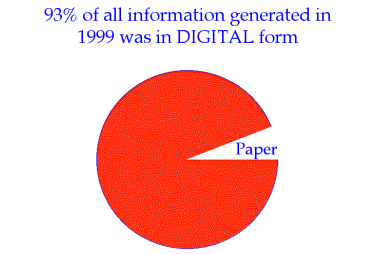
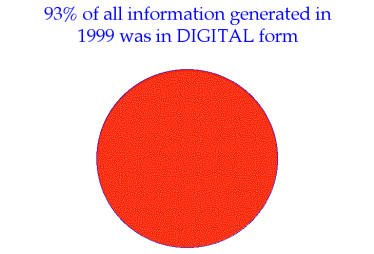
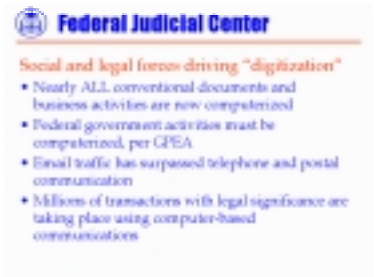
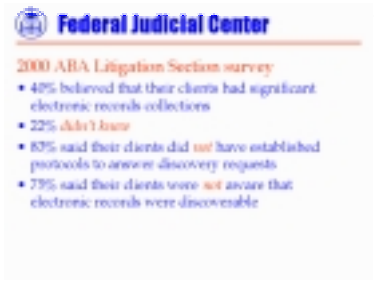
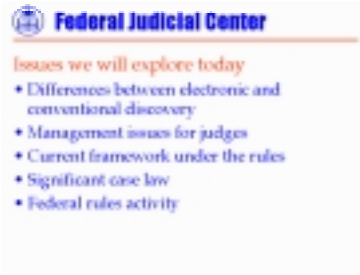
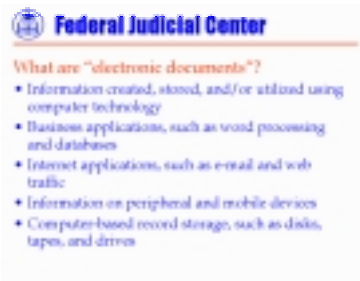
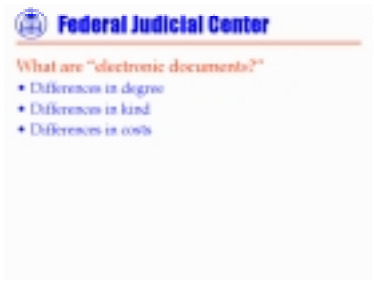
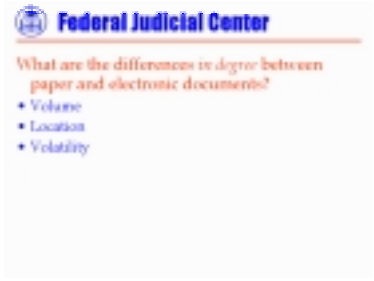
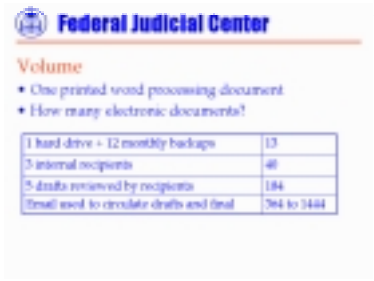
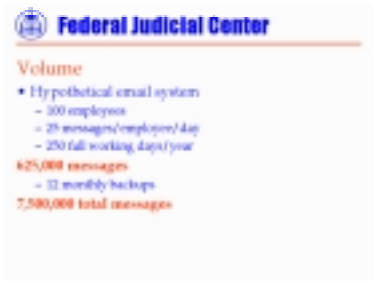
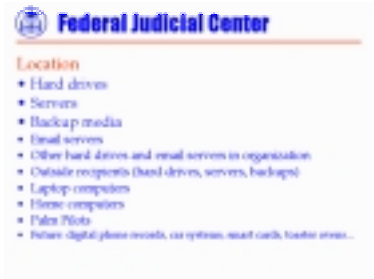
National Workshop for Magistrate Judges
"Electronic Discovery"
June 12, 2002

Slide 01  	[Title slide] *Caveat: Nothing that I say today should be construed as official policy of the Federal Judicial Center, the Judicial Conference, or any other agency of the United States Courts.*
Slide 02  	Recently researchers at the University of California at Berkeley announced that according to their studies… * 93% of all information created during 1999, the last year for which they had complete information, was generated in digital form, on computers of some sort. * That means that only 7% was generated using other media, like paper, phonograph records, clay tablets or smoke signals.
Slide 03  	It is safe to predict that over the next couple of years, even that 7% will shrink. That's not to say that paper itself will disappear. Actually, sales of printer and copy paper are up. But we live in an age in which almost all the information printed on paper, like the conference packet in front of you, is just a manifestation of computer data.
Slide 04  	What is driving this increased "digitization" of society? * First, nearly all conventional documents originate as computer files. Nearly all business activities, from buying gas at the pump to international commodities trading, are transacted using computer-based business processes. * GPEA, the Government Paperwork Elimination Act,

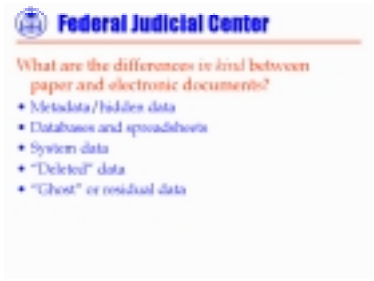
	mandates that all government agency business, to the extent practicable, be computerized by October 21, 2003. * Email traffic exceeded telephone calls and postal use a few years ago, with more than 3.5 billion messages exchanged daily in the US alone. * Millions of transaction with legal significance take place using computer-mediated communications, such as email, the Web, and file exchanges. Products are built and designed, orders are placed, payments are made, goods are shipped, people are hired and fired, all by computer. Everything has been automated, to the point if you complain about the lack of personal service, you send an email and get an automated reply.
Slide 05  	In spite of these sweeping changes in society, business, and law, you cannot expect that the attorneys in front of you are computer savvy. * Two years ago the ABA Section of Litigation conducted a survey of its members. * Four out of ten believed that their clients, mostly businesses, had significant electronic records collections. That seems low, until you understand that * 22% of those polled didn't know. You wonder what cave they've been living in for the past twenty years. * 83% said their clients did not have established protocols for answering discovery requests involving computer data, and * 75% said that their clients were not aware that computer data were discoverable until they faced a discovery request. Remember, Rule 34 was amended in 1970, 30 years ago, to include "data compilations" within the definition of a discoverable "document." So the lawyers are not themselves attuned to the changes that their own clients have wrought, and they have not prepared their

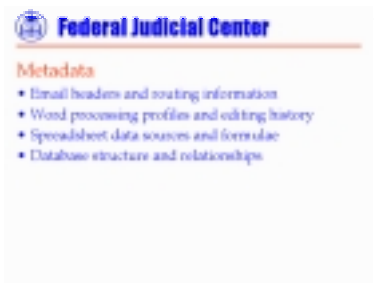
	clients for the legal consequences.
Slide 06  	So what are the implications for discovery? Couldn't we simply ignore the computers, print all the digital information relevant to our cases on paper and continue business as usual, treating electronic discovery just like paper discovery? Or is digital different? * In this session I want to explore with you some of the differences, * what those differences mean for judicial management of civil litigation * and talk about how our current rules apply. * If we have time, we can also look at some of the case law and I can recommend some further reading, and perhaps even * look at where electronic discovery currently stands in the rules process.
Slide 07  	Before I start, I want to make sure we're all on the same page when it comes to defining "electronic documents." By this, I am referring to any information * created, stored, or best utilized with computer technology of any sort. * This includes all the everyday business applications you might use, such as word processing, databases, and spreadsheets; * Internet applications, such as email and the World Wide Web * Devices attached to or peripheral to computers, such as printers, fax machines, pagers, wireless telephones

