JUDICIAL CONFERENCE
ADVISORY COMMITTEE ON THE FEDERAL RULES OF CIVIL PROCEDURE
CONFERENCE ON ELECTRONIC DISCOVERY

Fordham University School of Law
New York, New York
Friday, February 20, 2004

MORNING SESSION — 8:45 a.m.

DEAN TREANOR: My name is Bill Treanor. I'm the Dean of Fordham Law School. It's my pleasure to welcome you to the Civil Rules E-Discovery Conference.

As some of you may know, we at Fordham are getting ready to celebrate our centennial, and so I have been doing some reading into our history, and in particular the way the School got started.

Our first Dean, Paul Fuller, had a remarkable life story. He was orphaned at the age of three, grew up homeless on the streets of New York City, but despite the hardship and tragedy of his early years, became one of the great international lawyers of his generation, a senior partner at Coudert, represented President Wilson in his negotiations with Zapata, and he developed a profound and passionate commitment to the principle that well-constructed legal rules are the foundation of a just society.

It is because of that commitment that this Law School is here, and for a hundred years we have honored that commitment. We carry it forth in many ways. One of the most important ways is through the programs that are run under the auspices of our Philip Reed Chair, which was established in the name of civil justice reform. I would like to acknowledge Professor Capra, whom I will turn matters over to in a moment, who is our Reed Chair holder.

It is through the Reed Chair that we have invited judges and other experts from across the country to analyze the problems arising from electronic discovery and to explore whether a rules-based solution is required. We are also pleased that these proceedings will be published in the *Fordham Law Review.*

So I wish you well as you consider these important discovery issues.

Now I would like to introduce Professor Capra. Professor Capra, who has taught at Fordham since 1981, is one of the nation’s leading evidence scholars. He is the Reporter on the Judicial Conference Advisory Committee on the Federal Rules of Evidence and on the Judicial Conference Committee on Electronic Case Filing Rules. He is our Reed Chair holder. It is my pleasure to introduce Professor

Capra.

PROFESSOR CAPRA: I will not take up much time. This is such an august group that there is not much for me to say, other than thanks. I wanted to thank Judge Rosenthal for doing such a tremendous job, Judge Levi whose idea it was to hold this conference, Rick Marcus who put a lot of this stuff together, and Professor Myles Lynk who put all the agenda together and everything. It is a great job. I really didn't do very much other than to have the gracious assistance of Dean Treanor in providing the facilities. Helen Herman, thank you very much for doing all the groundwork for this.

That's all. I hope you all have a good time here, and if there is anything you need you should come and speak to me. Thanks.

JUDGE ROSENTHAL: Good morning. I am Lee Rosenthal. I am the Chair of the Civil Rules Committee. I am your hostess for today. I am very pleased that all of you are here. I want to thank in particular Dean Treanor and Dan Capra, who has been with his Chair just wonderful in their support of this conference, which made it all possible.

The purpose of these proceedings is to educate the

Committee members who are here, who are engaged in trying to determine whether electronic discovery requires an adjustment to the Discovery Rules or whether those Rules can accommodate these new media and new means of exchanging information. We don't know. We don't know how well the Rules are doing, we don't know how well the Rules will be able to continue to do, and we don't know if the problems that are present can be adjusted or addressed by changing the Rules. We very much hope that all of the experts that we have brought here can help us better understand those questions.

This conference as it is set up, and as I look around the room at the people who have very kindly attended, is an immediate source of frustration because we have such resources here that we cannot possibly mine them in the limited time we have available. So, above all, this conference is an invitation to each of you to continue to be engaged with us as we continue to examine these problems, which we suspect cannot be solved in a day and a half.

With that brief introduction, I want to turn the proceedings over to the Chair of the Conference, the Chair of the Subcommittee on Discovery for the Civil Rules Committee, Professor Myles Lynk, who with Professor Rick

Marcus has done the bulk of the work in putting the agenda materials together and in gathering all of the information that we needed to even begin this. Myles?

PROF. LYNK: Thank you, Judge Rosenthal.

While Professor Marcus and I were delighted to work together to put together much of the materials for this conference, credit must also be given to Judge Rosenthal whose careful hand both as an editor and an organizer and a final arbiter really made much of this conference possible.

Just before D Day, I have often heard the story that then-General Eisenhower went to visit the troopers from the 101st Airborne Division. He stopped before a young trooper and he said, “Young man, do you like jumping out of airplanes?” The trooper looked at the Supreme Allied Commander and he said, “No, sir, I don’t, but I like being around people who do.” Well, I do not know enough about electronic discovery, but I am going to enjoy being around people who do.

This program over the next two days consists of eight panels. Each panel will be moderated either by a Reporter for the Advisory Committee or a member of the Advisory Committee’s Subcommittee on Discovery. Each panel will discuss one of the important issues we have identified

in the area of electronic discovery as possibly appropriate for rule changes. At the conclusion of the panel presentations, we will open the floor for discussion and we very much look forward to a full discussion and participation from the audience. In fact, we hope that there will be a colloquy and interchange between the panelists and those of you who are attending this Conference over the next two days.

The last two panels will look at whether rule change is appropriate; and then, if rule change is appropriate, how such rule change should take place.

One of the interesting issues the Committee has to address is that the rule-making process itself is time-consuming, and so any rule that began to take shape by the end of this year would still be many years aborning before it was finally adopted. One of the issues we need to consider is that fact, that rule-making is a time-consuming process, how will that affect the rules that should be adopted.

Another issue the Committee faces whenever it looks at rule-making is whether or not the rule it adopts codifies existing best practices or whether it should adopt rules to define best practices beyond what is existing in

case law. That is sort of an interesting tension in the rule-making process and we hope to explore that tension today and tomorrow as well.

What that introduction, I would like to turn it over to Professor Richard Marcus, the distinguished Reporter for the Discovery Subcommittee of the Advisory Committee on Civil Rules. Professor Marcus.

PROF. MARCUS: Let me echo the thanks that the others have given to all of you who have given your time to be with us today. This is how we learn, particularly we in the academic biz, about what is really going on and what really matters, instead of the kinds of things that we sometimes become preoccupied talking about.

To talk about this, it occurred to me as I was considering what to say to you today, is sort of like something — it reminded me of a title of a very prominent book about changes in America that came out in 1930 [sic], called *Only Yesterday,* which some of you may recall was by a man named Frederick Lewis Allen,[1] which chronicled the changes that happened in this country for a variety of reasons during the 1920s. He found this made his career. He continued writing books like that, and he wrote another

book in 1950 [sic], called *The Big Change,*[2] describing changes in this country during the first half of the 20th century.

Well, it seems to me "only yesterday" is a theme that goes through my mind concerning the importance and pervasive influence perhaps of what we are here to discuss.

Only yesterday, we didn't worry about, or even know about, these things. Indeed, if you look through the compilation of cases at the back of your materials, you will see a 1997 decision involving Prudential, *In re Prudential Insurance Litigation,*[3] where it was a failure-to-preserve-records problem because the company foolishly thought it was sufficient to send out an email message to its thousands of agents across the country telling them to preserve records. That did not work because only 40 percent or so of them even had email and most of those did not know how to use it. Only yesterday, things were different.

As rule-makers, those who are charged with changing, adapting, improving the rules, one probably should be taking the long view. So whether what happened only since yesterday should be put into rules today is perhaps at

---

[1] FREDERICK LEWIS ALLEN, ONLY YESTERDAY:: AN INFORMAL HISTORY OF THE 1920S (Harper & Brothers, 1931).

[2] FREDERICK LEWIS ALLEN, THE BIG CHANGE (1952).

the heart of what we are talking about.

So what I want to talk about is the background mainly for that and then give you some observations about how we got to this point and what thoughts occur to me in a kind of spongy way ought to be in our minds as we go forward.

It seems to me using the "Big Change" as a theme, there are at least three big changes that form the backdrop for our discussion today.

• The first landed in the American litigation scheme in 1938 with the adoption of broad discovery in the Federal Rules of Civil Procedure. The last time this Committee held a discovery conference, Steve Subrin got up and told us about what a revolution that was. It was a remarkable change in the way in which litigation was handled, making this country unique in the world. Indeed, Steve Subrin has recently written an article on the same subject concerning our role or prominence in the world on this topic, entitled "Are We Nuts?"[4] So Big Change Number One is what happened before I believe any of us started being lawyers.

---

[3] In re Prudential Insurance Co. of Am. Sales Practice Litig., 169 F.R.D. 598 (D.N.J. 1997).

[4] Stephen N. Subrin, *Discovery in Global Perspective: Are We Nuts?*, 52 DEPAUL L. REV. 299 (2002).

• Big Change Number Two has to do with technology. If you think back to when Big Change Number One happened, technology was quite different, indeed not only in terms of the things that we deal with in today's and tomorrow's sessions, things like email, Blackberries, the Internet, even cell phones — none of that existed only yesterday — back when these Rules were written, there were not any laptops, word processing did not exist, they did not even have electric typewriters, and they used carbon paper to make copies because there were not any photocopy machines. All of those developments had an impact on litigation and discovery.

What they did not do in terms of Big Change Number One is have much of an impact on the Rules. Indeed, there does not seem to be a particular link between those changes and developments in the Rules.

Perhaps there is. Certainly it is not because Big Change Number One, the revolution in discovery, produced no concerns, controversy, or uneasiness. To the contrary, beginning in the 1970s, there was quite a lot of concern about the pervasive nature of possibly intrusive discovery. One began hearing frequently assertions that discovery costs too much, produces too little; that some parties seeking

discovery asked for far too much, they asked for everything. There were counter-assertions that those responding to discovery hid the ball, engaged in “dump truck” tactics. An overall criticism of the Rules was that there was no principle of restraint in the Rules; they simply invited as much as anyone wanted of discovery.

• Which brings us around really to Big Change Number Three, the rule-makers’ response to these other developments. Big Change Number Three started in the 1970s and went on for twenty-five years. It was in a sense a cycle perhaps — it’s hard to make predictions of this nature — a cycle that was completed with the Amendments that went into effect three years ago in December of 2000. It made a lot of changes. I will just mention some of them that seem to me pertinent to what we are talking about today.

It imposed numerical limitations on certain discovery devices; most recently, also a time limitation on depositions. It put a moratorium on formal discovery until the parties get together and discuss how they should handle discovery. It said that the court should order a time limit for the completion of discovery in almost every case. It addressed issues of misbehavior during depositions, in

withholding of allegedly privileged materials, instructions to witnesses not to answer in depositions. It put in place, and eventually in the National Rules without opt-out left in place, an initial disclosure provision that should advance discovery and advance proceedings in a number of places.

And most importantly perhaps, it added what is now in Rule 26(b)(2), what are called the proportionality provisions. In 1983, as that provision was first coming online, Arthur Miller stood before various groups and said that represented a 180-degree change in direction about discovery. Whether that 180-degree change in Rule direction resulted in a shift in the direction of the discovery ship is a question that one can debate, but certainly it seems that those provisions are receiving more attention, and in particular in relation to electronic discovery, than before.

So those are the big changes, the fifty-year-type changes that one might look at. The broad sweep of discovery expansion as a rule-making matter was followed by an era of discovery constraint — not abandonment, just constraint.

