As an example of how tedious this can be, a case may contain anywhere from a few hundred megabytes of email files to a case which may entail server hard disk images, resulting in terabytes of data. Regardless, the mere presence of this data requires that an attorney know something about technology or hire someone who does. This technology knowledge requirement has been slowly creating new problems for the legal industry as attorneys realize that mistakes in managing this information can cost them through sanctions.

The focus of this paper is on the dangers that face attorneys, as they must evaluate the electronic data in their cases. My own personal email contains more than 50,000 emails. Imagine the magnitude of email and other electronically stored documents in a company with ten or more employees! The task is frightening to manually review each of these emails, just to find the few that are relevant to a case – it's akin to the proverbial needle in a haystack. This need has spawned an entire litigation support industry as well as created sanctions due to improper handling of this torrent of data. Because of this overwhelming flood of data, the litigation support industry has developed tools to help shrink the data to a manageable amount.

One way that tools provide assistance is to filter documents by using keywords, dates, or concepts to shrink the pool of data into something more useful. It's the use of these filtering tools that has recently become more complicated due to improper application and sanctions. An example of how mishandling data can affect a case, we cite the [\*Qualcomm Inc. v. Broadcom Corp.\*](#)<sup>1</sup> case in which the attorneys for Qualcomm were blasted by the judge for 'aggravated litigation abuse.' The Judge's opinion, on top of previous opinions on electronic discovery in the last few years, is defining the course that attorneys must navigate in this swelling sea of data.

This growth in electronic information has invaded the legal market and caused long established, well paid attorneys, to choose whether to go back to school and learn about technology, or hire people to do the work for them, so they can concentrate on the law.

---

<sup>1</sup> [\*Qualcomm Inc. v. Broadcom Corp.\*](#), No. 05-CV-1958-B(BLM) Doc. 593 (S.D. Cal. Aug. 6, 2007)

## Problem Statement

In the world of litigation, one of the growing fears in the industry is the amount of electronic documents they will have to sift through for any case. It's estimated that in 2006 the 'digital world' encompassed 281 exabyte's of stored information, and it is projected that by the year 2011<sup>2</sup>, there will be ten (10) times<sup>3</sup> this amount of electronically stored documents in the world. These numbers include digital images and music, so not all of it will be relevant all the time, but this explosive growth means that electronic discovery will be a much larger part of any future case that attorneys must handle. The explosion of all this digital data in the world outside the courts has created hurdles for the legal system. Data must be collected in a fashion that the legal world is not familiar with - electronically. It must be kept pristine, to avoid accusations of data corruption or change. Data files on a computer are broken down into multiple parts, for brevity, I will only point out the two key parts; Metadata and Data. The Metadata is information about the data, and the data is the key information that must be looked at. Since the metadata is information about what we are looking at, it's imperative that it be reviewed as well. This, in effect, doubles the work for every relevant or irrelevant piece of data.

This long arduous process is daunting and can be expensive. The Sedona Conference has weighed in on the subject as recently as fall 2007 when they released an article entitled, "The Sedona Conference® Best Practices Commentary on the Use of Search & Information Retrieval Methods in E-Discovery."<sup>4</sup> This relevant article explains in more detail than we can in this paper the true dilemma that faces attorneys. The following is a simple comparison to illustrate why the digital age has confounded many attorneys. In 1990, the typical legal case would involve many paper documents, usually filling some amount of banker's boxes (10-50 for a small case). One gigabyte of electronic information can generate about 70,000 text pages, or 35 banker's boxes of documents (at 2,000 pages per box). A 100 gigabyte storage device (*e.g.*, a PC hard drive), theoretically, could hold as much as the equivalent of 3,500 banker's boxes of documents. By contrast, in 1990, a typical personal computer held just 200 megabytes of data - 1/500 the capacity of a typical hard drive today. Even if only 10% of a computer's available capacity today contains useful or "useable" information (as distinguished from application programs, operating systems, utilities, etc.), attorneys would need to consider and potentially review 700,000 pages per each device, and if a company has 20-30 employees (a small firm), the number of potential documents is staggering. The time and billable hours would make most cases too expensive and burdensome to even consider taking on without complete assurance of victory. The Sedona conference mentioned two additional problems:

- 1) Most electronically stored information today is in human language, which challenges computer search tools. These challenges lie in the language and our adaptation to short hand and slang for everyday items we become familiar with in a company or industry.
- 2) The efficiency of manual review versus alternatives is still in debate in judicial circles. Moreover, simple keyword searching, while valuable, is known to be inaccurate.

---

<sup>2</sup> IDC 2006 and subsequent 2008 update

<sup>3</sup> 1 Exabyte is equal to 1024 Petabytes which is 1,024 terabytes which is 1024 gigabytes

<sup>4</sup> [http://www.thosedonaconference.org/publications\\_html?grp=wgs110](http://www.thosedonaconference.org/publications_html?grp=wgs110)

These problems have caused the litigation support industry to develop new tools to assist an attorney in culling this data down from its tsunami like wave front. Litigation support firms use many tools to manage and shape the data. These tools can provide ways to eliminate obviously useless information such as the operating system (OS) files on a computer hard drive, the applications, spam, preference files and more that would typically not include relevant data. After this initial reduction of data, known as ‘culling’, an attorney finds that they are looking at a database of documents that may be relevant to a case. For an attorney to review 500,000 to 1,000,000 documents the task is overwhelming.

We see then, that the problem is the growing data pool that we, as a world, are generating with our computers. This exponential growth further spawns complimentary problems such as searching, cataloging, and backing up the data. There are numerous examples of how the technology is driving the cases, some cited in [Information Inflation: Can The Legal System Adapt](#)<sup>5</sup> or [Intellectual Foundation of Electronic Discovery](#)<sup>6</sup> which show us in cases like [Phoenix Four, Inc. v. Strategic Resources Corporation](#)<sup>7</sup> and [Residential Funding Corp. v. DeGeorge Fin. Corp.](#)<sup>8</sup>, where the attorneys involved were sanctioned because of improper or poor electronic discovery practices. Basic technology knowledge was lacking and it resulted in poor practice in dealing with the electronic information. There’s a growing sense that attorneys need to know more about technology, but this might come at the expense of their knowledge of the law. There have even been many attempts by clients to falsify data, as in the case [US v. Johnson](#)<sup>9</sup>, which forced an ethics decision by the attorney to withdraw from the case. This probably saved the attorney from sanctions later.

## Evolution of the Industry

In previous years, most attorneys were limited to inefficient searching and review of documents in a fashion that was very slow and tedious. This involved placing a database on computers and hiring banks of review attorneys or paralegals to pour-over the documents one-by-one and label (tag) the documents as privileged, relevant, responsive, or some other label. Subsequent this first review, another tier of reviewers would then look at these documents and filter the large batch of ESI into a small manageable wave of data to actually work with for the case. This was obviously expensive, time consuming and impractical for very large cases, and too costly in very small cases due to the cost of the paralegal’s time.

Many companies leapt into the market to provide solutions to the data wave problem. Many solutions were created and many failed. Several solutions had some common tools built in that have become accepted as standards. Under the new Federal Rules of Civil Procedure, it was acceptable to simply only make a ‘reasonable’ effort to search and identify relevant documents. Judges, attorneys, and their clients did not want to pay for the enormous labor involved in detailed searches. However, as new products are introduced and technology improves, the industry is seeing that technology can solve many of these hurdles and is solving them with greater efficiency and ability. The danger continues as many attorneys are failing to keep abreast

---

<sup>5</sup> [Information Inflation: Can The Legal System Adapt?](#) 13 Rich J.L. & Tech 10 (2007)

<sup>6</sup> [Intellectual Foundation of Electronic Discovery](#)