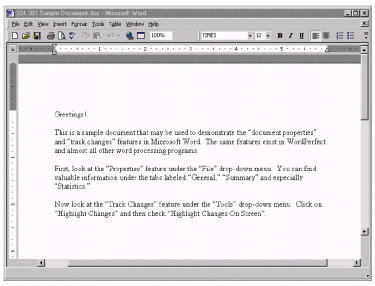
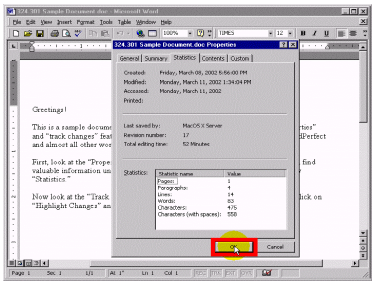
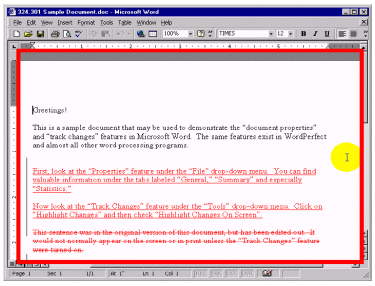
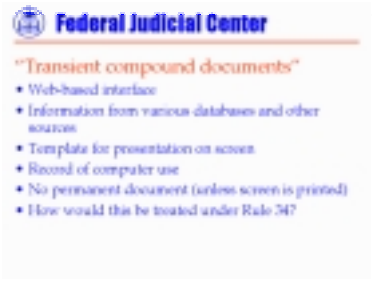
	* And various media used to store computer data, such as disks, tapes, removable drives, CDs, and the like.
Slide 08  	As I said, the first reaction to this by many lawyers and judges is to say, "so what, let's just convert everything to paper and proceed as usual." But it's not that simple. There are significant differences between conventional evidence and electronic evidence. There are differences in degree and differences in kind. And perhaps most important to litigants, there are significant differences in costs.
Slide 09  	Under differences in degree, there are * volume * location * and data volatility.
Slide 10  	Probably the most apparent difference is volume. Several legal scholars and some of the case law point out that electronic data is nearly always far more voluminous than conventional paper data. And it's easy to see why. * Let's say that a paper document, a word-processed memo, has been produced in discovery. * What might that one paper document, originating from a computer, represent in digital terms? * If the document from one desktop PC, there might be a few version or copies on that computer or elsewhere. But if it came from a corporate computer network, and was shared with other employees who commented on it, there could be several dozen to well over 1,000 copies or version of that document in the system.
Slide 11	That's one word processing document. In many investigations the real gold mine, or mine field, is email.

	* Let's say we have a small company with 100 employees regularly using email, and the responding party in discovery needs to sort through the email relevant to a particular person, topic, or date range. * Industry statistics tell us that the average employee with access to email at the workplace sends or receives slightly more than 25 messages per day. * Let's also say this company has 250 full working days in a year. * If we do the math, that makes 625,000 messages to wade through. * And if the company has kept 12 monthly backup tapes * That makes 7,500,000 messages. And to compound the problem, email is seldom organized in any way that makes it easy to search though, and as many observers gleefully point out, email files will likely contain embarrassing, stupid, inappropriate and irrelevant comments just waiting to be discovered in the process.
Slide 12 	Another difference of degree between conventional and electronic discovery is the number of locations that need to be searched when preparing disclosure or in response to a discovery request. We are used to conventional discovery, and having to look though central files, storage boxes, and people's desks. That's bad enough. But it's worse in electronic discovery. * We can start with the most obvious location for digital data, the computer hard drive. * But unless we're dealing with an individual, we will need to also look at the server or servers that computer is networked to,

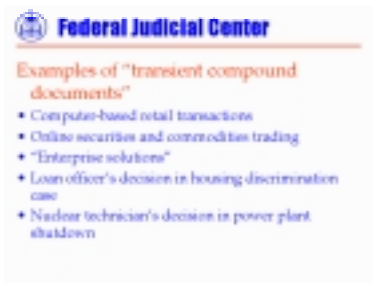
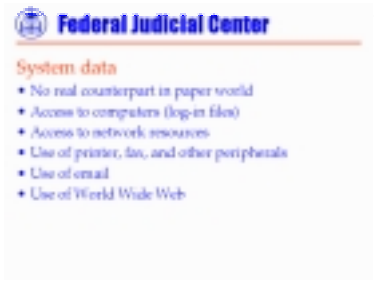
	* and any backup tapes or disks. * There will likely be separate email servers in any sophisticated office * And lots of other computers and servers that files might be copied onto. * If there has been any Internet activity, the number of possible outside computers, servers and backup tapes that may store relevant data explodes. * And then there are laptop computers that everyone carries around with them to conferences like this, * Home computers that people send things to so they can work through their evenings, weekends, and vacations * Palm pilots and other portable devices that contain discoverable data * And who knows what else. Everything has a computer chip and might store discoverable data. My Washington DC subway pass tracks my movements. Rental cars have "On-Star" systems that track your movements, even your speed. Even household appliances are becoming Internet-ready. Someday they may need warnings: "Whatever you toast for breakfast may be used against you in a court of law."
Slide 13 	There has always been a danger in conventional discovery of paper documents being damaged or destroyed or altered. * One of the wonderful things about digital documents is how easy they are to change without leaving any apparent evidence. Everything looks like an original. Which may give you pause.

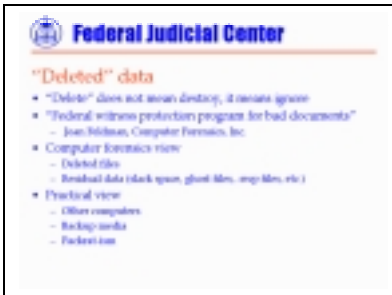
	* Digital files change all the time, even in routine use. Every time a file is opened and viewed open the screen, information about that file changes. Many files are set up to change the date every time they are viewed. * Computer system, whether they be stand-alone computers or vast networks, automatically recycle and reuse memory space, overwrite backups, change file locations, and otherwise maintain themselves automatically, which has the effect of altering or destroying potential evidence, without any human intent, intervention or even knowledge. * This is particularly painful in the discovery process, when well-meaning litigants attempt to secure evidence by downloading it or printing it, only to find that in the process they have destroyed it. * The Gates Rubber case is an example of how supposedly sophisticated computer technicians can inadvertently destroy documents.
Slide 14  	Volume, location, and volatility are differences in degree. You have faced similar problems with conventional paper discovery. * But increasingly you are going to have to deal with differences in kind, where the concept of a document with four corners, whether printed out on paper or reduced to a .tiff image, simply doesn't apply. I'll give you some non-technical explanations and illustrations of these, which include * metadata or hidden data * databases and spreadsheets * system data