Well, in 1996 the last episode of that discovery-constraint undertaking led to the Discovery Project that the Advisory Committee began in that year. The way to try to do

that was like what we are doing here today: "We need to talk to lawyers and find out what is going on."

Frankly, for myself, I would say I had some expectations about what we would hear and heard many of those things. But there is one thing we heard about repeatedly that was not something that we were expecting to hear about, at least speaking for myself that I was expecting to hear about, and that is what we are here to discuss today.

Beginning in 1997, and frequently, lawyers would tell us, "This is the problem: you are talking about yesterday's problem. Today this is our biggest problem, dealing with these issues."

One reaction to what we were hearing was, "Haven't we just been dealing with that for twenty years? You say that there is an awful lot of material and people are asking for too much. You say that it costs too much to find all this material and it is not worth it. Very little of the material is actually useful or used in the case." Those are the kinds of things that were grounds for objection to discovery for years and years."

So one reaction one might have had, only yesterday, was, "This is not new ground; this is just new

technology. The Rules have not been changed particularly to deal with new technology." So the question is whether they should be.

Something that caught my eye in the last advance sheets for the Federal Rules Decisions seems to me worth quoting at this point. A thoughtful judge dealing now in the present with an electronic discovery problem, carefully examining it, said as follows: "It can be argued with some force that the Rule 26(b)(2) balancing factors are all that is needed to allow a court to reach a fair result when considering the scope of discovery of electronic records," and went on to say, "The options available are limited only by the court's own imagination."

So that is a backdrop for what has been going on as a more careful matter in the last three or four years.

Beginning in 2000, the Discovery Subcommittee launched a careful examination of the issues we are here to discuss today. In the year 2000, we had too many conferences involving many of the people who are here today who assisted us then in evaluating these issues.

One bottom-line reaction that came away from that activity was that there was anything but a pervasive and unanimous view that "something should be done now and here

is exactly what it is," on either one of those points — "something should be done now in terms of rule changes," it was not clear; and to the extent people felt something should be done, it was not clear what that something should be.

With the passage of that time, and subject always to what we learn today and what we learn hereafter, it does seem that maybe there is something special about this form of information, this form of discovery. Only yesterday it doesn't perhaps seem that way, but today and tomorrow it will seem that way. Let me offer some reasons for thinking so.

First, the volume is astounding. Maybe with electronic evaluation and search techniques that doesn't matter that much, but it dwarfs what we have seen before, even though the numbers we saw before were large numbers.

There are things, there are creatures, that did not exist before and do not quite fit our expectations or descriptions. At least some databases, dynamic databases, designed to be manipulated, designed to provide information on request, are not exactly documents. They do not look like they are exactly suited to treatment under Rule 34 for documents or Rule 33 for interrogatories. They are a

special new creature.

Retention and spoliation may feel and look different and be different in this new world.

Inadvertent destruction takes on new meaning when pushing the “on” button on a computer may cause materials that were available, accessible, findable, usable before to cease to exist.

Automatic removal of excessive, useless, old materials may seem a reasoned response and one that machines can implement without fail where document retention policies regarding hard-copy materials often existed on paper but not in operation.

So gradually, those are just some illustrations of things that may well be distinctive and warrant attention here. Gradually, it became apparent that one should be more serious. So recently there has been an effort to try to put in words what might be in Rules, and those thoughts — they are thoughts — are included in your material as they have developed in the minds of the Committee while it looks at these questions.

I want to move into the last set of observations I have, hoping to leave our panelists plenty of time on the first panel to talk about what they came here to discuss and

to give you plenty of time to ask questions or raise additional points that occur to you. Let me make some observations that one might have in mind while looking at these topics today, tomorrow, and afterwards.

First, this is not going to go away. Computers are here to stay for the long term. They are as a central feature of our lives going to be a central feature of our litigation. Just think in Manhattan of recent famous trials — Frank Quattrone, email messages used as critical evidence; Martha Stewart, email messages being in dispute in evidence.

We are regularly told that 90-some-odd percent of business information or human information or some kinds of information exist only in electronic and not in hard-copy forms. Well, if that is correct — although I have always wondered what their counting method was to come up with that number — if that is correct, then it would be remarkable for discovery somehow to overlook this mountain while focusing on the smidgeon that is in traditional hard-copy form. So it is here to stay.

Let me offer some other observations.

I spoke of what I called, I believe, Big Change Number Three, changes in the Rules to respond to criticisms

of discovery, and I think it is worthwhile to keep in mind that not everyone seemingly was entirely satisfied with the extent or nature of Big Change Number Three. So I would call this first observation the "unfinished business" observation.

Some of the criticisms and concerns that have been raised sort of look like the same thing again — "Well, we didn't get there on making the responding party immune to the cost of discovery before, and we ought to be trying to get there, and maybe electronic discovery is a vehicle for getting there."

You can make some very interesting arguments, persuasive in many ways in other parts of the world, shocking perhaps in many ways to litigants from other parts of the world who encounter our discovery, that, "Gee, anyone who wants to put someone to effort to respond to discovery ought to have to reimburse that person for the effort involved." That is not our starting point in this country. It is a possibility, but as a background matter I think it is important to keep in mind that the retreat from the broadest version of discovery was not an overall retreat from the decision of 1938.

Second observation: In some ways, perhaps newness

is the problem, unfamiliarity is the problem, and that is not necessarily, as I will reiterate, something that Rule changes can cure. After all, as I said, there are lots of technological changes that happened and the Rules were not changed to take account of those.

And, understandably, we have heard repeatedly that lawyers say, “We don’t know what to tell our clients. What exactly will judges do? What exactly do the clients have to do?” Well, a couple of reactions to that.

One, the new and unfamiliar and important, probably inevitably, has some aspects of that kind of uncertainty associated with it.

Two, related to that, if you look at Rules like the Civil Rules, I think it may be clear that often they focus on standards of reasonableness, and it is not clear that those sorts of Rules can provide the kind of certainty that would be a good thing if it could be provided but maybe will come only with experience.

A third, related observation is that a lot of the things that seem to be very important concerns are not related to court rules, the Federal Rules of Civil Procedure — for example, what I might call the “loose lips sink ships” phenomenon: people put things into email messages

that we certainly wish they had not. *Fortune* magazine recently had a story, for example, that said that email is "the corporate equivalent of DNA evidence, a legal albatross." *The Economist* a couple of years ago in commenting on some Wall Street activities said, "To put in a near-indestructible email the sorts of comments you might give vent to around the water cooler is to invite trouble." And so *Fortune* described some companies that were trying to school their employees in wise use of email. That is not a Rules problem. It may be a real problem, but it is not a Rules problem.

Second, again just a sense I get, that perhaps will not be reflected or agreed to by our first panel, and that is the orientation of lawyers and IT people is different. Now, that does not mean either is wrong, but if the reality is that IT people are the world's greatest packrats and they keep everything and are proud of it, then a different reality is that lawyers may find that a challenge to deal with. That is a background for rule-making and discovery decisions, but just a background.

Third observation: another problem that seems to exist from reported cases, some of them described in the materials we gave to you, maybe many of them familiar from

experience to some of you, is a communications problem. What do you do if you are the outside lawyer and you are supposed to tell the court what is available, what has been discarded, what can be done in discovery? You go and talk to the client, and the client says, “Here’s what we can do and here’s what we can’t do.”

There are a number of cases in which the “client” is somebody who does not really know but tells the lawyer what can and cannot be done, and the lawyer turns around and says, “Judge, here’s the situation.” Later on, during the deposition of the IT person, it turns out that the stuff that we told the judge was not available actually was and it isn’t anymore. That is one of the times trouble begins to brew.

What is the solution? Well, I would think on one level the solution is lawyers — and hopefully clients — have to learn about the problem and how to solve it. Perhaps there are rule-making solutions. One can find in the District of New Jersey’s local rule, and I believe in the one from Wyoming, that are in the materials rule-making reactions to that: “You must talk to your client and find out about these things before you establish a discovery plan.”

Next, a couple of observations on preservation, a big topic, less clearly so in the past.

On the one hand, preservation does not mean production. The fact that something is potentially available does not necessarily mean that it has to be obtained by moving mountains. But the flip side is if it is not available, then the decision that it would be worth obtaining it cannot be made at the time it should be made. That is one observation about preservation.

A flip-side observation is preservation above all would be a crippling thing to pursue. If you cannot turn on your computers because you might change something, then how is the world going to move forward?

Two more observations and then I will wind up so that we can turn to our panelists.

One is that judges are actually very smart people. This is not just currying favor; this is my mature view. And I am not going to necessarily say that is true also of law professors. They are practical, they understand generally what is going on, and often they understand what they do not understand. They have been learning, they are still learning, and they are telling others what they have learned. That is what the case law collection that is

described at the back of your materials shows.

So a question in the background is: have they been getting it wrong? Do we have reason to think that there are too few of them who have gotten far enough along the learning curve? Can we teach them through Rules what perhaps they do not really need Rules to know and to do? So is it wrong? Is the case law something we should change? Is it insufficient because not everyone has gotten the word?

That gets me around to something I believe Myles mentioned earlier, and that is changing Rules is hard. Some people think it happens too often. Practicing lawyers do not like it for good reason. It may seem a trap for the unwary. In my judgment, it is not true, as one of our panelists said in a recent article in *Litigation* magazine, that the Federal Rules of Civil Procedure are amended as often as the telephone book, but that was the opening line in an article in the most recent issue of *Litigation* magazine.

It is a cautious, time-consuming, difficult, challenging process. It is by nature all of those things. And just to emphasize what Myles said, if we got the clarion call from this group here and from the larger world and we could see what the goal was immediately, and immediately

begin to pursue it starting next Monday, then the soonest Rule changes dealing with these things could be in effect would be December 1, 2006, basically three more years from now, by which time somebody might be saying, “Only yesterday when we met in Fordham, we thought this was the problem, but now we see that there is this other problem, or that this solution will not be satisfactory.”

So I guess to wrap up what I am going to say and open the first panel with my introductions, the question might be phrased in terms of The Big Change. Do we need big changes to deal with what we are here to discuss? Can we figure out what those changes might be and be confident that they will produce the results that we want and not produce results that we deplore but discover only after the fact?

Surely this is a highly important issue. The Federal Judicial Center has been keeping track of CLE-type sessions on this subject and finds that they occur, and have for two or three years now occurred, at the rate or two or three per week, week in and week out, year-round. So surely there seems to be a groundswell of concern.

Whether that is a groundswell for making Rule changes is a different question. If one wants to know whether it is, at least somebody as technically challenged

as me has to find out from somebody who knows a lot more about these things. So our starting point — and I am shifting gears now — is to talk to some people who know a lot, probably more than most of the rest of you do, about these subjects.

PANEL ONE:

TECHNICAL ASPECTS OF DOCUMENT PRODUCTION AND E-DISCOVERY

Moderator
Professor Richard L. Marcus
University of California, Hastings College of Law
Civil Rules Committee

Panelists
Joan E. Feldman
Computer Forensics, Inc.