<sup>7</sup> [Phoenix Four, Inc. v. Strategic Resources Corporation](#), 2006 WL 1409413 (S.D.N.Y. May 23, 2006)

<sup>8</sup> [Residential Funding Corp. v. DeGeorge Fin. Corp.](#), 306 F.3d 99 (2d Cir. 2002)

<sup>9</sup> [US v. Johnson](#) (2008 WL 2060597 (E.D. Va May 15, 2008))

of the technology improvements, to their own peril. The litigation partner's job appears to take the heat off the attorney to learn technology, thus exposing them to the potential for sanctions without realizing it. In the case [U.S. v. Michael John O'Keefe](#)<sup>10</sup>, U.S. Magistrate John Facciola ruled that the adequacy of a keyword search:

*"...is a complicated question involving the interplay, at least, of the sciences of computer technology, statistics, and linguistics. . . Given this complexity, . . . [t]his topic is clearly beyond the ken of a layman and requires that any such conclusion be based on evidence that, for example, meets the criteria of Rule 702 of the Federal Rules of Evidence."*

Rule 702 is what authorizes the use of expert testimony in Federal cases.

An even more direct case, that has shaped the future direction of searching data, is the [Victor Stanley, Inc. v. Creative Pipe, Inc.](#)<sup>11</sup> case. The judge in this case, Judge Grimm, provides a 43 page opinion that has become a guide on how to properly detail to the court the process used for searching electronic documents. He further allowed the threat of sanctions for the attorney if they fail to follow his guide. Another relevant case that has shaped the industry is [Equity Analytics, LLC v. Lundin](#)<sup>12</sup> In this case, Judge Facciola lays down the charge that attorneys don't understand the technology and need to educate themselves.

It's with these cases and the realization that the future of electronic discovery has become a minefield for litigants that we begin to explore solutions. Attorneys must prepare or take the risk of sanctions. Potential clients must adapt or drown in the flood of data in their corporations. This same flood will become their worst nightmare in a situation involving litigation. Traditionally there has been three main duties for Attorneys; Zealous Advocacy, Duties to Court, and law competence. I would advocate that there is now the need for technology competence. Most seem to be aware that the technologically gifted tend not to become attorneys. In researching this paper I came across several laments about the lack of technologically gifted attorneys (outside the area of Intellectual Property or Patent law). This seems to be a key competency that is lacking in the industry as it moves forward into the sea of data that confronts it. Obviously, partnering with firms that provide litigation support is a key method to overcome this lack of competence, but it still leaves the attorney at the mercy of their choice of partners.

## Solutions Developed

Using existing technologies, the litigation support industry has begun to educate itself on the best way to manage the growing flood of data. These firms are doing this out of self preservation, as they have to handle storing and converting this data. Gone are the days where a few dozen copy machines and handling the resulting deforestation were the main job of these firms. Now they are working with hard drives from litigant computers and servers. These hard drives are full of potentially relevant data that must be culled and prepared for the attorneys. At our company, we use a strategy of multiple best of industry tools. The purpose behind this philosophy is to keep all options open, and to have the latest and best tools available to clients. Some firms have developed their own proprietary tools. This has many dangers as the tools tend to be developed towards certain 'problems' and as the problems that must be solved change, and

<sup>10</sup> U.S. v. Michael John O'Keefe, 2008 WL 449729 (D.D.C.)

<sup>11</sup> Victor Stanley, Inc. v. Creative Pipe, Inc., 2008 WL 2221841 (D. Md., May 29, 2008)

<sup>12</sup> Equity Analytics, LLC v. Lundin, 248 F.R.D. 331, 331 (D.D.C. March 7, 2008).

the amount of data continues to grow, the required changes take time and energy and resources to fix. By keeping our resources tied up in the expertise of using the tools and researching all possible tools, we feel that we are uniquely positioned to provide the best possible solutions to any potential problem that clients may need to solve.