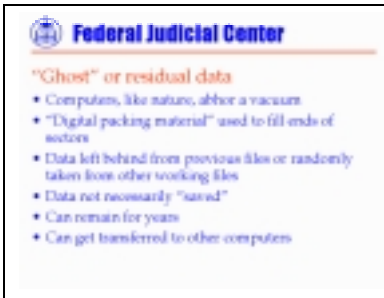
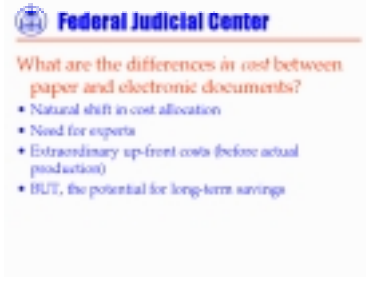
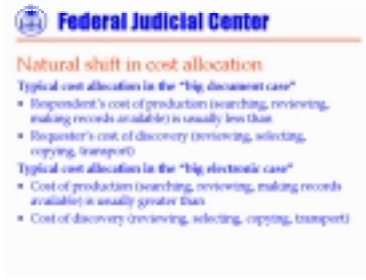
	* so-called “deleted” data * and ghost or residual data residing on a computer hard drive.
Slide 15 	The difference between conventional and electronic discovery cited most often by commentators is the existence of “metadata,” or “information about information.” Metadata is information embedded in an electronic file about that file, such as the date of creation, author, source, history, etc. This information seldom appears on the screen or in a printed version of the document, and often is generated automatically by the software application the author is using, without the author’s knowledge or intent. * Metadata is essential to the proper routing and handing of email. * It is found in nearly all word processing files, * and is the *sine qua non* of spreadsheets * and databases, which are collections of information with no structure or meaning without the metadata. To illustrate, let’s look a little closer at email and word processing.
Slide 16 	Here is an email message, as it might look on your screen or printed on paper. It has what purports to be the author, addressee, date and subject line at the top, and a signature block at the bottom.
Slide 17	Here is the same message, viewed in a text editing program. Now you see a complete header, with lots more information about the history of this email message, including the various servers it went though. It didn’t

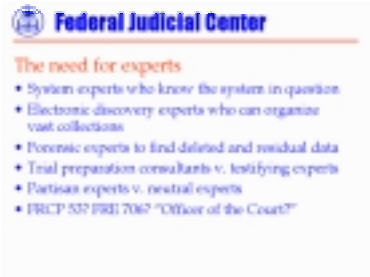
	originate from the Federal Judicial Center, as you might think, but from a different server. And this complete header contains a unique message ID number, which can be used by a computer investigator to track this message throughout the Internet.
Slide 18 	Turning to a word processing document, here we have a short memo in Microsoft Word. What I'm about to show you can be done with a WordPerfect document or almost any other word processing program.
Slide 19 	The document has a profile, automatically generated, which can tell us who the original author was, when it was created, when and how often it has been edited. Most computer users never see this.
Slide 20 	Slightly more frightening is the fact that the document's history is meticulously preserved. Here we see all the editing that has been done on this document in red. Material that was removed is crossed out, and material that was added is underlined. This is true whether or not the user has consciously selected an option under which they can view this themselves. It is almost always there for someone to recover.
Slide 21 	Another difference in kind between digital discovery and conventional paper discovery is the digital transaction that never creates a permanent document, in electronic or any other form. Bob Williams, a consultant with several federal agencies currently converting their business to computers, calls these "transient compound documents," which is a mouthful. These really aren't documents at all, but they are business electronic processes, poised to take over the ordinary course of business, private and public.

	* How many of you flew to this conference? And how many of you booked your ticket on a web site? If you did, you filled out all sorts of information on a screen, or clicked on menus of options, right? * The information that you entered was then matched up with other information from a variety of sources: databases, other web sites, what-have-you. * It was then placed in a blank template and presented to you on a screen: flight number, time, price, tax, credit card information. * There is a record of your having logged in, and a record of this transaction, placed into another database in some format, * but what you see on the screen doesn't exist in any permanent form and will disappear as soon as you move on to another screen, unless you print it out. * So a later document request for your e-ticket will be pointless. There is no such document. There are only integrated databases containing bits and pieces of millions of transactions. If you want to get a feel for how a document request for your e-ticket might be received in some later litigation, just take your printout or your email confirmation to another airline's ticket counter, and see what kind of reception you get. E-tickets are part of a growing class of business processes that have been designed for electronic commerce with very little regard to conventional concepts of record keeping or document exchange.
Slide 22	* The e-ticket is one example that you've likely experienced first hand. Most electronic commerce works in a similar way. *

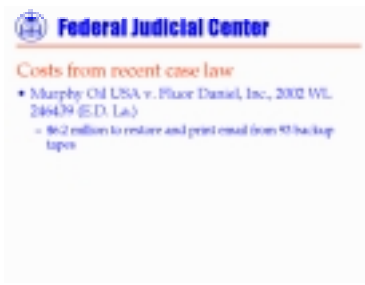
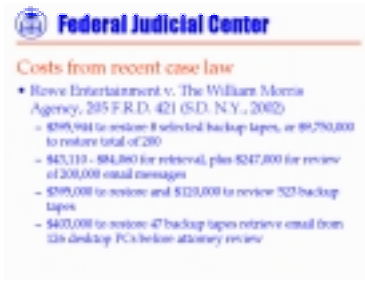
	Online securities trading works much the same way as the e-ticket. * On the wholesale level, giant corporations have computer systems called "enterprise solutions," which manage sales, deliveries, manufacturing, order parts from suppliers, everything in the chain of commerce. * In the services field, a mortgage loan officer makes a decision on a mortgage by looking at a number of windows open on her computer screen. She is networked to databases containing credit history, neighborhood turnover, property tax records, insurance claims, police reports, what-have-you, and maybe has a program that calculates risk based on these various factors. She makes a decision based on an exhaustive report that is never reduced to one computer file, let alone printed on paper. * Or consider a technician in a nuclear power plant, who sits in front of a terminal and makes the decision to close the plant down based on the transient, compound data presented to him on a screen. There is no record of the decision in the conventional sense.
Slide 23 	I mentioned that in an e-commerce transaction, the computer system itself will keep records of your computer use. * These are not detailed records of the transaction itself, but records of your activity on the computer or network, generated usually without your knowledge, by the system itself. When you log on or off, your use of various applications, the web sites you visit, the passwords you use, whether you print or fax a document and what document it is, all are little facts recorded by the computer. Imagine in conventional paper discovery if the file cabinets could talk, if they could be deposed and asked who was "in their drawers," prior to the litigation, so to speak. This is system data.
Slide 24	* By now we should all know that the "delete" keys on our computers do nothing of the sort. They simply rename the

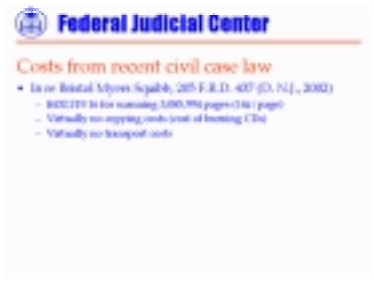
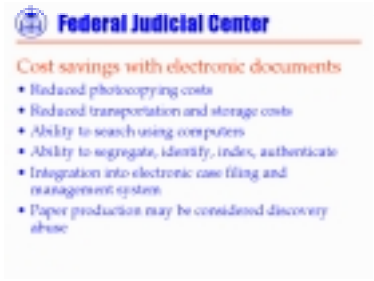
	file and mark the physical space that the file takes up on the hard drive as available for overwriting later, if the space is needed. But changing the identity of a file does not get rid of it. As Joan Feldman, president of Computer Forensics in Seattle likes to say, the delete function is * like a "federal witness protection program for bad documents." * There are two ways to approach this problem. One is the "computer forensics" view, which * claims that all files can be recovered, albeit sometimes at great cost, * from unallocated space, slack space, "swap" files, and other places that the experts can explain to you. * Then there is a practical view, which says that you don't need to be an expert to understand that if dozens of copies of a file exist, at least one copy can always be found * on another computer * on a backup tape * or on a floppy disk kept behind someone's desk. That's not to say that finding a copy of that file will be any cheaper or easier than recovering the deleted version from a hard drive. Either way, the possibility that a deleted file could be recovered or retrieved is a temptation to engage in electronic discovery on a much broader scale than is usually contemplated in conventional paper discovery.
Slide 25	Another aspect on electronic discovery that grabs headlines is the ability to recover information from areas of computer memory that most people don't know exist. This is the really sophisticated part of computer forensics, and hopefully is something you won't see too often in civil cases. Here's a simple explanation of what "ghost"