George J. Socha, Jr., Esq.
Socha Consulting, LLC

Kenneth J. Withers, Esq.
Federal Judicial Center

PROF. MARCUS: Panel One deals with technical aspects of e-discovery. I am going to introduce our heavy-hitter panel on this subject. We have three who are with us today. They have asked that you hold your questions until they are all done. We were going to have an illustration of technical difficulties, but they were removed supposedly, maybe as a transitory measure, but this magical item in front of me is supposed to assist them in making their presentations. I am going to introduce each of them and mention that my role here is as traffic cop, to impose time limits on each of them so there is time for you folks to ask questions or make comments.

First up will be Ken Withers, who is now with the

Federal Judicial Center, where he has been since 1999. He went to Northwestern Law School, spent about ten years in private practice in Boston, and then worked for the Social Law Library in Boston before joining the FJC. So far as I am able to determine from the résumés that I have, he is the only panelist on this panel who has received the Lord Lloyd of Kilgerran Award, which had to do with his graduate studies, I believe, in a related field in Great Britain. His topic is going to be, as I understand it, essential practical differences between electronic discovery and traditional hard-copy discovery.

Then we turn to Joan Feldman — and I should have said this of Ken; it is even more true of Joan — who many of you I suspect have heard before. Both of them regularly speak on these topics. She is the Founder of Computer Forensics, Inc. She has a background in records management and a background in litigation services, beginning, as I recall, as a paralegal. She must be one of the most sought-after speakers on this subject. Her résumé lists more than seventy such presentations over the last three years, so that's two or three per month. Her topic will be current practices in e-discovery, what is actually going on.

And then finally, to try to stare hopefully into

the future, George Socha will talk about what the future presently seems to hold. George is a graduate of Cornell Law School, spent fifteen years in practice in Minneapolis, increasingly addressing issues of electronic discovery in large-ticket litigation. Last year he started his own consulting firm on that topic. He is also a frequent contributor to events on electronic discovery. He will be our final speaker.

After those speakers are done, we will hopefully have substantial time — I think I am coming in ahead of time — for you folks to ask questions, and of course during the rest of the conference and after it is over, further suggestions/reactions are welcome.

I have completed my welcome to you. I would therefore thank you for your attention and turn the floor over to Ken Withers. Ken.

MR. WITHERS: Not only did Professor Marcus come in ahead of time, but he also said everything I wanted to say, so we will save a lot of time here.

My mission in the next ten minutes or so is to spell out the differences between conventional discovery of paper documents and the emerging world of electronic discovery, discovery of information that is created, stored,

or best manipulated and viewed using computers or computer media.

There are differences in degree and there are differences in kind. But first, and probably the most important, has already been alluded to by Professor Marcus, and that is a difference in degree that dwarfs all other, differences in degree and kind alike, and that is volume, the sheer volume of information.

Professor Marcus alluded to statistics from the University of California, his own school, which claimed that 92 percent of all information being created in the world today is created and stored in digital form on magnetic media — that is, on computers and disks and tapes. George Socha at the end is going to go into a little more detail on what that statistic really means. I simply want to demonstrate a few of the ways that this has occurred.

The fundamental difference between the way people create and communicate information on paper and on computers is that computer data is not tied to any artifact, like a piece of paper or a clay tablet. Computer data is digital, it's a sequence of zeroes and ones, positives and negatives, ons and offs, a stream of energy. When it is transmitted, there is no transmission of a physical object, like a piece

of paper, but of energy, which takes patterns from one medium and places them on another, like a computer hard drive or a disk. No physical object is moved.

This replication results in the buildup of massive volumes of data, mostly redundant but often containing subtle changes made by the people or the automated systems along the way. That is why one printed document that may surface in conventional discovery, if it is for instance a word processed document or the result of some other automated system, may represent hundreds of copies or versions to be found on computers and on network servers and on disks and on tapes.

The fact that data can be sent to the next cubicle or around the world, to one person or to a million people, with the same click of a mouse creates a buildup of data entirely unlike anything that we have seen in human history. But computers have created whole new categories of data that do not have easy comparisons in the paper world.

The first one that I want to mention briefly is metadata. Metadata is a made-up Greek word. Roughly translated, it is information about information. It is essential for the functioning of a computer. It is contained within each computer file. It tells the computer

such things as the file's creation date, the location of where it was created, how often it has been edited and on what other computers, the date and time it was last viewed or altered. Metadata is usually generated automatically, although it can be designed and manipulated by humans.

[Slide] It is not difficult to view. This is an example of a word processing document in Microsoft Word and what is called the "properties" window, which tells some of the metadata that is available.

[Slide] But computer files themselves may contain data which was never printed on paper, is never viewed on the screen. This is an example of the same word processing document showing the editorial changes that were made, what we call embedded edits.

When one looks at the data on a computer hard drive not through the lens of the operating system, which arranges it much like physical documents in a file cabinet, but through the lens of computer forensic software, we see a totally different world. We can see documents that have been supposedly deleted. References to that file, that document, have been removed from the visible operating system, but the data is still present and still intact on the hard drive.

[Slide] This is a quick example of a document, called Ericaperlvocalresume. The indication using this computer forensic software is that it has been deleted. However, the bottom half of the screen shows the document in its raw form complete with all of the formatting commands that would normally make this look like a very pretty résumé.

Because of the almost magical nature of digital data, that stream of energy, to be transmitted into any medium, we have many more places in which data relevant to discovery or an investigation can be found. And also because of the magical ability of digital data to transform itself in the process of attaching to these different media, we have any number of formats in which the data can be found, as though we have to conduct discovery simultaneously in a number of countries and in a number of languages.

While the volume of discovery increases on a macro level with the number of places, the number of formats, and the sheer numbers of documents that need to be looked at, it also increases on a micro level as each electronic file becomes in essence a little database unto itself.

The typical word processing or email file or other electronic file contains of course the visible data, the

things that one can see if the file is printed out or is shown on a screen. But below that there is another strata: the metadata that we have seen, the formatting commands, the formulas used to create the spreadsheets, the hidden and embedded edits that may be contained within that file.

And below that there is yet another strata: the bedrock on which that file rests, which is the hard drive or the medium itself, which may contain residual data from past files; it may contain what one of our speakers here, Dan Regard, in the past has called “digital packing peanuts,” which is data that is used to fill out the sector on a hard drive. These are the bedrock elements underneath any particular file.

If we are simply looking at paper or the electronic equivalents of paper, what we call PDF or TIFF images, all we see is the visible file. If we look at data in its native format, in the way that it is kept in the normal course of business and manipulated and used, then we see the second strata, the metadata, the formatting, the formulas, etc. If we take the step of going to on-site inspection of the computer media itself — the computers, the disks, the tapes — or we take what the forensic scientists call a bitstream image or a bit-by-bit copy of

the data, then we have the ability to look at the residual data.

Each way we view the computer file reveals a different layer, and this may go to the question of relevance. At what stage does this become irrelevant?

But documents themselves as tangible objects are actually disappearing. Today most commercial, governmental, and even personal communications and information are not reduced to immutable physical objects, like paper. When we conduct discovery, we are actually querying databases to generate selected data which we then arrange and present in a particular way. We are no longer looking at existing objects, like paper and file cabinets.

So the primary focus must be on the relevance of the questions being asked and the efficacy of the process being used to obtain the answers, not on the nature of the physical documents involved.

With that, I would like to turn it over to the next speaker. Joan?

MS. FELDMAN: Thank you, Ken, for laying out some of the primary technical issues — I always like following Ken because he carries the heavy load — and Rick as well for giving us that chronology.

As was pointed out, twenty-five years ago I started out in the paper cut brigade reviewing what was considered to be huge volumes of material, courtesy of another big technological change, the photocopy machine.

We'll fast-forward to twelve years ago, when I began trying to explain to people that a deleted file could be restored and that there were such things as embedded data and so on.

I believe that today we are in the middle of the next revolution in electronic discovery, and it concerns the overwhelming volume of material that we are facing. There has been a lot of focus on this issue, and for good reason.

[Slide] I like to term it "the tsunami effect." I also know that George when he follows me today will be talking about the fact that this current problem is actually only going to grow and continue to grow.

I would like to talk to you today about what people are doing in the real world to deal with electronic discovery. To that end, again, I want to encourage you to have the mindset that there is an enormous amount of material out there, that it is often difficult to identify where the real value is going to be in going through that material. That is not an idle subject because there is so

much information out there and there is so little in a way that actually turns out to be truly responsive.

There has been a big push in applying technology to solve this problem. There have been some amazing developments in tools for sifting through the electronic documents, for acquiring it more easily, for doing text searching and concept searching. All of these are most helpful because we are trying to deal with a tsunami.

I would just like to tell you that in many ways it is like having a snorkel when you are out in the ocean, that it is a good tool, but we are dealing with a huge volume of material and it is growing.

[Slide] Let me put this in context for you in a real-world example. We were recently called upon to be a mediator in a case involving a large Fortune 100 company. The special magistrate was at a stalemate with both parties. At issue was a huge volume of material. A well-respected large litigation support company, a well-respected large law firm, had assisted the Fortune 100 company in identifying the documents to preserve and produce. They came up with a total volume of 42,000 backup tapes — that's another issue — and they identified twenty hard drives. For the judges sitting here today, I think that you might be familiar with

hearing this type of number brought before you.

What was at issue? Plaintiffs dug their heels in and insisted that 42,000 backup tapes be restored and reviewed. Defendant producing the documents said, “It’s expensive, it’s a fishing expedition, it’s not going to yield anything.”

Between the two of them, they probably spent over $75,000 just on motion practice, and the lucky judge got to hear the debates about what a tape was, what was on the hard drives, embedded data. At the end of the day, nothing really had been produced, nothing really had been reviewed.

“Break the deadlock,” that was our charge. As expert witnesses often do, we gently guided the court, and although our mandate was to actually see if it was really worth the money to look at 42,000 backup tapes, we suggested something else. We suggested that they begin focusing on where the evidence might actually be. “Well, we have 42,000 backup tapes and we have twenty hard drives. How much is it going to cost?”

We gently suggested that they needed to focus on where the evidence might actually be. We applied some techniques that the attorneys sitting here today are familiar with. We questioned witnesses. We went from the

point of departure as to where the evidence that they were looking for might actually be stored, in whose hands; what was the evidence — it was a trade secret theft case; who worked on the documents that might actually have something to do with that; who worked on the product at question.

In one and a half days of interviews with the systems people and with some of the key witnesses who had actually created some of the documents that we felt would be at issue, we actually located another server that had not been disclosed that had been set up by the engineers, as they often do. We like to refer to engineers as our “rogue” folks because they often set up their own systems. They had established their own system, including their own email server. It’s like Mount Everest — you know, it’s there.

We located this fact that there had been a server. By the way, in the course of the discovery in a six-month period, they had disabled and mothballed the server. Actually they destroyed the server because they wanted to use it for some other application. They had not been informed by the attorneys or didn’t pay attention to it. But, as engineers will often do, they were also packrats and they had created two backup tapes for that server. That is actually where the evidence was. It had been previously

unidentified.