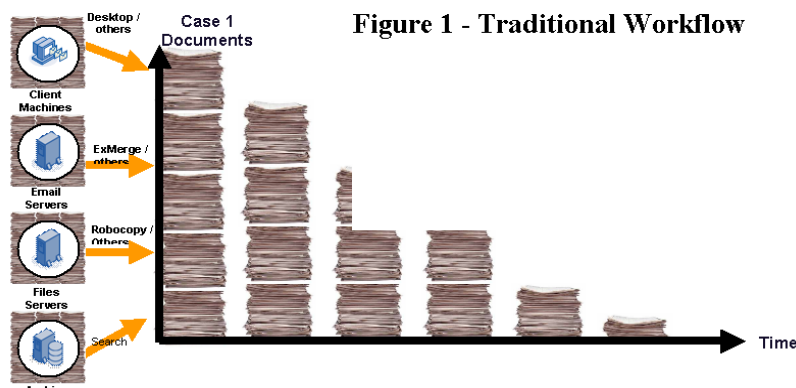
A recent acquisition by our company (ALC) is the new Clearwell Systems computer appliance that allows for culling of email and loose files. This ‘box’ is designed to be used early in the case management process. In culling early on in the process, it frees up the attorneys time so s/he can focus on the relevant documents in the case and not the traditional daily chit chat that most emails are full of. It further allows for the latest in keyword, content and discussion searching of these documents to further remove decrease the document corpus. This is a key change that many are just beginning to realize to be the fastest way to cull the data. To search for what is relevant leaves you searching large amounts of data, with long search times and much frustration. To search for irrelevant data turns up larger pieces of the pie to cut off until you are down to the final small piece that contains what you really need.

Traditional review is the type of process that lends itself to many hours of work, missed documents that should have been produced and sanctions. Non-attorneys are doing the bulk of the grunt work and the possibility for sanctions and production of privileged data or non-responsive data is high. As the industry develops tools to assist attorneys, their *appropriate* adoption will be key in making up for the lack of technology competency, as well as finding partners who can work with the attorney as a consultant to ensure that the process is accurate and meaningful for all parties involved.

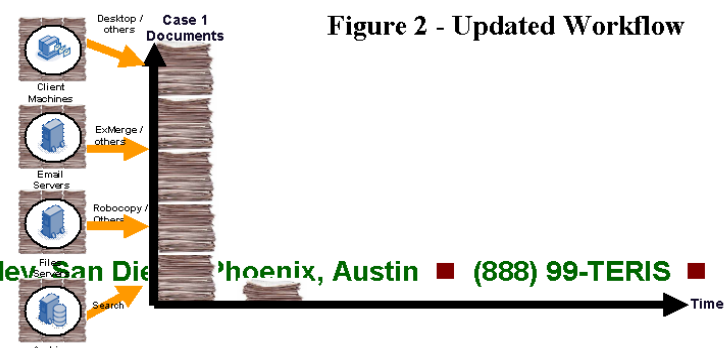
### Practical Tools Help

In utilizing the latest technological tools, we typically take what is an impenetrable tsunami of data and siphon it down to a manageable trickle that can be dealt with in any of the current or old style methods that attorneys are used to. With fewer documents it also lowers the cost of the litigation and risk associated with mistakes. The traditional process looked something like Chart 1 in Appendix B.

Or it could look like Figure 1., where the case documents are slowly reviewed and culled over time resulting in the final pile of produced documents.



With the new technologies available today, we move from the traditional approach, where culling the data takes large amounts of time, to something that is much more efficient as represented in Figure 2.



## Efficient Searching

With the smaller set of documents, you can quickly search for relevant terms and phrases and keywords. These will result in more accurate information coming back and less accidental returns that must be tossed out. This saves time and money and lessens risk for all parties involved. The review products we utilize allow for nearby, fuzzy and relevancy. These review tools then allow the attorney to be more efficient in their time and fine tune their searching by having only a few documents to look at instead of the 100,000's that could easily be returned without culling. And in looking up relevant data on patent litigation, who wants to find out that the '415 patent' also returns all the 415 area code documents? This is an area where the attorney must also partner with the litigation support team and not gloss over it. The failure in the Creative Pipe case was all about poor searching. If the attorneys had spent some more time examining their search terms and consulted with the case management team at their litigation company, they might have not turned over privileged documents.