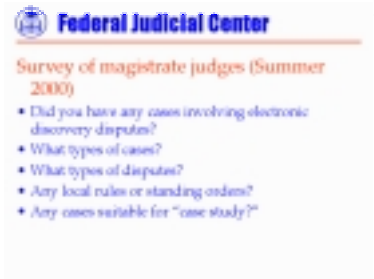
	or residual data is. * Computer hard drives can't have any space on them that is truly empty; that goes against the laws of physics. Data is stored on hard drives in little containers, all the same size, called sectors. A particular computer file might take up many sectors, which have a tendency to appear all over the hard drive, but seldom to they fill every sector completely. There is always a little leftover space at the end of the last one. That leftover space needs to be packed with something, and what the operating system does is find some random data to fill it up, like you might fill out a box of books with some old newspapers. That data could be something left over from a previous file, or something randomly grabbed from somewhere else. It could be something the user never intended to save. It could remain for years, and if the data is transferred wholesale onto a new computer, it might come with it. Have you ever unpacked a box and found that the old newspapers crumpled up inside are really more interesting than the other contents? That's what hidden or residual data is.
Slide 26 	Aside from differences in degree and differences in kind between conventional paper discovery and electronic discovery, we have differences in cost. Whether these are differences in total cost, or differences in the allocation of costs or the timing of costs remains up for debate, and depends on the case and the management skills of the parties. But lets explore some of these cost differences and see where they come from.
Slide 27 	* There is no hard data on this, and commentators can disagree, but the general consensus among practitioners in conventional discovery was that costs in the "big document" cases shifted to requesting parties, when responding parties would simply make warehouses full of documents available for the requesting party to go through. * While there were costs to the responding party in

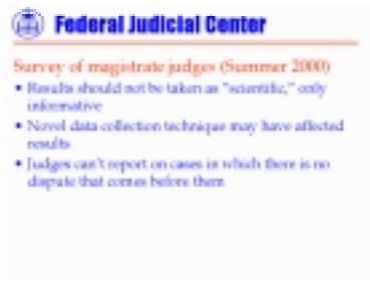
	organizing material and pulling out privileged matter, * the high cost of selection and copying borne by the requesting party served as a natural break on the volume of discovery. * In electronic discovery, this paradigm is reversed. * The cost of locating, reviewing, and preparing vast digital files for production is perceived to be much greater than in conventional discovery, * while technology has dramatically reduced the cost of searching and virtually eliminated the cost of copying and transport.
Slide 28 	* Nearly all electronic discovery involves experts of some sort and employment of experts adds to costs But we need to be careful to identify who these experts are and what their role is, because that can have a tremendous impact on the bottom line for the parties. * At the operational level we have systems experts. These are the techies who know the computers, software, and files at issue in this case. Most likely they are employees of the parties, the IT or MIS people. They must be involved, just as the records keepers and file clerks needed to be involved in conventional discovery. * Then there are outside experts who are brought in to conduct electronic discovery. This has become an industry. Their role is to take the data collections, convert them into indexed and reviewable files, and make them ready for production. In the process, they have to solve a number of logistical problems. They are more expensive than the in-house systems people. * And then there are the forensic examiners. They are brought in to find the smoking guns: the deleted

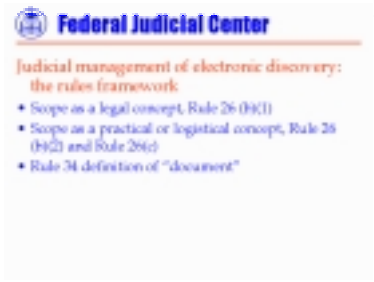
	document, the missing email, the hidden files and ghost data. They are very expensive and highly specialized. A judge may be wary when a party announces at the outset that it has retained a computer forensics expert. That could be the first sign of trouble, when costs could get out of control. * We need to make a clear distinction between the consulting experts who are assisting counsel with discovery, and possible expert witnesses, who would be subject to discovery themselves. Whenever possible, we want to clear the way for the systems people and consulting experts on both sides to be able to communicate directly with each other. Most of the problems with electronic discovery are technical and logistical problems, and the worst way of dealing with them is through attorneys. Often when we remove the attorneys from the process, the techies can work things out amongst themselves. This may require some judicial supervision of the process and some protective orders so that no privileges are waived, but the judicial time and energy saved will be tremendous. * Some judges have opted to appoint a neutral expert, especially where the parties are contentious and want access to each other's computer system. Using a neutral may be the cleanest way to deal with privilege and privacy issues when ordering an inspection of computers or hard drives, and may result in cost containment for the parties, but it isn't absolutely necessary. * If neutral experts is appointed, their role needs to be narrowly defined. They may be considered "special masters" under Rule 53, but if they are likely to testify on some evidence issue, they become witnesses under Evidence Rule 706, which opens them up to discovery. I have seen orders in which this issue has been dodged by appointing a neutral "officer of the court" without reference to any rule.
Slide 29	Let's look at a couple of recently-reported civil cases for concrete examples of electronic discovery costs. *

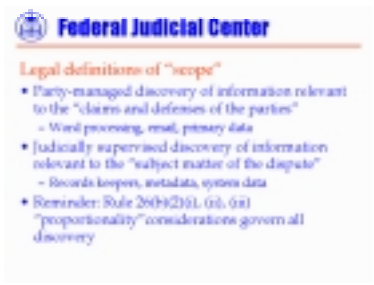
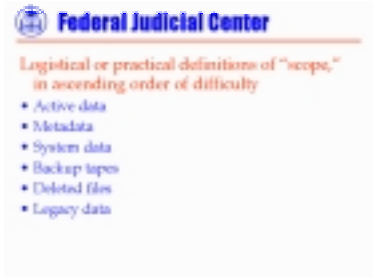
	In Murphy Oil, 93 backup tapes were involved. Backups need to be restored to a workable computer system, then converted into some format so that the individual files can be organized, searched, and read, and then reviewed by attorneys for relevance and privilege before production. An electronic discovery expert, and is essential to this process, and occasionally a forensics expert is needed as well. * That process was estimated at $6.2 million BEFORE attorney review of the resulting files for relevance or privilege. Normally, this would be a cost borne by the responding party, but the court intervened to reallocate costs.
Slide 30 	The Rowe Entertainment case from the Southern District of New York is quickly becoming a classic on the issue of discovery cost allocation. The opinion of Magistrate Judge Francis summarizes estimates from computer experts for electronic discovery of four defendants. I should note that Judge Francis' opinion was affirmed by District Court Judge Patterson about a month ago. * For the first defendant, complete restoration of 200 backup tapes would cost $9,750,000, but restoration of eight randomly selected tapes, just to see if there was any evidence at all on them, could be done for a mere $400,000. * For the second defendant, 200,000 email messages could be retrieved for between $43,000 and $84,000, and attorney review was estimated at $247,000. * For the third defendant, with 523 backup tapes, restoration was estimated at $395,000 and attorney at $120,000. * And finally, restoring 47 backup tapes and retrieving email from 126 desktop computers was estimated to cost just over $400,000. Judge Francis articulated eight factors to consider in allocating those costs between the plaintiff and the defendants, and I recommend reading the case if you

	haven't already.
Slide 31 	Those seem like some pretty bad examples of costs, but they are not atypical. On the other hand, if we aren't engaged in restoring backup tapes or searching hard drives for lost email, the cost picture can be quite different, even in very large document cases. * In the Bristol Myers Squibb Securities Litigation, over three million pages of pre-existing paper documents were scanned and converted into computer images for a cost of $432,000. That may seem like a lot, but that's only 14¢ per page, and by doing so * further copying costs could have been virtually nil * as well as transport and storage costs. The defendant chose not to make these electronic files available to the plaintiff, and thereby artificially drove costs up.
Slide 32 	Which points to the possibility that electronic discovery could be very economic, if handled properly. The discovery and use of electronic evidence can greatly reduce costs, reduce time, and facilitate the pretrial preparation process. * The more discovery is conducted in electronic form, the lower will be the costs for photocopying, * transportation and storage. * Computer file searching technology allows us to find key words, dates, names, and other textual items in a matter of seconds, even if the data collection is the equivalent of millions of pages of paper documents. You've all experienced this first-hand using Westlaw and Lexis. * When properly managed, electronic discovery allows us to segregate, identify, index, and even authenticate documents in a fraction of the time and at a fraction of the cost of paper discovery.