What about the 42,000 backup tapes, the subject of much fevered debate about cost and cost sharing? Were they impenetrable? Was it this monolithic dataset that was going to cost at a conservative estimate $4-to-$5 million? Through questioning of the IT staff, we were able to find the Rosetta Stone that helped us begin to prise apart that monolithic dataset to identify particular tapes that might actually contain evidence.

Through a closer look at some of those tapes as we began this process, we were able to narrow the 42,000 set to thirty-seven tapes — thirty-seven backup tapes, previously undisclosed data — not as a result of some technological marvel or breakthrough in text-searching technology. Despite the fixation with blue screens as solutions for electronic discovery issues, I would just suggest to you that a good background in technology, an understanding of how enterprises use their computers, and the same principles that guide experienced litigation attorneys and jurists in their decision-making process, in terms of finding and refining and looking for responsive information, is critical here.

There is a dynamic tension in my field these days

because there is such an emphasis on the ability to process massive amounts of data. That is fine. I am not saying that there — thirty-seven backup tapes was a lot of data; it was good to have those tools — but the volume is increasing and we do not necessarily see a corresponding interest in just understanding some of these basic fundamentals.

So that is one issue that we are dealing with and, conversely, all of you are dealing with.

[Slide] There are a few ways to begin chipping away at these issues.

- You must start at the preservation phase because you are going to have to make some decisions about what needs to be preserved; and if you do not and you are continuing to overwrite tapes or reuse or format hard drives, you are going to destroy critical information. So you have to start there.

- You have to learn how to distinguish what kind of data you are looking for. Are you looking for Word documents? Are you looking for email? Are you looking for database types? This question needs to be answered early on.

- Data elements. Ken does a masterful job of explaining things like metadata and embedded data. These are data elements that you may be concerned with. Or you may not; the parties may make a decision that they are not, they don't care, they just want what is on the face of the documents or that compilation. That's fine, but those decisions have to be made.

- Common terminology needs to be developed between the parties. We suggest the adoption of a glossary of terms and that they agree to it, so that you do not have this shifting target as you move through as to what is a database, what is a relational database, what is a file. You need some basic terminology that you agree upon.

- And you also have to make some decisions even at the earliest stages as to how you are going to produce that information. Mention was made of a TIFF image versus a native file. There is a big distinction there. TIFF images do not contain embedded information; they do not contain the original metadata. When you are producing those documents you need to have some idea of what it is you are going to be producing to each other and, unfortunately for everybody, you have to make those decisions early on.

[Taping malfunction — short gap — need “hog farm” quote at end and beginning of Mr. Socha’s presentation.]

MR. SOCHA: Now is where […]

Next is the question of more data. The volume of data is expanding rapidly.

[Slide] Here is a little bit more detail from the 2003 study done by the University of California at Berkeley. That followed up on a 2001 study, I believe, where the authors made at that time what they considered to be outrageous projections as to the growth in the volume of material there in electronic form. In the executive summary to the 2003 report, the authors said: “We had no idea. What we thought was outrageous didn’t even come close to what appears to have happened.”

And, importantly for this discussion, is the row on magnetic. Now, they are talking here about 4 million terabytes of data. That is a volume that I think none of us can even begin to conceive of. There is nothing like that in paper out there. So we have got this enormous volume of information that we potentially need to deal with. If we keep trying to buy hog farms instead of just ham sandwiches, we are going to be in a lot of trouble.

[Slide] There are also more types of data and in more places than I think many people really recognize.

- Of course we've got email. There is a lot of discussion about that. That is what captures people's attention. That's the easy picking, though; that's the low-hanging fruit. Email is almost like paper in many ways. You can pull it up on the screen, most people now are used to dealing with it, and you can read what is right there.

- Instant messages, though, it is predicted will be equal in volume to email within a couple of years, perhaps sooner. That is a much more difficult medium to deal with for electronic discovery purposes.

- Text messages, such as the ones sent back and forth by cell phone users, are rapidly growing in use.

- Relational databases, while they have been around for a while, have not for the most part been the subject of discovery requests. I think lawyers have avoided going there because relational databases are simply too esoteric, too complicated, too confusing for most people who have not had to deal with them in some other aspect. If I were to bring up on the screen — and I will not do this to you — the plan for a relatively simple relational database,

what you would see would be lots of boxes over the screen. It's like the anthropology class I had as a freshman in college. The professor put up boxes on different subjects and then started drawing lines to each other. By the time he was done, there were about thirty boxes on the wall and lines from everything to everything.

If you look at that on the screen as a user would, you will see something that is coherent and makes sense, provided they built the relational database property. If you try to go in without knowing what that database is and without the benefit of that user's experience and expertise, you might find yourself just with gobbledygook. But that is where a huge amount of data is stored these days.

- XML datasets. The word processing document we see today is nothing like what you think it is. It may look like something that just gets printed out on a sheet of paper. There is not just metadata there, though. It may be broken up into all sorts of constituent parts that are not even part of that file but elsewhere. So the information is all over the place.

- Digital photos.

[Slide] And then, with expansion also comes better processes and tools, some of the stuff Joan was

talking about. People are learning how to do this better and how to move forward with it. Well, with expansion then follows routinization. We get used to this stuff. It becomes part of what we are doing. Bigger projects can be done than ever before.

In 1996, I handled what was probably one of the largest tape cases that year, with 461 backup tapes. I had a hard time finding any vendor who could handle that work. The largest backup tape case I know of from last year involved 10,000 tapes. Now, most vendors cannot handle that, but there are some who can.

The amount of data we are dealing with in the 1990s — in the late 1980s we were talking about kilobytes; in 1996 10 gigabytes was huge; now 10 terabytes are not unusual. We have to figure out how to deal with that volume because it is only going to get larger.

[Slide] But with routinization also the impossible becomes possible. We discover that we can do things now that we simply could not handle a few years ago. There are vendors out there offering services that were unimaginable six or seven years ago. The data that was essentially inaccessible not long ago is routinely available now, and that is only going to continue to change and be so

moving forward.

And then, finally, with this routinization our expectation level goes up. Because we can do this much, we want to be able to do this much. As we demand more, the people who are providing the services and capabilities turn around and, so far continuously, have been able to offer more to us, which then takes us right back around to expansion.

So looking into the future as best one can at this point, there is an enormous growth in the volume of information we have to deal with, a growth so far beyond what we are capable of doing that we cannot even really begin to imagine in some ways now how we are going to be handling this information in a few years, except to know that most of the issues we are dealing with right now are going to at a very technical and detailed level be yesterday's news, at best.

Thank you.

PROF. MARCUS: Thank you.

This is the time when we hope to be receiving the benefit and insight of the presence of all of you. There are obviously questions one could ask these panelists. If no one else wants to, I will.

We have a question from Jim Rooks.

QUESTION [James E. Rooks, Jr., Center for Constitutional Litigation, Association of Trial Lawyers of America]: For Joan Feldman. You described that representative case. Can you fill in a couple of additional details? You said one of the parties was a Fortune 100 company. Generically speaking, what other kinds of litigants were involved?

PROF. CAPRA: Would you identify yourself when you ask a question?

PROF. MARCUS: We are trying to keep track of who is participating, so if you can say into the mike who you are, then we can keep track of that.

QUESTIONER [Mr. Rooks]: Excellent idea. I'm Jim Rooks for the Association of Trial Lawyers of America.

In the representative case you described, you said at least one of the litigants was a Fortune 100 company. I'm wondering if you could describe generically the other litigant or litigants.

Next, were you brought in as a court-appointed expert? I'm assuming you charged a fee for your services. Was it split among litigants or paid by one side?

MS. FELDMAN: I'll answer the first question. The

other party was a smaller company alleging that the larger company had appropriated their trade secrets, so they were not of the same size as the larger company.

Your second question was the —

QUESTIONER [Mr. Rooks]: Were you court-appointed?

MS. FELDMAN: Yes, I was a court-appointed expert. Actually I was asked to assist on assessing the cost issue, and that's how I was brought in, because the parties were at a deadlock.

QUESTIONER [Mr. Rooks]: Were your fees split among the parties?

MS. FELDMAN: My fees were split between the parties, yes.

QUESTIONER [Mr. Rooks]: By agreement or by court order?

MS. FELDMAN: By court order.

PROF. MARCUS: Down here.

QUESTION [Stephen D. Susman, Esq., Susman Godfrey]: Steve Susman, Houston. My question for the panel is whether anyone is aware of any studies or surveys that have been done of lawyers who have retained these forensic consulting firms to see whether they really believe anything that they discovered was outcome-determinative. I mean, was

it worth doing, or is it just a huge waste of money? And can you cite me to such surveys?

MR. SOCHA: No such surveys that I know of, so there is nowhere I can cite to. The best I think any of the three of us can do is provide anecdotal information.

I would say that in my experience the electronic discovery activities tend to be very costly. Much of the work is fruitless, which is the case with any discovery activities. I mean, isn't it true? But there have been certainly pieces of information that have come out of electronic discovery that have been critical to the outcome of the cases.

MS. FELDMAN: Mr. Susman, I would say that in approximately 45 percent of the cases that we have worked on the evidence that was disclosed turned out to be critical to the case. I do believe that the focus needs to be narrowed, and it has been our role in guiding our clients to try to help them reduce the scope of their review.

QUESTIONER [Mr. Susman]: Of the lawyers here, is there anyone who can give a testimonial to the fact, or give us an example of how something discovered electronically that couldn't have been discovered in hard copy made a difference?

PROF. MARCUS: Is that responsive to his question?

PARTICIPANT [Richard Seymour, Lieff Cabraser Heimann & Bernstein]: Yes.

PROF. MARCUS: Okay. You are going to be next, but maybe if we've got a response to the question? We also only have one microphone. Sorry.

PARTICIPANT [Mr. Seymour]: My name is Richard Seymour. I handle employment discrimination class actions. I have been discovering electronic databases since 1971. I would have lost virtually all of my cases without it because everyone always denies discrimination. It's the analysis of what took place inside the company that makes the difference. In some instances, we have had to create a database that contains all the job movements of all the employees over a period of a decade and a half in order to be able to do that. But our clients would have lost their cases without that. It would kill us if we didn't have it.

I have to say that there are very few problems with it. Sometimes you have people being unreasonable, the same way that you have with paper discovery, but I have never yet come across something that reasonable people could not sit down and resolve in a very short period of time.

PROF. MARCUS: Thank you. The next questioner is

unfortunately on the other side. There may be a chance — I have no idea of our technical capacities — to have a second microphone.

PROF. CAPRA: We have stereo now.

PROF. MARCUS: Ah, you have two mikes. I was going to suggest that. Go.

QUESTION [James. L. Michalowicz, Tyco International (US), Inc.]: Jim Michalowicz from Tyco, formerly of du Pont.

A question for, I guess, the whole panel. I think the root cause of a lot of the subjects we are going to talk about — spoliation, cost shifting — the root cause seems to be these backup tapes. My question is: do you see any change in terms of technology on the backup tape front, or also on the thinking in terms of the retention of backup tapes, that may go ahead and reduce some of that root cause?

MR. WITHERS: I'll take that first. The short answer is that backup tapes are a technology that will be disappearing quickly, I think, but that is driven mainly by business reasons and not necessarily by discovery costs and risks, although it is somewhat.