It will be in this area that the technology gap of most attorneys will come into glaring focus. The simple misunderstanding of how a 'simple search' can return so much irrelevant data to still have to cull. The 'technology incompetence'<sup>13</sup> that this shows ends up creating a legal incompetence or failure to fulfill the duties to client & court. This is exemplified in the Qualcomm case where the attorneys are vilified by the Judge for their incompetence and lack of ethical behavior because of this technology gap.

## Litigation Support

The litigation support company is a law firm's friend. These firms have focused their business on Information Sciences and technology. By working with them, the attorney creates a team approach to the problem that will overcome the information hurdle far easier than working alone or without an expert in the field. It has been suggested by Ralph Losey (who created the triangle picture on Chart 2 of Appendix B)<sup>14</sup> that the field of law is evolving and that new attorney's who fail to understand technology will fail at law. It is his concept that for any case, an electronic discovery team must be implemented so that you have all the appropriate knowledge for successful electronics discovery, Law, Technology & Information Science. This could occur either in a corporate setting, or for litigation. The key being that with the tidal wave of data, only one branch of information will not stand the tide.

The idea is that the three disciplines that must be forged together to survive the tide. Law, Technology, and Information Sciences. We show this in a Pyramid diagram (Chart 2 , Appendix B). The Pyramid represents the three disciplines setup for successful electronic discovery. Without any one, the whole is not functional. We have seen that attorneys in cases (Qualcomm, Phoenix, US v Johnson) can be held liable for this lack of competence with technology as it relates to discovery. It would tend to make some sense to implement a strategy to avoid this pitfall.

---

<sup>13</sup> Ethics and Professionalism in the Digital Age: Foreword, *60 Mercer Law Review* \_\_ (forthcoming 2008) (with David Hricik)