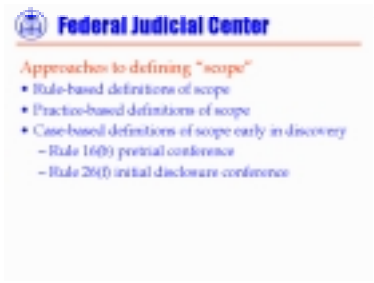
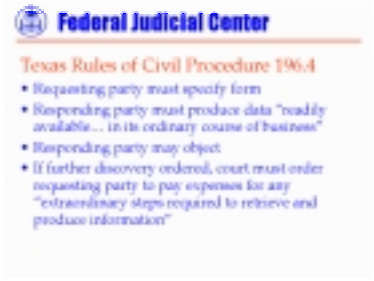
	* When well planned, electronic discovery leads naturally into electronic case preparation and electronic trial presentation, which statistics tell us reduces trial time by as much as one-third. * These cost and management advantages are so great, that there are now civil cases in which an insistence by one side or the other to engage in paper discovery is considered an abuse.
Slide 33  	Let's move on to the types of cases most likely to involve electronic discovery, and the types of management issues you are most likely to face. * The FJC Research Division conducted a survey of United States Magistrate Judges in the summer of 2000. The questions we asked were: * Did you have any cases involving electronic discovery disputes? * If so, what types of cases, and * what types of disputes? * Were you operating under any local rules or standing orders that governed electronic discovery? * And finally, did you have any particular cases that might make good subjects for an in-depth study of electronic discovery?
Slide 34	While the results can't be considered authoritative or scientific, they are informative. On the question of the types of cases in which you found electronic discovery disputes, the results were a little surprising. The highest percentage reported disputes arising in individual employment cases, which usually are not "big cases" in conventional discovery disputes, and the lowest

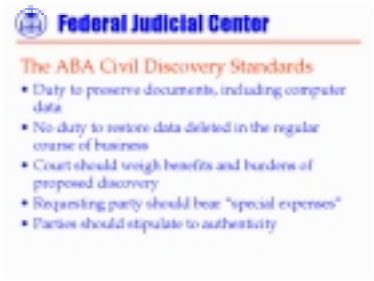
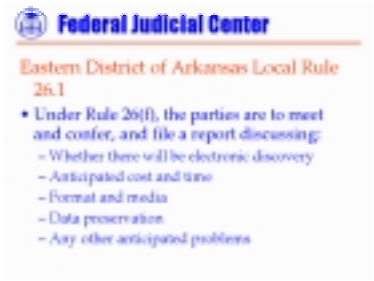
Types of cases in which judges report electronic discovery disputes | Case Type	% of judges reporting at least one case involving electronic evidence
Employment – individual plaintiff	59
General commercial litigation	55
Patent/Copyright	44
Employment – class action	25
Product liability	24
Other	23
Construction litigation	10
Securities litigation	10
Antitrust	8		percentages reported disputes in the antitrust and securities areas, which are nearly always document-intensive in conventional discovery. It is difficult to compare this to the overall federal docket, because we don't have statistics about conventional discovery. But when we look at these numbers and numbers on a comparable subset of the overall federal docket, what is striking is that there is no significant disproportionality. In other words, don't assume that just because a case is small it won't have a electronic discovery dispute, or that a large case necessarily will.
Slide 35 Types of electronic discovery issues judges experience 	Issues/Experiences	% of judges with at least one case with this issue	% of total cases reported involving this issue
---	---	---
Involvement of computer experts	69	25
Privilege waiver	49	15
On-site inspection	48	15
Sharing of discovery costs	48	15
Alleged spoliation	47	13
Data preservation order	35	10
Sharing of production costs	35	9
Increased efficiency	21	13		On the question of what types of electronic discovery disputes we are seeing in courtrooms, we should first note that 39% of judges who responded reported NO electronic discovery experience. This chart reflects the experience of remaining judges, two-thirds of whom reported that electronic discovery involved the use of an expert or consultant hired by the parties. About half reported experience with four types of situations: possible waiver of privilege for inadvertent production of privileged material, a request for on-site inspection of computer hardware or media, the sharing of discovery costs, and allegations of spoliation of discoverable data. About one third reported consideration of a data preservation order and the sharing of production (as opposed to discovery) costs. Only 21% reported seeing any cost savings as a result of electronic discovery.
Slide 36  	Again, * the results of this survey should not be taken as "scientific," but only as "informative." * The FJC used the Internet to administer this survey, and that alone might have skewed the results. The FJC certainly had a lower response rate using the Internet than conventional paper-and-pencil surveys, and we had to resort to a conventional follow up to get a useable number of responses. * More importantly, the judges couldn't report on cases in which no dispute came to their attention. For all we know, the majority of cases involving electronic discovery

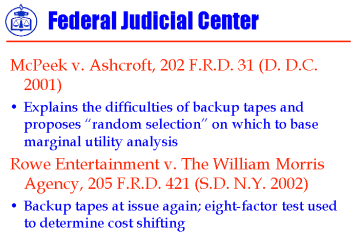
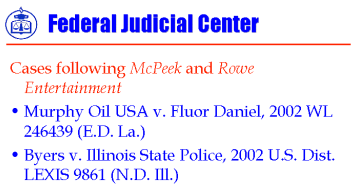
	run smoothly and require no judicial intervention. We would need a different type of research project to answer that question.
Slide 37  	What is the framework within which you can manage electronic discovery and resolve disputes? * The central issue in almost all discovery management is the determination of scope. In electronic discovery, we have two ways of thinking about scope: * The legal scope, governed by Rule 26(b)(1), * and the practical scope, which is often expressed in cost/benefit terms under Rule 26(b)(2) or Rule 26(c) * I should just mention, more or less as an aside, that the definition of a document under Rule 34 does not act as a limitation on the scope of electronic discovery. While academics and pundits may debate whether metadata, deleted files, or backup tapes are or are not "documents," no case since 1970 has construed the definition of "document" to exempt any form of computer data from discovery. Judges have questions from the start of the case about preserving computer data and preventing (or occasionally sanctioning) spoliation. There is the question of what has been called "heroic" data retrieval, which is related to practical scope. There is the question of protecting data after discovery so it can be used as evidence at trial; chain of custody, stipulations as to authenticity, and the like. There are heightened concerns for privacy and privilege in the computer context. And underlying it all are concerns for cost and cost allocation. But all of these issues tend to be manifestations of disagreement or confusion over scope.
Slide 38	Turning first to the legal definition of scope under Rule 26, as we all know the scope of discovery under Rule 26 has been bifurcated by the amendments of December 2000. *