The purpose of a backup tape — and we must make

this distinction between backup tapes which are used to capture all of the data on servers and systems for the purpose of disaster recovery — if the system is struck by lightning the next day, then all of the data can be immediately restored, but it is restored in an unstructured way. Once the data is put back onto the computer system or the media, you must have the original operating system and the original configuration to bring it back to life, to make it usable. So backup tapes are not archival media, they are not used, or they should not be used, for the retention of data that needs to be accessed in the future.

The problem has been digital packrat'ism, and that has been alluded to before. People keep backup tapes long after their logical life, which should be twenty-four hours or until the next backup tape is made. There is no reason to have more than one backup tape for any particular system.

Now, the problem has been that people have been keeping these backup tapes. And indeed, if you ask any good IT person in any company, "Have you ever been asked by the executive suite to recover a lost email from a backup tape?" the answer will be "yes." People have been using them for that purpose. They are not really very good. They are very costly.

What is happening is that that technology is being replaced by parallel processing, by having dual systems operating, often on different continents, simultaneously so that this reduces risk tremendously. If one system goes down, the other system is up. That is the reason why on September 12, 2001, the financial markets in the world continued to operate out of Charlotte, North Carolina, because that is where a lot of the backups were.

So backup tapes as such will probably be disappearing. We hope so. But it will be driven by business purposes.

PROF. CAPRA: Rick, we've got somebody over here.

QUESTION [Chris A. Seeger, Esq., Seeger Weiss]: Hi. Chris Seeger, Seeger Weiss. I think one of the reasons we hear — and this was a question that was asked by Steve Susman — is "why don't hard copies suffice?" I would like to hear you guys address what actually never makes its way into printed form so that hard copies wouldn't suffice? I have heard reports anywhere from only 10 to 40 percent of information that is created by companies ever finds its way onto paper.

MS. FELDMAN: I'd be happy to answer that question for you. Who here has used an Excel spreadsheet? When you

print out the spreadsheets, are your formulas printed out? No. The formulas are embedded in the field. So if you give me a printout, I will not have the complete document.

Probably the other issue that gets a lot of attention is what is referred to as metadata. You can find metadata in any of your documents by opening the "properties" feature. It will show you the date of creation and give you some statistics. Again, not normally printed out, only available in that electronic format, in that native format. Sometimes it is captured during a conversion process to TIFF, but not always.

QUESTIONER [Mr. Seeger]: So if you do not do electronic discovery, that is information you would never get?

MS. FELDMAN: If you do not do your electronic discovery to get that version of those documents, that embedded information is not available to you.

MR. SOCHA: The other piece is a relational or other database. If you just try to print out a database, you will most likely get many, many boxes of garbage and nothing else.

QUESTION [Michael Arkfeld, Assistant U.S. Attorney, Arizona]: Thank you. My name is Michael Arkfeld.

I am an Assistant U.S. Attorney from the District of Arizona.

The question I have for the Rules Committee, and for the panelists especially, is: presently or in the future, with the increase in costs of electronic discovery, will there be a corresponding decrease in the cost of paper production?

MS. FELDMAN: [Nods affirmatively.]

MR. WITHERS: My view is no because people insist on printing out all of their email. I do not know why they do it, but we have many, many cases around the country where people are routinely converting large amounts of digital information into paper for processing. It is becoming astronomical.

As we have mentioned before, things like metadata or the formulas used in spreadsheets — the reaction of many attorneys when they realize that they are missing this information is to ask for it in paper form.

QUESTION [Michael R. Nelson, Esq., Nelson Levine de Luca & Horst]: Mike Nelson, the Philadelphia law firm of Nelson Levine. I defend insurance companies against class action lawsuits.

A quick question for you, Joan, and then a

question for the panel. That cost that was split, was it split 50/50?

MS. FELDMAN: Yes, the split was 50/50 in that case.

QUESTIONER [Mr. Nelson]: As I watched the presentation this morning, I started to think that a lot of what we are going to get into at this conference is about this concept of what is reasonably producible. It sounds like as technology catches up with us that that concept is more or less going to become outdated because everything is going to be "reasonably accessible." As litigation goes on, a lot of these backup tapes are going to be more or less made available anyway as part of this. So as you look at the ever-expanding universe of what is out there, does it make sense to stick to tenets that are in some of the opinions we have now that says "if it is reasonably accessible, it should be produced," because then it gets into a much broader universe as time goes on?

MS. FELDMAN: I think that a definition of "reasonable" is what is critical here. To use a contemporary analogy, you can have a satellite view of what you are going after, which is global and all-encompassing, or you can use on-the-ground intelligence to begin to

identify what actually may be real and valuable.

When you are looking for electronic documents that are responsive to your case, you can of course deal with the universe of those documents in virtual repositories, on backup tapes, in databases, throughout an enterprise, but understanding how people create that information and how they use it, the people that are particular to the issue, and then extrapolate from there I think is the only way people are going to get out of this in terms of what is reasonable, where is your trajectory when you are looking for responsive information, when you are dealing with this huge volume of material that is out there. So I do not think it is purely a technical issue.

MR. WITHERS: I would like to amplify that a little bit. It is indeed true that as the technologies related to search engines and to discovery improve, then more and more things become accessible. The bottom line is that if something is in digital form and it is online, then it is possible to search for relevant names, dates, and relationships with other information, and it would be possible to take a large corporation in the future and place a search engine over top of its entire enterprise-wide data collection and search for everything.

The question is not accessibility; that is a threshold question. The question goes back to Rule 26(b)(1). Accessibility is a 26(b)(2) issue. But is it relevant and do we want to concentrate more on the process by which those searches are conducted and what they are going after?

Then we have the cost beyond that. Once we have gotten down to the relevant information, it still has to be reviewed by attorneys and used in some way or another. That is where the real costs come in, and no technology is going to replace hapless first-year associates.

PROF. MARCUS: Just as a housekeeping matter, I've got five folks who have indicated an interest and then we may be running out of time after that. I'm not sure. We've got a sixth and a seventh.

MR. SOCHA: I was going to say one quick coda on that. You may have to use technology to replace first-year associates if you truly do have to deal with the volume of data out there. You cannot put enough bodies in front of enough boxes or enough screens to in a reasonable amount of time review 10 terabytes of data. It is not possible.

PROF. MARCUS: I've been told we really ought to cut it at five. There will be plenty of chance for others

to participate. We've got five, I think. Anyway, you are next.

QUESTION [Charles A. Beach, Esq., Exxon Mobil Corp.]: Chuck Beach, Exxon Mobil, Irving, Texas.

Let's assume that a lot of critical information does come out of emails. Have there been any studies that show whether that is email that is accessible in the normal course of business or do these critical emails come off of backup tapes?

MR. SOCHA: I know of none.

MR. WITHERS: I know of none either.

PROF. MARCUS: Okay.

QUESTION [Francis J. Burke, Esq., Steptoe & Johnson]: Frank Burke from Steptoe & Johnson in Phoenix, Arizona.

My question for the panel relates to your advice on taming the tsunami. My take-away from Joan's remarks was that perhaps the greatest computer you have access to is the one in your head, and perhaps the way to start is by looking at the corporate relationships and who was working on the data. She took 42,000 tapes down to two or thirty-seven. Judge Scheindlin, I guess, taught us that maybe the first step is sampling, start with a sample and then go from

there. I just wondered if the panel can give us some other ideas. When you have 42,000 tapes, how do you tame the tsunami?

MS. FELDMAN: I think sampling is a great idea. There are different ways of sampling. We use a little different terminology, and that is actually trying to get the catalogues off the tapes so you can do a more effective sampling. But I think you have to use every tool available to you. You have to use your head, you have to use technology, you have to use sampling. I don't think there is one solution. I think the danger right now is focusing only on the technical as a solution. That's my main point. So I respect the different ways of attacking this problem.

PROF. MARCUS: Okay. We've got a Committee member and then we've got three more and then I think we have to take our break.

QUESTION [Hon. Brent H. McKnight, U.S. District Judge, North Carolina (Western), Civil Rules Committee]: Brent McKnight, Western District of North Carolina, Charlotte.

If lawyers come to me early on in a case with claims of spoliation in electronic discovery contexts and ask me to issue a preservation order or have an early

spoliation discovery conference of some sort, knowing that, at least under the current version of the rules, you need some kind of order unless you are relying on the inherent authority of the court to enforce sanctions, I guess my question is: what could I anticipate that technology can tell us about whether and to what extent there has been spoliation to make such an order something that I should issue?

MR. WITHERS: This is a very fact-based question. It depends on the nature of the spoliation that has been alleged. But certainly through computer forensics technology we can determine for the most part whether files have been deleted, whether scrubbing mechanisms have been used. We have several case examples, the *Kucala* case and such, where technologies were used to attempt to destroy evidence.

But you have to have more to start than simply an allegation that spoliation may have occurred because it is theoretically possible to destroy documents. You need something to go on to determine what is going to be the appropriate way to approach this. My view is that early in the case a question of spoliation arising had better be grounded in some evidence outside the speculation that it is

possible to happen before an order can arise.

PROF. MARCUS: I think this is proving that this is a format that works, but there is limited time on the first panel. I think I am going to have to cut it. I thought John Carroll was next — well, golly, I thought that's where next — and then the gentleman on the aisle, then Laura Ellsworth I think were the three that I had noticed. I apologize to those I overlooked. My colleagues will do better.

QUESTION [Hon. John L. Carroll, Dean and Professor, Cumberland School of Law]: John Carroll, Cumberland School of Law, Stanford University.

We are talking about a Rules process that takes three-to-five years to change. You are talking about technology that is happening overnight. What is your view on whether technology has outstripped the Rules process and that any Rule change that might be implemented is obsolete well before it becomes law?

MR. WITHERS: Any Rule that attempts to address a particular technology, such as email or instant messaging or backup tapes, is doomed. It has to be very general, and it really should concentrate on the discoverability of relevant information in whatever format, using whatever technology.

QUESTION [John Vail, Esq., Center for Constitutional Litigation]: This is a specific question for George Socha. This is John Vail from the Center for Constitutional Litigation.

George, you put up a graph of costs. Can you describe the data sources for that graph? It sounded when you were talking that the graph was actually a graph of money spent on consulting businesses. You tell me.

MR. SOCHA: Very quickly, we did a survey to find out a number of things about electronic discovery, but in particular for the costs. We conducted telephone interviews of, if I remember correctly, fifteen or sixteen of the better-known service providers, of maybe nine or ten of the law firms that we knew were most actively involved in electronic discovery activities, and similarly of folks at five or six of the corporate legal departments. We then gathered other data from various third sources. I forget — I ended up with 15 gigabytes of data or something like that, pulled this all together.

We were asking for these costs what people were spending or thought they would be spending for electronic discovery activities that ran the spectrum from the initial preservation and collection of information through the

processing and production of that information to the other side.

PROF. MARCUS: This is going to be our last question of this session. There will be questions in every session, there is time between sessions. We want to hear from everybody who has got something you want to communicate. But for this session we have one more comment and then we will be taking our break.

QUESTION [Laura Ellsworth, Esq., Jones Day]: I'm Laura Ellsworth. I'm with Jones Day.