<sup>14</sup> See footnote 4 on Bibliography page

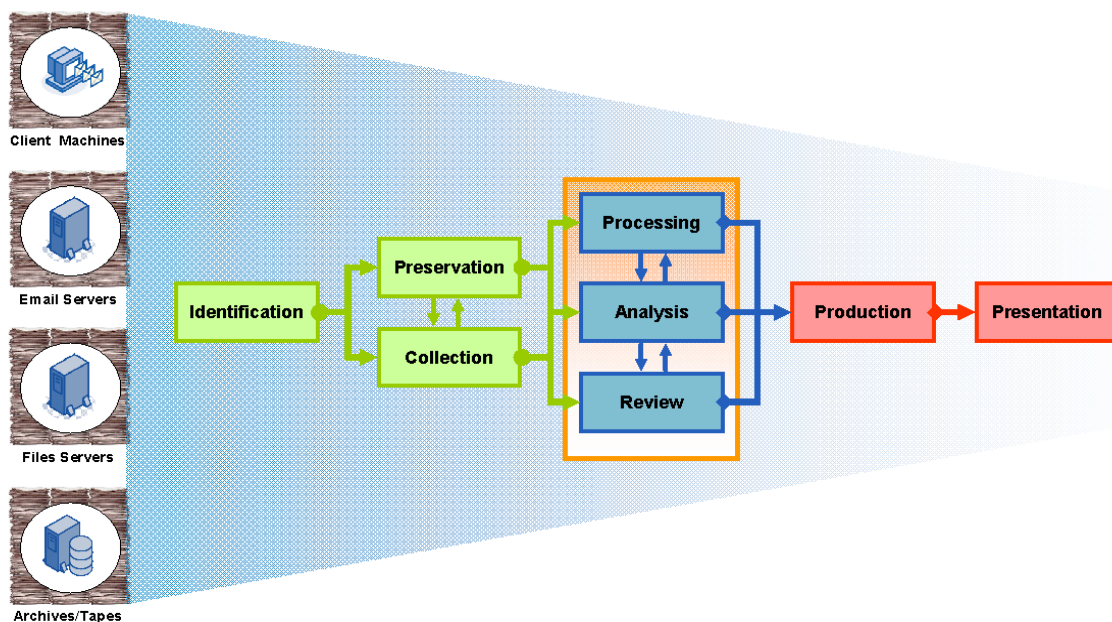


Our thought is that having an expert, as is customary for discovery, makes more sense than having attorneys return to school to learn skills as a technology guru. I think the better course is to allow qualified partners to supply this knowledge and gain whatever can be gained through the course of performing your role as an attorney. Self education would be a highly desirable tact moving forward, but the burden should be on having enough information to properly understand the technology, without the depth to do the daily tasks.

### Don't Drown!

This is the key aspect for this paper; without proper implementation, we are all awash in the flood of data that has risen to engulf us. Perhaps by now, you have probably caught on that we are painting a picture of oceanic proportions. It's our opinion that with the court rulings, attorneys must either learn technology, or partner with experts in technology and information science. They will either have a lifesaving partner, or begin to drown in the data that will be coming their way. It's as inevitable as the explosion of digital media on consumers. As people use digital technology at home, it moves to corporations. As corporations embrace technology, so we in the litigation market must embrace it.

Our proposal is simple, as an attorney begins to assess their case and workflow, represented here:



They involve an electronic discovery expert. As they begin the preservation & collection phase of the process, already have one of our experts working with them. This person's responsibility will be to guide and assist the litigants in best practices for the data collection, and setup for processing. We would help with early case planning and assist with guidelines for keywords and helping test these on samples of the data to ensure that terms are actually useful as keywords or concepts. The new Rule 26(a)(1)(E) calls for a conference to discuss planning for discovery. This is a good time to be preparing by discussing, with our case managers, the kinds of tools that are available for the case, what methods might help, and if the other party is claiming unreasonable burden under rule 26(B), we can work with both parties to ensure an equitable fair confirmation of this. ALC strives to be an impartial worker for the cases we work on. Our job is

to process the data in the most efficient and effective manner. This saves our clients money, time and energy they can use on other things.

## **Conclusion**

The pitfalls are clear, failure to understand technology can leave you open to sanctions due to fraud by your clients, failure to produce, production of privilege documents and just appearing incompetent before the court. These are a clear and present danger and, depending on the judge, could be viewed anywhere from simple incompetence to an ethical violation inviting sanctions. And avoiding these sanctions requires a response. Just moving forward with the status quo implies that you are either pretending that technology will go away, or you turn your back on it and don't see the tidal wave of data coming at you.

In this paper, we have looked at the problems of the data flood, the growing annoyance of judges with attorneys who abuse the courts through poor discovery practices, the growth of the litigation support industry into technology, and looked briefly at the evolving process that will soon be common place to form discovery teams to ensure the technology is applied correctly. The conclusion we would like you to derive from all this, is that short of going back to school for a computer science degree, you need to take steps to find a competent and savvy technology partner. This partner will provide you with the expertise the judges and your clients want you to have. This is the role that firms such as ALC must take moving forward; we will be the technology/information service provider/Consultant/Expert to assist the attorney in their cases. With the assistance of competent and knowledgeable experts, attorneys are far less likely to have sanctions imposed on them due to their own lack of technological savvy.