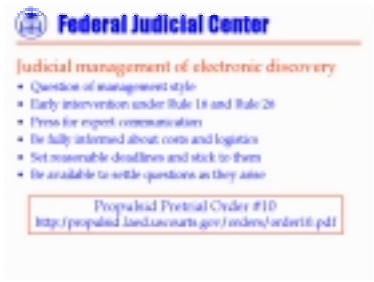
	We have party-supervised discovery, which is limited to information relevant to the "claims and defenses of the parties," and * judicially-supervised discovery of information relevant to the "subject matter of the dispute." * But it should be borne in mind that under either standard or procedure, the judge always has the power to intervene in discovery and apply the "proportionality" considerations of Rule 26(b)(2)(i),(ii) and (iii), whether or not any party has applied for a protective order or filed a motion to compel. The judge can establish procedures or set limits before things get out of hand.
Slide 39 	Whether or not the particular computer data in question should be subject to discovery, and therefore to preservation or recovery, goes to the second scope consideration, which is logistical as opposed to legal. Given that there may be an abstract right to discovery, should the judge allow it, applying the cost/benefit analysis of Rule 26(b)(2)? * Here I've set up a hierarchy of different types of computer data. Assuming that the information being sought is relevant to the claims and defenses or the subject matter of the dispute, you would probably have no problem at the top of this list, requesting and obtaining the production of active data, available to the responding party in the ordinary course of business. * And a strong argument can be made for metadata, for while it might not be as obvious as active data, it is usually available in the ordinary course of business, and costs little to produce. * System data may get a little more costly and remote. One needs computer operators to locate it, and there are questions about how to render it in an understandable and useful way. *

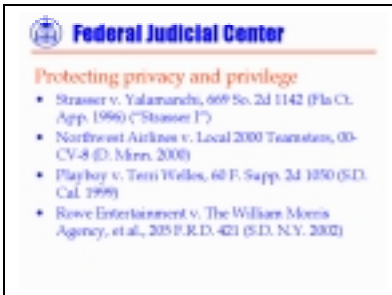
	We get into real cost and logistics problems with the discovery of backup tapes, which are designed for restoring computer systems in the event of disaster, and not for the retrieval of individual files. Each tape must be restored to a computer system with the appropriate operating system and software, and then searches must be performed to find individual files. The restoration itself is costly and time-consuming, before you can determine if the files are truly relevant to the discovery request. * As I said, deleted files can be recovered, but this takes the active intervention of computer forensics people and is a costly and speculative enterprise. We are getting further from data available to the respondent in the ordinary course of business. * And at the bottom of my list is legacy data, which I want to spend a minute on. This is data that for whatever reason has been saved, although it was created on obsolete computer systems using obsolete operating and application software. Let's use a simple example. How many people here remember 5-1/4" floppy disks, the floppies that actually were floppy? Now, if you found a 5-1/4" floppy disk today in your desk, what would you do with it? Any ideas? You'd first have to find a computer with a 5-1/4" floppy disk drive. Then you'd have to guess what the operating system would have been for that computer, perhaps some version of MS-DOS. Then if you found there was a file on it, what would you do if that were a WordStar document? Believe it or not, many large corporations and government agencies have tombs full of this kind of data, waiting for the day when a digital archeologist will dig them up and bring them back to life, like the Linear B clay tablets in the basement of the British Museum. I my experience, I've stumbled on 8" floppies, reels of corroded magnetic tape, and shoeboxes full of IBM punch cards labeled "Retirement Plan Annuity Records: Retain for 50 years."
Slide 40	To cope with the question of the scope of electronic discovery, different courts have come up with different approaches, and these approaches may be instructive. What is a rule in one jurisdiction may be a valid judicial management strategy in another.

	* One approach is to define the practical scope of electronic discovery by rule. * Another approach is to look outside at standards of practice or protocols developed by bar associations, or something like the Manual for Complex Litigation. * And another approach, much more in keeping with our conventional view of party-controlled discovery, is to have the parties define scope on a case-by-case basis, as part of the Rule 26(f) disclosure and consultation and the Rule 16(b) pretrial conference.
Slide 41 	One state court has developed a specific civil procedure rule on electronic discovery, Texas. * Under the Texas rule, the requesting party must be specific in the request as to the form in which they want production. * The responding party's obligation is limited to producing data "readily available... in its ordinary course of business." This language has yet to be construed by any Texas court. Presumably it excludes backup tapes and deleted data. Whether metadata or system data is "readily available in its ordinary course of business" is an open question, but the clear intent of the Texas rule is to restrict the scope of electronic discovery on the basis of accessibility and cost. * The responding party may object to any further production, * and if further production is ordered by the court, the requesting party must pay for any "extraordinary steps required to retrieve and produce information." Again, this language has not been construed by any Texas court.
Slide 42	The American Bar Association has weighed in with recommended standards of conduct for lawyers and

	judges in electronic discovery. Standard 29 of the ABA's Civil Discovery Standards states * that attorneys have a duty to preserve computer data pending discovery * that there is no duty to restore data deleted in the ordinary course of business, presumably prior to litigation * that the court should exercise its powers under Rule 26(b)(2) and its inherent power to weigh the benefits and burdens of proposed electronic discovery, * that the requesting party should bear "special expenses," much like the Texas rule, and * that the parties should stipulate to the authenticity of data produced in discovery.
Slide 43 	The two Federal District Courts in Arkansas have adopted matching local rules governing electronic discovery. They take a different approach than Texas, in part because local rule-making authority in the federal courts is restricted. * Under the Arkansas local rules, the parties must meet and confer regarding electronic discovery under Rule 26(f). They must file a report with the court, stating * whether there will be electronic discovery * the anticipated cost and schedule * the format and media for production * any efforts taken to preserve data pending discovery, and *

	any other anticipated problems.
Slide 44 	The Federal District Court in Wyoming has a similar rule, but it targets email as a particular problem. * Under Wyoming local rule 26.1, the parties must meet and confer regarding * data preservation * the scope of email discovery * how to deal with inadvertent production of privileged email * whether they plan to discover deleted data, and * whether they plan to discover backup data.
Slide 45 	[text accompany this slide was not transcribed]
Slide 46 	[text accompany this slide was not transcribed]
Slide 47	[text accompany this slide was not transcribed]

	
Slide 48 	The rules of thumb we derive from the case law and from informal discussions with the judges who have been through these battles are: * It's always a question of individual management style. * But the key is early intervention, using Rule 16 and 26. * Get the parties' experts involved in working out the discovery plan. If you are going to need to determine scope as a logistical concept, they know the logistics * Be fully informed as to what the costs and procedures will be. * Set reasonable deadlines and perimeters, and don't waiver. * Make yourself available to answer questions and settle disputes as they arise. You might find that problems disappear when the parties know the judge is willing to step in and decide immediately. * For an example of a very thorough electronic discovery order, albeit from a large and complex case, see the order issued recently in the Propulsid litigation. While most cases won't need this sort of exhaustive treatment, it can serve as a pattern for a more tailored approach.
Slide 49	Now I want to move into the case law, but with a few caveats. First, the vast majority of electronic discovery cases reported by the legal publishers are decisions at the



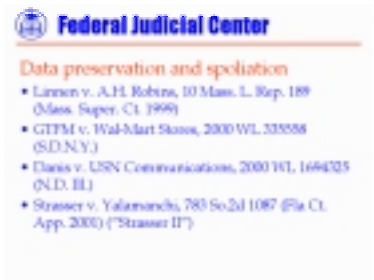
trial court level, and of those, the vast majority are memoranda and orders on non-dispositive procedural issues. Therefore we have precious few circuit court decisions and only one Supreme Court decision that touches on electronic discovery, and no real binding precedent out there, and don't expect any. The cases we will talk about serve as examples of how other judges have dealt with the issues before them. We can learn from their experience, but we're not really talking about common law development here. Second, I never want to put myself in a position of second-guessing any magistrate or district court judge who is working in the trenches. You have the parties and the facts before you, you are listening to the arguments and you have to render a decision. So for my money, you are always right. With the advantages of 20-20 hindsight, great distance, and total ignorance, others might come to different conclusions, but I'm in no position to criticize and no criticism is intended.