Most of what you all talked about had to do with production obligations, what we do to search and produce from a known body of data and information. My question goes to the preservation obligation, which I think is a thornier problem and one which is not as susceptible to resolution by Rule 26(b)(2), as many of these issues of what do you do once you have grabbed everything and you're figuring out what you have to turn over.

My question is this: what technological solutions are you aware of that permit us to deal with the problem of how you preserve all the information for purposes of 26(b)(2) analysis and how do you deal with what Professor Marcus identified as you can't cripple the company? But the

thornier issues are if you don't preserve it, you are into the spoliation issues, and that is the larger area of uncertainty for practitioners, I think.

MS. FELDMAN: I think before you can begin effective preservation you have to attend to the identification stage in terms of trying to locate what may be responsive, Laura. So there you might be a bit more broader based.

In terms of the technical ways of preserving it, if you have identified a particular server out of fifty or 400 servers that might be key, you can make a "snapshot" backup, you can make selections of what you are going to do, you can store that information on tape. That doesn't mean that is how you are going to produce it.

In terms of hard drives, you can preserve hard drives forensically for selected drives. But again, it is difficult to preserve documents until you know what you need.

Let me give you the analogy of the warehouse filled with paper that most large companies have. They have destruction schedules. They know to instruct the warehouse to not move the boxes off the shelves for destruction even though it's their due date, for example.

So there are methods that people have developed in corporations over the years. They can apply these methods using good data stewardship to their electronic documents, and many companies are starting to embrace this so that they can more quickly identify, and therefore preserve, responsive information.

PROF. MARCUS: Okay. I am going to have to stop the first session, and only the first of eight, there. We thank our panelists. Thank you.

[Break: 10:40 to 11:00 a.m.]

PANEL TWO: RULES 33 AND 34 — DEFINING E-DOCUMENTS AND THE FORM OF PRODUCTION

Moderator
Hon. Shira Ann Scheindlin
United States District Judge,
Southern District of New York

Panelists
David R. Buchanan, Esq.
Seeger Weiss LLP

Adam I. Cohen, Esq.
Weil, Gotshal & Manges LLP

Hon. James C. Francis IV
United States Magistrate Judge,
Southern District of New York

Paul M. Robertson, Esq.
Bingham McCutchen LLP

JUDGE ROSENTHAL: The next panel is moderated by Judge Scheindlin. In this group Judge Scheindlin needs absolutely no introduction.

JUDGE SCHEINDLIN: Thank you, Lee, for that wonderful introduction.

As the Moderator of this panel, I am going to keep an uncharacteristically low profile. I intend to do nothing but introduce the speakers and the topic. I also will use demonstrative evidence. This is the only piece of demonstrative evidence I am going to use. I am going to cite you to the pages that you might want to look at as we

cover certain of our assigned topics. So if you have it with you, it may be time to make it available.

I am going to begin by introducing the panelists. Everybody on our panel has a title.

Our judge — Judge Francis has been a Magistrate Judge in the Southern District of New York since 1985. He is sitting on my right. A graduate of Yale College and Yale Law School, he clerked for Judge Robert Carter in the Southern District of New York and then joined the Civil Appeals and Law Reform Unit of the Legal Aid Society. He is currently an Adjunct Professor at this very Law School where he teaches constitutional torts. Judge Francis is the author of another case that needs no introduction, *Rowe Entertainment v. William Morris Agency,*[5] the leading case on cost-shifting in the context of e-discovery.

Our author — our author, Adam Cohen, is the author of a leading treatise in this field. He is on my far right. He is the co-author of *Electronic Discovery: Law and Practice,*[6] which recently came out, fall of 2003. He is a partner in the Litigation Department of Weil, Gotshal &

[5] Rowe Entertainment, Inc., et al. v. The William Morris Agency, Inc., et al., 51 Fed. R. Serv. 3d 1106; *aff'd.* 53 Fed. R. Serv. 3d 296 (S.D.N.Y. 2002).

[6] ADAM I. COHEN & DAVID J. LENDER, ELECTRONIC DISCOVERY: LAW & PRACTICE (Aspen Publishers 2003).

Manges. His practice areas include intellectual property and commercial matters for clients in the technology, media, and entertainment industries, with a focus on Internet- and computer-related issues.

On my left, David Buchanan is a partner in Seeger Weiss LLP, specializing in representing individual and corporate plaintiffs in complex litigation, including securities, consumer fraud, pharmaceutical torts, products liability, and pension claims. He is currently involved in some of the largest MDLs pending around town, including *In re Rezulin Product Liability Litigation*, *In re IPO Securities Litigation, In re Delta ERISA Litigation,* among others. Before joining Seeger Weiss, he represented defendants while associated with Fried Frank, a large New York firm representing primarily corporate clients.

On my far left, probably inappropriately far left, is Paul Robertson, who is a Litigation Partner at Bingham McCutchen. He represents clients in many practice areas, including bankruptcy, construction, mergers and acquisitions, directors' and officers' indemnification, products liability, general commercial disputes. Paul is a member of the Defense Research Institute; the Sedona Conference Working Group, and in that capacity he

participated in drafting the recently issued “Best Practices on Electronic Discovery.”

Now I am going to tell you very quickly what our topic is and what is the format that we hope to follow.

We are going to begin roughly at pages 5 through 7 of Exhibit A, which I previously showed you. The topic there is to briefly discuss the definition of e-data, and we are going to do that very briefly; that’s a three-minute segment of this show.

We will then turn to the question of whether Rule 34 needs revision in order to refer to “data” or “information” rather than “documents,” which as you just heard may be a passé concept in the 21st century. Listening to the last panel, I must say that relational databases and formulas for spreadsheets do not entirely sound like “documents.” In any event, that will be found at pages 14 through 15 of your Exhibit A.

Questions that we will cover in that segment will include such things as: In producing data stored on electronic media, should that production include all data stored or maintained as part of the electronic record? — just to whet your appetite.

Our next topic, our third of four, will be the

form-of-production question. The question there is: should Rule 34 require the requesting party to specify a particular form for producing the requested data; and should the Rule also talk about the grounds on which a producing party might object, such as inaccessibility? That will be found at pages 16 through 20 of your Exhibit.

Finally, we will briefly turn to whether Rule 33 needs to be amended to specify that interrogatories may be answered by the production of electronic data; and, if so, what responsibility might the producing party have to produce that data in some way that is actually usable?

Now, on each topic we have decided to go in this order: our author will go first, Adam will try to give us the very briefest of backgrounds; David and Paul are set up a little bit to be sparring partners, a little bit of plaintiff/defense viewpoints, will then go next; and, as is always appropriate, the judge will get the last word on every topic, and of course that is Judge Francis, not your uncharacteristically quiet moderator. At the end of the session, we will leave hopefully fifteen full minutes for Q&A. We hope to stop at 11:45 and do the Rick Marcus show, with the “you there, you there, you there” part of this.

We are going to begin. We are ready to go. We

are going to talk a little bit about the definition of e-data as a topic. I think we are going to start with Adam on that.

MR. COHEN: Yes. I think what I am going to do is I will put up on the screen a definition that Paul has suggested and let him explain the reasoning behind his wording, and we can talk about it as a panel.

MR. ROBERTSON: Sure. I guess a couple of thoughts just before I start this. The dichotomy that has been set up is between defendants' and plaintiffs' bar. As David and I talked in preparation, we found that on a lot of stuff there is some agreement here on the result that should be reached. We really wanted to make sure that we kept both questions in front of us at all times.

The first one was: Is there a problem; is there something that needs to be fixed? Only then did we get to the second one: Okay, if there is a problem, what is the proposed Rule change? In all instances, even if I thought that there wasn't necessarily a problem, I thought it was important to at least propose a suggested fix, some suggested language. To the extent that a proposal was put forth, at least we had something to talk about.

JUDGE SCHEINDLIN: Some of the Advisory Committee

language on this subject is found at page 6. If you want to look at what is a beginning — not a proposal, but a talking point, a thought, an idea — if you look at page 6, there is one definition, but now we are going to hear another.

MR. ROBERTSON: In this instance, the first question — "Is there a problem?" — the issue here is if we are going to put in some language in the rest of the Rules to talk about electronic discovery, do we need to define what the subject matter is at the starting gate?

If you take a look at some of the other states and federal district courts that have put in rules, none of them did so. None of them defined electronic discovery. I think that looking through it, my thought after looking at what some of the other jurisdictions have done, and the general premise that definitions are not favored in the Federal Rules, I did not think that a definition was necessarily appropriate.

I think that if you talk to folks in the places where it has been put in place, when you talk about electronic discovery, most folks do not need to run to a dictionary to find out what it means.

I thought that to the extent, though, that we use a definition, I thought about the one that had been

proposed, and I thought it was an excellent start, and I molded mine working with that one. I had a couple of comments to it, though.

One, it talks about whether the information is "created, maintained, or stored in a certain capacity." I thought that it's okay to just simply say that it is stuff that is in a digital format.

I thought, too, the final part, the attempt to try to identify some of the sources from which this information could come, the definition was "computers, telephones, PDAs, media players, media viewers, etc." I thought maybe that might suffer from the fatality that Ken Withers had identified, that things move so quickly that if you talk about a PDA, in five years folks are not going to know what that is. You know, the techies tend to change these definitions before you have taken the equipment home.

So I tried with my definition "electronic data is recorded information" — and I thought it should say "recorded" because there is a danger I think that, although some of this stuff is becoming more abstract, that the abstraction shouldn't be removed from having it tangible. It is something that is kept somewhere, as compared to something that is an ethereal idea in a witness's head —

"that is readable and available only through the use of electronic or other technological means." I put the "other technological means," and I thought that as we are moving along, maybe we do not want to limit it to electronic means, that for example biological and chemical data, although it sounds awfully farfetched today, I think some of the things that we talk about today sounded farfetched ten years ago.

So that was the proposal that I thought of.

JUDGE SCHEINDLIN: Since this really is our three-to-five-minute segment, does anybody want to say anything more about that, or should we get right into Rule 34 and documents? Anybody want to comment on this one?

MR. COHEN: Just a couple of quick comments.

One, there is a problem with including documents that were created electronically as electronic information because that can be converted into paper and then it is not what we are thinking of as electronic information.

Also I just want to point out there is a very interesting issue in terms of what is tangible when applied to data. Some of you may be familiar with all sorts of different cases, cases applying the "trespass to chattels" theory to documents, to electronic information; cases dealing with whether insurance policies cover electronic

information. So that is something we may all have different theories about in terms of use of the word “tangible” with respect to electronic information.

I think what is clear is that we are not talking about paper, we are not talking about oral testimony, and we are not talking about things like the cow in the “Replevin for a cow” case that we all read on the first day of law school.

JUDGE SCHEINDLIN: Okay. I think we should probably turn to the big topic that we have for our panel, which is Rule 34. Take a look at pages 14 and 15 for the introduction to that topic.

The question that we are really going to begin with, in the order that I mentioned earlier, is: do we need to revise Rule 34 at all to define “data” or “information” and turn away from the concept of “document,” which may be creating misunderstandings and causing problems? We are going to address that in the order we said. Adam, do you want to give us a start?