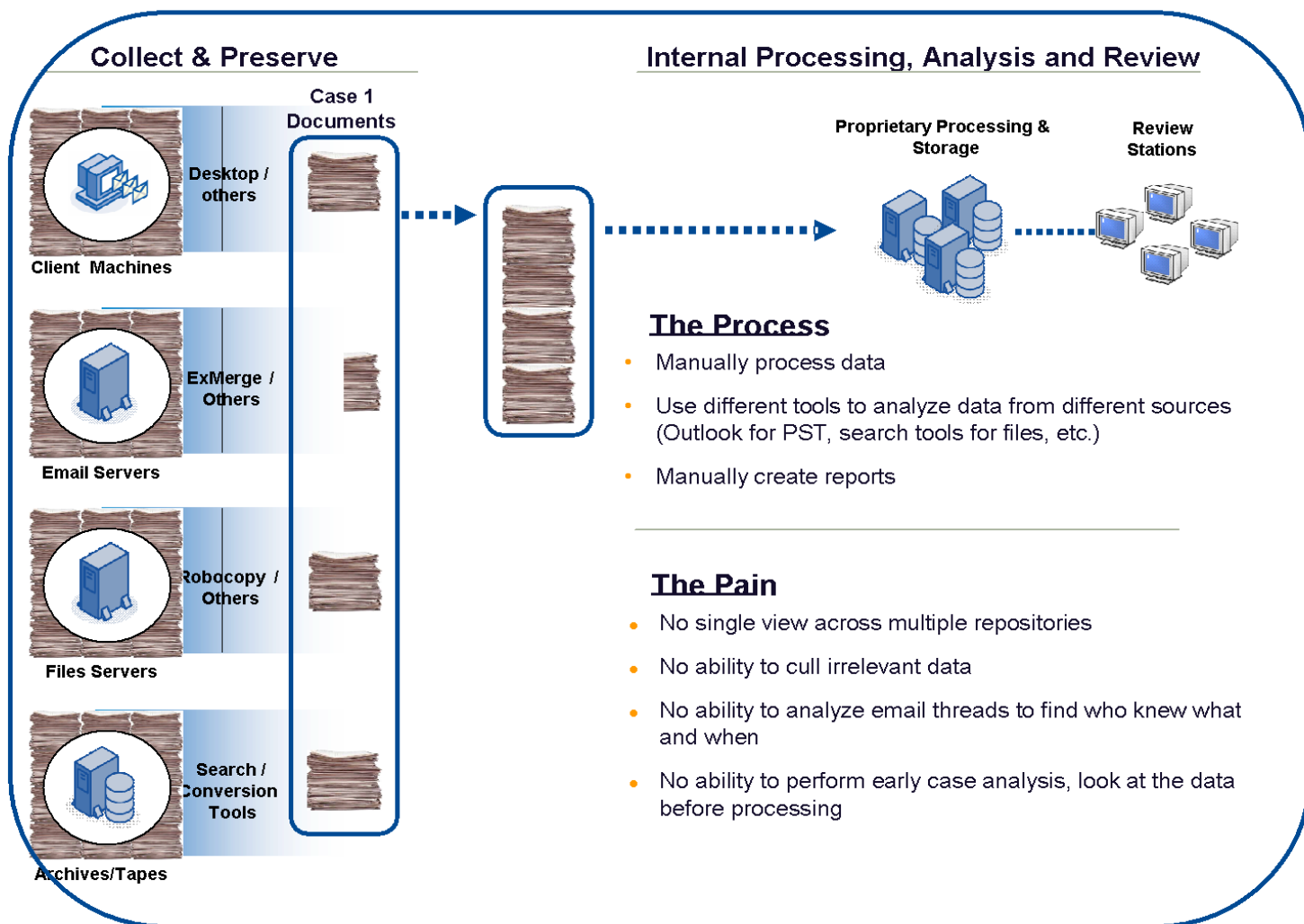
In short, working with a litigation support company like ALC, is the best first step to avoiding sanctions and nightmares. Judge Faccioli & Judge Grimm and other have laid down the stipulations on attorneys that they become experts or hire them. In hiring ALC, you have the experts to avoid the problems and pitfalls of not being a computer geek. Early case planning is key to avoiding the possible sanctions that can come without the planning and verification of the search and cull of your data. We are here to help.