*

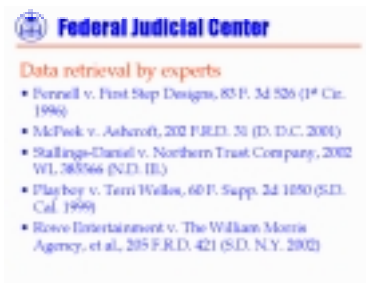
The first case I want to mention is actually a state court case from Florida, but it is cited in many federal cases for the proposition that broad discovery of computer data, such as allowing the requesting party direct access to computer hard drives, has the potential to compromise privilege and invade privacy. Strasser involved the breakup of a medical practice and a suit between two doctors, one of whom requested access to the opponent's computer hard drive, which also contained confidential patient records. The court had to fashion a rather strict protective order.

*

Protecting privilege and privacy in discovery is always an issue, but more so in electronic discovery, where people use their computers for all manner of personal as well as business dealings, and particularly use email in ways that make us all shudder then they are exposed to the light of day. We shudder because we're all guilty of the same behavior, not because the respondent is particularly stupid. Northwest Airlines accused its flight attendants' union of organizing an illegal sick-out over the New Year 2000 weekend, and discovery zeroed in on two union activists in particular. Northwest asked for an order allowing them to go into the defendants' homes, make

	complete, forensically-sound duplicates of the hard drives of their home PCs, and then review the contents for relevant documents. The Magistrate Judge ordered a search protocol involving a neutral expert, but the press called this an outrageous invasion of privacy on the Wall Street Journal, which isn't known for sympathizing with union activists. The discovery resulted in nothing and the case was dismissed for lack of evidence, but not before raising a lot of questions about what differences there are between electronic discovery and conventional discovery that might justify this sort of procedure. * In Playboy v. Terri Welles, a home computer was also at issue, and the court developed a careful protocol in which a neutral, court-appointed expert would conduct the discovery and turn over potentially responsive files to the respondent's attorney, who would have an opportunity to review them for privilege and relevancy before production. * In Rowe Entertainment, the protocol was switched. The neutral expert was to conduct discovery, and turn the potentially responsive files over to the requesting party's counsel on an "attorney's eyes only" basis. Counsel would then select the documents relevant to their discovery requests, and present them to opposing counsel, who could then raise any objections they thought appropriate. Although I think this procedure is rather extraordinary, it has the benefit of significant reducing costs for the producing party.
Slide 50 	The next issue is data preservation, and its flip side, spoliation. The most cited case in this area is Linnen v. A.H. Robins, which while being a Phen-fen diet drug case, luckily was not a federal one. There the defendant's attorney didn't know that their client had several hundred backup tapes preserved from prior litigation, and that the client was in the process of destroying them during discovery in this litigation. The state court judge was understandably upset at the repeated representations of counsel which turned out to be baseless, and sanctioned the defendants with a spoliation inference. The case settled shortly thereafter. *

	GTFM was another case in which counsel was ignorant of the client's activities, but slightly worse, in that the client may have misrepresented the facts to counsel. Early in discovery, the defendant responded to interrogatories stating that computerized sales records were routinely destroyed after only a few weeks, and therefore would not be available for the relevant time period. Later in a deposition, it was revealed that they were routinely kept for a much longer period, and would have been available early in discovery, but had now been destroyed. More sanctions, and more lessons for counsel to prepare discovery and data preservation plans early in the case. * Danis v. USN Communications was a case in which the defendant had no coherent electronic records management system in place in the ordinary course of business, and compounded the problem by having no coherent procedure for preserving data during the pendency of litigation. The case was further complicated by plaintiffs who kept asking for the same data, not knowing it had already been produced, because it didn't understand the production. Both sides spent 3/4 of a million dollars litigating this discovery dispute, and both sides lost. Perhaps more importantly, the CEO of the defendant corporation was fined personally for his failure to take reasonable measures to preserve data. * And finally in this category we have the second chapter of the Strasser case, which I cited before. It seems that while the parties were spending several months working out the details of how to conduct computer discovery to preserve the confidentiality of third party information on the computers, the defendant was using this time to systematically destroy the data. The judge was not pleased and issued stiff sanctions. The opinion is well worth reading, particularly since defendants always like to cite the first opinion and forget this one.
Slide 51	By "heroic" data retrieval, we mean the efforts by computer forensics experts to recover deleted data, restore backup tapes, or resurrect legacy data to obtain information that is not available in the ordinary course of business. The question for the judge is whether this expensive and intrusive form of discovery is justified under Rule 26(b)(2)'s balancing of benefits and burdens,



*

The First Circuit weighed in on this question in Fennell v. First Step Design, which is often cited for the proposition that the mere assertion by a requesting party that such efforts might result in the discovery of relevant information is not enough. There must be some proof or at least a solid theory that discoverable information will be obtained. The problem with this case is its procedural posture. The plaintiff made her discovery request after discovery had closed, in the face of a summary judgment motion. So strictly speaking the holding was not based on Rule 26, but it is cited often in relation to Rule 26 and judicial discretion to limit digital fishing expeditions.

*

McPeek v. Ashcroft is less authoritative, but much more solidly on point. Judge Facciola gives an excellent explanation of why backup tapes are such costly and speculative sources of otherwise allowable discovery, and orders that discovery in the first instance be limited to a small sampling to determine its value. A very practical exercise of judicial management powers.

*

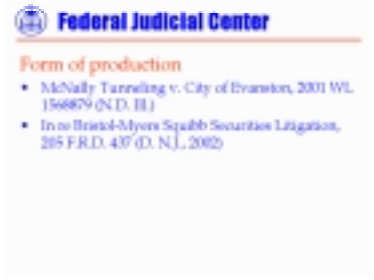
Stallings-Daniel is a very recent case that presents similar facts as Fennell v. First Step Designs, but without the procedural quirk. Again, the relevant inquiry is the likelihood of discovering relevant information, weighed against the cost and intrusiveness of the proposed discovery. Mere speculation that there may be discoverable data isn't enough.

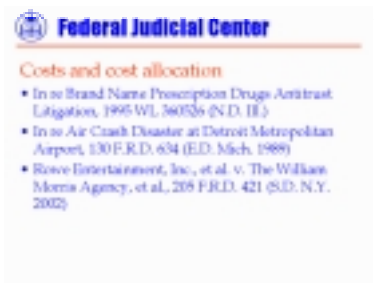
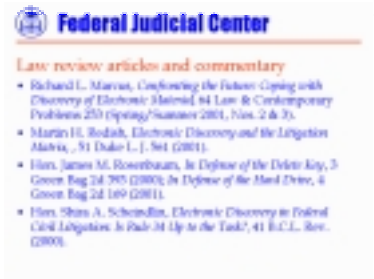
*

I mentioned Playboy before. Here the operative fact was that the defendant admitted that she had deleted relevant email, so "heroic retrieval" was justified.