MR. COHEN: Okay. I am just going to try to set up some of the issues here and give a little bit of context.

[Slide] The current Rule talks about “data compilations,” which to us today probably sounds like a

little bit of an odd formulation. It is not a phrase that we tend to use, although in 1970 it probably sounded almost like science fiction.

[Slide] If you look at the notes where that phrase was imported into the Rule, it is actually quite prescient, I think, in terms of recognizing changing technology, the requirement of using devices, which is similar to what we were just talking about in terms of electronic information, needing to use some kind of technology to look at it.

The last sentence is kind of funny in the conclusion there. It's sort of what Ken was talking about, taking all the emails and printing them out. I think the way we look at this has changed.

[Slide] There is also a recognition of the potential need to check the source itself, so even in 1970 recognizing that there may be information that you do not see when you print this stuff out.

[Slide] I just want to point out that some of these local rules and state rules address whether electronic information is included within the scope of what is normally considered a "document" and whether it presumptively is or it is not. You have these rules in Texas and Mississippi

where you have to specifically request electronic information and it will not be presumptively considered a document.

[Slide] In Virginia, you have this rule dealing with subpoenas. It requires you to produce what they call a "tangible copy of electronic information."

[Slide] The central problem that I see, which was pointed out by the prior panel, is this issue of: do we talk about "medium" or do we talk about "information" whatever the medium? There was a suggestion in the materials of a limited change, adding "data" or "data compilations in any media."

Then there is also a talking point to address the issue of metadata and embedded data, as to whether those are included in the definition of a "document." You have the language there in the materials.

I noted that in one of the footnotes a sort of unintentional suggestion of a definition is a definition of electronic information.

JUDGE SCHEINDLIN: Okay. Dave?

MR. BUCHANAN: I guess when asked to consider the proposed amendment, the first thing that occurred to me is: what do I think we would all agree is information that

should be disclosable in litigation? The last panel I think was pretty instructive in guiding us about the types of information that parties are wrestling with in terms of discovery disputes, and then, once we understand what we think should be disclosable in litigation, then make the definition fit the types of categories to make sure that we are at least broad enough.

The things that came out in the last panel were databases, relational databases, email, spreadsheets, PowerPoints, embedded data, metadata, backup tapes. These are all things that we are talking about as being sources of electronic data that may be disclosable.

Now, I am certainly not advocating a laundry list in a Rule — I think that would be problematic — but the definition I think has to encompass those. The definition should not strike a balance between the relative burdens among the parties in terms of identifying or producing certain information. I think that is an important issue. That is an issue that needs to be addressed, though, elsewhere in the Rules, perhaps in Rule 26, or by the court in applying Rule 26.

The definition of “documents” has not caused problems for me in getting all the electronic data that I

have needed. It has included relational databases, emails, metadata, embedded data, in very large litigation. So I think the Rule has been extended in such a way so that the definition encompasses those items.

That having been said, there are two items, embedded data and metadata, that present the thorniest issues under the current Rule. I would submit — and we'll talk about it in a little bit — that those should be items that are presumptively documents but perhaps not something that you get in every case.

But in thinking about what a "document" is, it certainly includes everything within the file. It includes the creation date, the edit dates, who did it, all that information that's all within the native file. It includes the embedded information within the file. I think it is the wrong place to strike the balance in Rule 34. If there are any issues of burden, that should be addressed elsewhere.

I could certainly address a proposal for the Rule, if you want to do that now.

JUDGE SCHEINDLIN: If you're staying in this part of it, sure.

MR. BUCHANAN: There has been a suggestion, and I think Adam highlighted it, that we should be talking about

"information" or "data" that is "fixed in a medium." I think that eliminates the ethereal concept that we spoke about a moment ago, information that just crosses the wires, doesn't really register in any system, but yet it preserves the real object of a "document." There is something tangible, there is something physical. It is "information" or "data" that has been "fixed in a medium." Even if it changes over time, it has been fixed.

JUDGE SCHEINDLIN: All right. Paul?

MR. ROBERTSON: I guess I am in general agreement with David on this, that from the defense perspective we do not see that much of a struggle over whether a particular electronic piece of information is considered a "document." The struggle is always whether it is relevant to a particular case.

There are two issues that I think have been identified — and I look back. The first time I saw them identified is in the article that Judge Scheindlin and Jeff Rabkin did four years ago, which was extremely prescient in nailing some of these issues.[7]

One is: is there a need to untie this to documents? A lot of his stuff doesn't really fit our old

definition of "document." Things like cookies and other embedded information, does that really fit into the definition of "document"; shouldn't it be called "information"?

The second excellent point was: look, these are data compilations. A compilation, if you look at the definition, is a heaping together, a collection of information from other places. Much of this data is not a collection or a heaping together; it is created in the first instance. I think of a cookie again as an example.

But I think that again, although those are issues that have been identified, neither the practitioners nor the bench struggle with them. If you take a look at the *Anti-Monopoly v. Hasbro* case from several years ago, it really sets forth the law here, and I quote it: "It is now black letter law that computerized data is discoverable if relevant." I think that has really become the issue.

So I do not see the need for a fix, even though there is a little bit of a discrepancy between what is being done in practice and what is actually written in the Rules. Given that everybody accepts that the definition described in the Rules today includes not only compilations of data

[7] Hon. Shira A. Scheindlin & Jeffrey Rabkin, *Electronic Discovery in Federal Civil Litigation: Is*

but also data itself, there is not really a need for a fix.

To the extent of getting to the point if there were language to be included in the Rules, I think that adding the word “data” before the word “data compilations,” so you simply say “data and data compilations,” would serve that fix. I do not think that it would do any harm.

I do not think that you will find that it is a big-ticket item for either the defense bar or the plaintiffs’ bar or the judiciary, but it would perhaps make the Rules consistent with what everybody’s understanding is and it would clean up that confusion.

JUDGE SCHEINDLIN: All right. Judge?

JUDGE FRANCIS: I think as a judge one of my primary concerns is conflict avoidance. One way to avoid conflicts is to have clarity in the Rules, and particularly in the definitions.

I think that while there has not been a massive problem with the definition of “documents,” for the reasons that my colleagues have described, I think it may well be advisable to bring the definition into conformity with actual practice, particularly because the definition of “document” basically creates a default position. In the

---

Rule 34 Up to the Task?, 41 B.C. L. REV. 327 (2000).

absence of judicial gloss on this, people look to the Rules. "Document" I think suggests paper, and I think it may be helpful to expand that.

I think it has implications for other parts of the Rules. For example, when a party is going to respond to a document request, are they going to search for everything but then respond in paper because the Rule currently talks about "documents"? So I think in order to provide some clarity and to bring things in line with real practices.

And also I think to anticipate the future. We may agree that everybody understands now that computerized information is a "document," but when we go on beyond computers and we talk about biological information and so forth, is that going to be encompassed within the information that would be discoverable under Rule 34? I think we should adopt to that as well.

JUDGE SCHEINDLIN: Before we turn to our next topic, which is metadata, let me just ask you all one question. There is information or data that is stored and never reduced to a document, such as transient information, like spreadsheets, and they change every time the parameters are changed, or a daily example might be an e-ticket that is never a document unless it becomes printed. So there is

information in data that is simply stored on a medium but is not yet a document. Does that question make you think that that needs to be addressed in this definitional Rule 34?

MR. BUCHANAN: The important point I think with an e-ticket, for example, is there is a database behind that e-ticket that contains all the parameters. There is something electronic in nature that has been fixed in a form that contains all the parameters of that e-ticket.

The same with the spreadsheet that you highlighted. While it may change day to day and you have multiple versions of the document, the formula, for example, within the document is the same perhaps, or maybe that changes over time too, the resulting numbers.

JUDGE SCHEINDLIN: But as the last panel said, if you printed it out, you would never see that formula. So the question is: can you obtain that data when you think of the term “four-cornered document?” That is the question I am asking.

MR. BUCHANAN: I agree. I think that is more of a production issue in my mind, the format in which it is delivered to the other side.

JUDGE SCHEINDLIN: Okay. Anybody else want to address that?

MR. COHEN: I just want to say that it seems that with the types of electronic information that we have these days and that are becoming more and more prevalent, such as transient data, instant messaging, digitized voicemail, we are moving closer and closer to what is more like oral communication in how evanescent it is.

We might ask ourselves: if we are going to require data like this to be captured and produced, does this mean now that when we have oral conversations about a case when we are under a duty to preserve we should be recording it all?

JUDGE SCHEINDLIN: Let's turn to an issue that at least in the Advisory Committee we spent a lot of time thinking about, and that is the question very specifically now of the production of metadata and embedded data. I shouldn't have thrown in the word "production." Put that aside for a minute. Just whether Rule 34 conceptually would call for the production of metadata and then later embedded data. I would like to take those separately because they are different concepts.

Let's talk for a minute about metadata, starting with Adam.

MR. COHEN: Okay. Just to set up the issue the

way it is set up in the materials, the issue is: do you make this a routine requirement of production; do you make it a permissive requirement?

What are the positions on either side? You know, on the one hand, opposition to routine requirement would be based on the notion that there is not really a likelihood that it is going to be terribly material. It is going to add costs. On the other hand, there are situations where you are going to be adding more costs by stripping that data our — and believe me that happens a lot in real life, oddly enough.

On the other hand, you might need the metadata to facilitate the searching, the manipulation, the kinds of litigation databases that people use right now to handle large amounts of documents. Some of the formats that people produce their documents in, these image formats without the metadata, require a lot of work before they are they actually usable in one of these databases.

So the question becomes: should this be presumptively something that gets produced or is it only available by special permission? We have some positions on that that have been taken by members of the bar and the judiciary.

[Slide] The Sedona Conference document shows a position where this type of information is presumptively not something that is included in a production unless there is separate analysis on a case-by-case basis.

[Slide] The ABA talks about “duty to preserve” in a very broad way, specifying it at “media” rather than the type of information.

[Slide] That brings us to form of production, so why don’t I let the panelists talk about metadata?

JUDGE SCHEINDLIN: We are going to hold off on form for a little while. Let’s just talk about the concept of metadata as something that ought to be produced with the information, or not. Dave?

MR. BUCHANAN: Again, I think, focusing within Rule 34, the conclusion I reach is that this is supposed to talk about the types of forms of documents you can request or the types of forms of information that can be requested in litigation.

I think metadata unquestionably can be relevant to a claim. We have seen — well, how about in paper productions of years past a file routing slip on the top of a document that showed when a document went to somebody, when it moved to the next person; a revision history that

tracked changes to a contract over time? These types of things were discoverable. They were affixed to a document or to a file.

Now we have electronic documents that have different flavors of similar concepts. Rule 34 needs to contemplate that those types of documents are documents or information discoverable in litigation.

If there needs to be a balance struck, it should be struck elsewhere.

One thing that I think is important to note is there are other proposed changes in Rules 26 and 16 that require the parties to talk. What I heard from the last panel, and I think it is an important issue, is that the parties need to talk. I would expect that metadata and embedded data would be something that would be discussed during those early planning conferences both privately and with the court.

So I think Rule 34 is not the place to limit this. Rule 34 should be encompassing, though, of metadata and embedded data. The question is: is Rule reform necessary to accomplish that?