## Bibliography

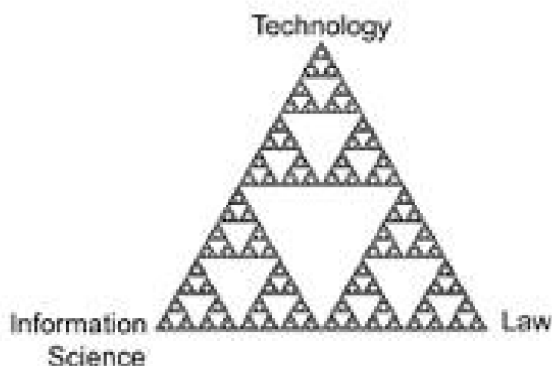
- 1 [Qualcomm Inc. v. Broadcom Corp.](#), No. 05-CV-1958-B(BLM) Doc. 593 (S.D. Cal. Aug. 6, 2007)  
[http://ralphlosey.files.wordpress.com/2007/08/case-qualcomm\\_8\\_6\\_07\\_order\\_on\\_remedy.pdf](http://ralphlosey.files.wordpress.com/2007/08/case-qualcomm_8_6_07_order_on_remedy.pdf)
- 2 IDC 2006 and subsequent 2008 update view here:  
<http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>.
- 3 1 Exabyte is equal to 1024 Petabyte which is 1,024 terabytes which is 1024gigabytes
- 4 [The Sedona Conference® Best Practices Commentary on the Use of Search & Information Retrieval Methods in E-Discovery](#)  
[http://www.thesedonaconference.org/publications\\_html?grp=wgs110](http://www.thesedonaconference.org/publications_html?grp=wgs110)
- 5 [Information Inflation: Can The Legal System Adapt?](#) 13 Rich J.L. & Tech 10 (2007).  
<http://law.richmond.edu/jolt/v13i3/article10.pdf>
- 6 [Intellectual Foundation of Electronic Discovery.](#),  
<http://ralphlosey.wordpress.com/2007/06/07/intellectual-foundation-of-electronic-discovery/>
- 7 [Phoenix Four, Inc. v. Strategic Resources Corporation](#), 2006 WL 1409413 (S.D.N.Y. May 23, 2006)  
<http://ralphlosey.files.wordpress.com/2008/05/phoenix.pdf>
- 8 [Residential Funding Corp. v. DeGeorge Fin. Corp.](#), 306 F.3d 99 (2d Cir. 2002)  
<http://ralphlosey.files.wordpress.com/2008/05/residential.pdf>
- 9 [US v. Johnson](#) (2008 WL 2060597 (E.D. Va May 15,2008)
- 10 U.S. v. Michael John O'Keefe, 2008 WL 449729 (D.D.C.)
- 11 Victor Stanley, Inc. v. Creative Pipe, Inc., 2008 WL 2221841 (D. Md., May 29, 2008)  
<http://ralphlosey.files.wordpress.com/2008/06/victorstanleymomay29-08final.pdf>
- 12 Equity Analytics, LLC v. Lundin, 248 F.R.D. 331, 331 (D.D.C. March 7, 2008).  
<http://ralphlosey.files.wordpress.com/2008/06/equity-vlundin.doc>
- 13 Ethics and Professionalism in the Digital Age: Foreword, *60 Mercer Law Review* \_\_\_ (forthcoming 2008) (with David Hricik)  
<http://www.law.mercer.edu/faculty/bio.cfm?staffid=240>  
<http://ralphlosey.wordpress.com/2008/09/28/lawyers-behaving-badly-understanding-unprofessional-conduct-in-e-discovery/>

## Appendix B

### Graphics & Charts



**Chart 1**  
*Typical data workflow*



**Chart 2**  
*Proposed Pyramid of support*

■ Seattle, San Francisco, Silicon Valley, San Diego, Phoenix, Austin ■ (888) 99-TERIS ■