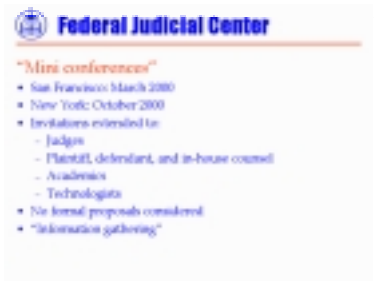
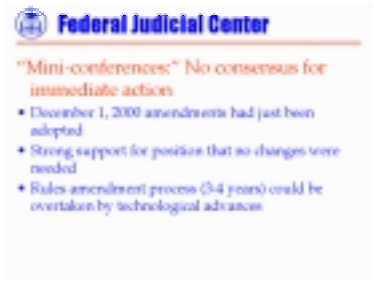
*

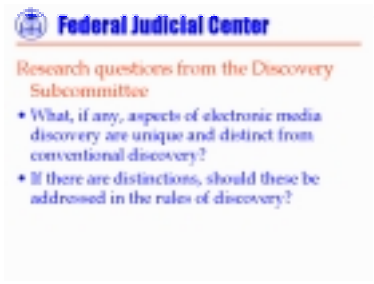
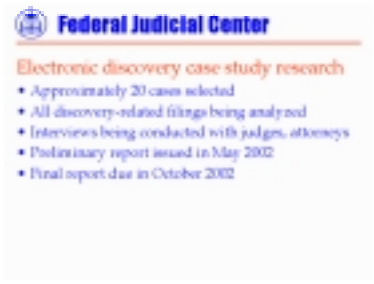
And in Rowe, the defendants first objected to the costly recovery of email from backup tapes, claiming that one the one hand, email was not used for relevant corporate communication, and on the other hand, there was so much that it would be too costly to recover. The judge thought this was a contradiction. If email wasn't used for ordinary business, why was there so much of it? So while he pared down the plaintiff's request and shifted costs, the

	restoration of the backup tapes was ordered.
Slide 52 	In McNally Tunneling, the judge concluded that authority is split on whether a party is entitled to discovery in electronic form in addition to paper form, citing prior cases that came on both sides of that issue. So the judge exercised discretion and denied the defendant's request for computer files to supplement the plaintiff's paper production, as not supported by any demonstration of need. * Early in the Bristol Myers litigation, the parties had agreed to paper production and a per-page price for photocopying. However, the defendant did not disclose that the documents had been scanned, were being "blown back" in paper form at a cost below that of photocopying, and were available in electronic form for considerably less money. The court held the parties to the agreement to produce paper, but at the lower cost of the "blow backs," and ordered that the electronic versions also be produced, at the nominal cost of duplicating compact disks. The court rejected the defendant's argument that the plaintiff contribute to the cost of scanning the documents, as that action was taken unilaterally by the defendant, who didn't inform the plaintiff, for its own purposes. I think what the court really objected to were the games that the defendant was playing with the form of production, which had no purpose but to drive up the plaintiff's costs.
Slide 53 	Gates Rubber v. Bando Chemical, an old case by our standards; 1996. Here the parties didn't use a qualified expert to handle the evidence, and regretted it. On behalf of the requesting party, a well-meaning amateur incorrectly installed some commercially available software to recover supposedly deleted files from the respondent's computer, and ended up destroying 7 or 8 per cent of the potential evidence on the hard drive instead, as well as hopelessly compromising the ability to authenticate any remaining evidence. The judge declared that there was a duty for those getting into electronic discovery to use the best available methods, and thus was born the electronic discovery consultant industry.
Slide 54	Closing out this review of the case law with the issue of costs and cost allocation, we have what looks like a confusing picture in the case law, and the pundits claim that the case law is contradictory. I don't agree.

	* On the one hand, we have cases like Brand Name Prescription Drug, which stand for the proposition that if a business decides to use computer technology to create and store business records, it should anticipate the cost of using that technology to respond to discovery requests. * On the other hand, we have cases such as the Detroit Air Crash case which stand for the proposition that if a plaintiff asks for computer data that isn't readily available in the ordinary course of business, or in a format solely to suit the needs of discovery, the plaintiff should be willing to pay for that. * In Rowe Entertainment the issue is email, and particularly who is to bear the cost of searching through vast email files for relevant messages. First, Judge Francis reviewed in detail the cost estimates of the experts, which is very educational. It really gives us an excellent view of how extraordinary these costs are. Then Judge Francis applies an eight-factor test to determine the most appropriate cost allocation, based on a comprehensive review of the case law. Recommended reading for all.
Slide 55 	Legal academia has been paying attention to the problems of electronic discovery, too, and several articles on this topic have been published. Highlighting just a few, we have * Prof. Marcus' article last summer in Law & Contemporary Problems. Prof. Marcus is reporter for the Discovery Subcommittee * Prof. Redish's article last November in Duke Law Journal, which proposes a comprehensive new regime for electronic discovery * Judge Rosenbaum's two articles in Green Bag, which highlight the privacy aspects of electronic discovery * and Judge Scheindlin's article nearly two years ago in the

	Boston College Law Review, which proposes language changes to Rule 34.
Slide 56 	I have been working on electronic discovery and related issues for several years now, and in my copious spare time have put together an unofficial, non-sanctioned web site where I post articles, seminar proceedings, PowerPoint slide sets, and other resources. Among the items you can find there are * some of my own articles and presentations * former federal judge Barbara Caulfield's recent on cost bearing in electronic discovery * Redgrave and Hiser's article on the ever-widening scope of electronic discovery * a debate on whether the federal rules should be amended to specifically address electronic discovery * and a model state electronic discovery rule currently being advocated by some members of the defense bar.
Slide 57 	In the very few minutes that we have left, I'd like to turn our attention to the activities of the Advisory Committee on Civil Rules, which has a long-established Discovery Subcommittee. In October 1999, the Subcommittee turned its attention to electronic discovery. It did not have any pre-established goal or agenda. The idea was to explore the issue to see if rules changes were necessary or desirable. * The first set of activities were to hold two "mini-conferences" on electronic discovery, which I'll discuss momentarily. * Then they commissioned the Research Division of the FJC to do some surveys, which I reported on earlier *

	And finally, they have asked the Research Division to study a small number of cases in depth.
Slide 58 	As I said, the Subcommittee's first activity was to hold two "mini conferences," * one in San Francisco * and one in New York. * Invitations were extended to judges, attorneys, academics, and technologists. * The mini-conferences didn't have formal agendas or proposals to discuss. * They were more educational, information-gathering and sharing conferences, designed to help the subcommittee members understand the issues.
Slide 59 	Although there was a wide range of opinions presented, the consensus that came out of these two mini-conferences was that no immediate action on rules amendments was necessary or desirable. There were three reasons for this. * First, the December 1, 2000 amendments to the discovery rules had just been adopted. In fact, they hadn't even taken effect yet. And there was little enthusiasm for initiating yet another round of amendments. * Second, although again there were differences of opinion, the view particularly among the judges was that no changes were needed, and that individual judges had the ability under the current rules to deal with problems as they might arise. * Third, there is the reality of the amendment process, under which a rule change takes at least three years from inception to implementation. Since technology is moving

	so fast, any proposed new rule or amendment might be mooted before it is codified.
Slide 60  	Given the consensus that immediate action was unnecessary, the Discovery Subcommittee, with help from the Research Division of the Federal Judicial Center, undertook further investigation of the issue. They were looking for answers to two fundamental questions. * First, what, if any, aspects of electronic media discovery are unique and distinct from conventional discovery? * And second, if there are distinctions, should these be addressed in the rules of discovery?
Slide 61  	The FJC is continuing their research to assist the Discovery Subcommittee. About 20 cases have been selected for in-depth study. For these cases, the court filings related to discovery are being studies and analyzed. Interviews are being conducted with the attorneys and the judge in each case. A preliminary report was be presented to the Discovery Subcommittee in May, and we expect a final report at the Subcommittee's meeting in October.