This is the only area in my practice I think where there is any debate with defendants about whether metadata

or embedded data is a "document." So I do believe that clarification would be helpful in that regard, but it should not be on a showing of good cause within Rule 34.

JUDGE SCHEINDLIN: So in other words, you don't think it's second tier; you think it is presumptively part of the data?

MR. BUCHANAN: It is. It's within the file wrapper.

JUDGE SCHEINDLIN: Okay, it's within the file wrapper. A little patent law. Okay.

Paul?

MR. ROBERTSON: I think that the first thing to do — you know, the issue of whether there is a problem, I think there is a need to unpack embedded data and metadata for a second because they really are different things.

Folks talk about metadata and they quickly say it's information that is embedded in the document. If you look at some of the articles and some of the writings on this, the excellent article by Judge Scheindlin and Jeff Rabkin talked about embedded data as being metadata; they used the terms interchangeably, as a lot of folks do. Sedona talks about metadata being embedded data.

But they really are different things. I think

that the metadata is the information about the information in the document — things like in an email the code that tells how the email is to be delivered, how it is to be routed; the information in a Word document, paragraph shifts; information in a spreadsheet about how calculations are to be made.

I think, on the other hand, when we talk about embedded data, it is a very different animal. It is typically user-created edits or information that is put into the document purposefully — things like track changes; things like a sticky note that you put underneath; things like other versions of the document that are hidden beneath it. I think that those are very different things.

I think that when you are talking about embedded data, the way that I understand it more easily is to think about embedded edits. I think that edits to a document certainly are in certain circumstances presumptively discoverable as a type of draft of the document.

I think, on the other hand, in 99.9 percent of the cases metadata is irrelevant because it is not even the envelope that you are sending the email in — and most of this stuff, by the way, is about email — it is not the

envelope information, who the email is from, to whom it is being delivered — but instead it is instructions that you have given to the mailman about how to take it, how it is to be routed, and then information about how that email was actually delivered.

In most instances you don't keep the FedEx package, you don't keep the instructions telling the FedEx man or woman to go to this certain place. It is not typically relevant. It is the equivalent to having to, after doing a document production, to go back and say, "I want to fingerprint your data room to find out who was in there and who was not."

That said, I would certainly agree that in some cases it is very relevant. Martha Stewart is an example of a case where you wanted to find out about who edited this document and when.

But the question I have next is: is there a dispute about whether that is considered a "document"? I think that again the Rules do a very handy job of this. I don't see any cases out there where a court has said, "You can't have it because it's not a document." The issue becomes, "You can't have it because it is not relevant."

Even in those cases where you do need metadata,

information about information, it is usually targeted to a very few spreadsheets, a few emails, and in most cases a requesting party does not want to get with each document sometimes 800 pieces of information about that email that neither do they need nor they understand.

So I think that — is there a problem? — I don't think that there is a problem with respect to metadata. I think that most folks understand that it is a "document," but the question is whether it is relevant.

JUDGE SCHEINDLIN: And the judge?

JUDGE FRANCIS: I love seeing so much agreement between plaintiffs and defendants.

I think there is agreement that metadata and embedded data are information and that they are at least potentially relevant and therefore come within, or should come within, the definition of "document" or "discoverable information," however we characterize it.

I think the tougher issue is whether there should be some good cause requirement imposed before a requesting party has access to that information. There I would point out that as a judge one of the values that I try to embody is doing justice, and that means being able to adapt the law to the facts in a particular case. The more constraining

the Rules are, the more difficult it is for me to do that adaptation.

If there is a good cause requirement, it is a thumb on the scales of justice, and somebody is going to have to overcome that presumption in order to get what may ultimately be relevant discoverable information.

I think it is preferable to leave that to be determined on a case-by-case basis. I think the ABA's approach to putting the burden on the requesting party to ask for that kind of information is fine, but to place a burden of persuasion on that party I think would probably be a mistake.

JUDGE SCHEINDLIN: I think we are going to talk more about metadata and embedded data when we move to form of production, so don't worry that we have left it behind. We are going to get a second round of hearing about it.

We are now going to turn to the form-of-production question. You can look at pages 16 through 20 of your exhibit, where the Lynk/Marcus memo discusses the form-of-production question as possibly addressed in Rule 34(b).

[Slide] The question there is whether Rule 34 should be amended to require, either permissively or mandatory, but that the requesting party state the form in

which the e-data is to be produced. If so, should that request be as simple as “I want paper” or “I want an electronic mode of production”? Or should it be more complicated, such as, “I will be satisfied with a mere TIFF image” or “I want a PDF searchable” or “I want native digital information produced in a specific format, like a DVD, and it has to be compatible with my Windows operating system”?

So the question is: what level of specificity should the requesting party have to express if they should have to make a choice at all?

[Slide} Then, of course, the flip side of that question is: if they don’t specify, is there a default mode of production?

The third question, I suppose, is: the producing party, what is the ground of objection there? Can the producing party say, “I shouldn’t have to produce it all because it is inaccessible”?

So it is sort of that series of questions that we are about to address. And I think, inevitably, in addressing those we are going to get back to the metadata and embedded data because how you produce it may mean whether or not you include those types of information.

So with that quick background — maybe I did too much — Adam?

MR. COHEN: Okay. I think that is right. There is the segue right there — and we should probably talk to some of the technology people about this — but the parties' determination of whether they are going to resist production of metadata or embedded data may depend largely on what format they have their documents in and they are ready to produce them in. If they have gone and printed out all the emails and scanned them in and created electronic images that are stripped from the metadata, then they are not going to want to produce the metadata.

Why require or permit a specification of the form? Well, if you ask for documents in a certain form, this should preclude you later on from coming in and saying, "No, no, I want something different." On the other hand, making it optional may make sense because at an early stage in the case when you are formulating your request you may not know what format you need or what the other side uses or what is going to make the most sense in general. In any case, there is always going to be a need to balance the burden of producing in a certain form against the utility to the other

party.

Another issue that has come up — I don't know how common this issue is — is that there are certain proprietary aspects to certain formats. For example, parties have data that can only be viewed with certain proprietary software and generally will resist producing that type of software. At any rate, this does seem like it would make sense to discuss it in the initial conference.

[Slide] If you look at the Sedona Conference, the position that they represent, they talk about the importance being the substantive information content, that you should not have to produce documents in more than one format. They suggest that "production of electronic data in a commonly accepted image format should be sufficient." Now, that has implications of course for metadata and embedded data. "Data that is not ordinarily viewable or printed when performing a normal print command need not be produced."

[Slide] At the same time, there is a recognition that:

- Electronic formats may be preferable in many cases;
- Whatever format is chosen should deal with the

genuineness/authenticity issues;

- That there should never be a requirement to produce in both hard copy and electronic form. I know this is something that is often the subject of debate based on the case law that is out there already.

[Slide] The ABA has said that you should consider asking for production in electronic form, you should consider asking for production in a form that gives you the ancillary information.

[Slide] And then there are some of the cases that were talked about in the material that deal with these issues in different contexts.

Here the *Bristol-Myers* case[8] shows what happens when a party goes ahead and scans all these paper documents into images and then they want to produce them back in paper as per the ancient past. This was a case where I guess no one had said anything about the fact that these documents were available electronically and were trying to get somebody to pay the cost of a normal paper production when that wasn't really necessary. In that case, they were required to produce an electronic format.

Interestingly, and probably most controversially,

[8] Bristol-Myers Squibb Securities Litig., 205 F.R.D. 437 (D.N.J.

there was no requirement that the other side, the requesting party, pay for any of those costs that were involved in creating the electronic format.

[Slide] The issue of the proprietary format came up in this case that is mentioned in the materials, the *Honeywell* case,[9] where PriceWaterhouseCoopers stated that production of these documents in a usable form would require the use of proprietary software or large cost. The court basically gave them the option of either producing the proprietary software, the proprietary format, using the protective order, or pay for it themselves.

[Slide] And then finally — and this shows another aspect of this issue — the *McNally* case,[10] which shows no presumption that you get the computer files when you've got the paper production because you need to show some sort of special basis for it.

JUDGE SCHEINDLIN: Just a quick moderator comment. At Tab 9 of your materials all of these cases and many others are summarized for you, so it is a great resource in the back, the annotated case law.

2002).

[9] In re Honeywell International, Inc. Securities Litig., 2003 WL 22722961 (S.D.N.Y.).

[10] McNally Tunneling v. City of Evanston, 2001 WL 1568879 (N.D. Ill.).

Okay, Dave?

MR. BUCHANAN: I suspect this will be more of a point of departure between the plaintiffs and the defendants, and that is the form of production.

There is no question that plaintiffs prefer as a general matter native production of electronic files. That provides all the embedded data, the metadata to the extent it has been appropriately preserved. It gives you the opportunity to quickly search for terms. In short, it puts you on the same playing field as the defendants, or the company at least, in accessing their own data. Those are the arguments plaintiffs use to get native production.

But you may not want a native production in all cases, and that is why it is important I think for there not to be a presumptive production format of native, because we talk about proprietary formats — or even if we're talking about relational databases, if I have to receive all of your databases in a native format, I may not have the capability of rebuilding that, as opposed to me meeting with you and discussing the appropriate searches to run on the data, extracting the data, running it in reports and producing the electronic versions of the reports that I can then load into my database.

So I think again this is something that is in Rule 34, but I think it is something that will be addressed quite specifically by the parties at their 26(f) conference and at the Rule 16 status conference as to how to treat non-paper documents: how are we going to treat electronic data; how are we going to produce it; how are we going to preserve it; what are we going to do with the embedded data; what are we going to do with the metadata?

The Rule needs to contemplate the production of native data. That is the most easily usable form for litigants as a general matter. That statement can be thrown completely out the window, though, when it came to large proprietary systems where a smaller plaintiff, or even a large plaintiff, didn't have access to the software to view it.

JUDGE SCHEINDLIN: Let me just ask you a few quick questions. So do you favor a Rule that mandates the requesting party should select the form of production it wants? You are usually a requesting party. Should you have to state what you want?

MR. BUCHANAN: I will, and I do, and I will do that in the Rule 26 conference and I will do it in the Rule 16 conference. I think making it permissive to do so and

making it permissive for the other side to object to the form requested is fine. But I think having a presumptive form of production would tilt the scales in favor of something that may not work across a large-scale litigation.

JUDGE SCHEINDLIN: That was only my first question. Should you have to specify to avoid confusion?

MR. BUCHANAN: I think it should be permissive.

JUDGE SCHEINDLIN: Permissive, okay.

The other question I have for you is: should the Rule talk about "the data should be produced in the form in which it is created, in which it is ordinarily created"? Should that be the fallback, presumptive form?

MR. BUCHANAN: Here's what I want. I will let people who are good with language and the Drafting Committee tell me the best way to implement it. What I want is information that is as accessible or as usable as on the defendant's system. I mean that is what I want. In many cases that is native files. In other cases with complex databases, it may be an extract of the data from the databases.

JUDGE SCHEINDLIN: Well, we have those words "created, stored, maintained" that you will see throughout this memo. Which of those words do you like, or all of them

— in the form in which it is created, in the form in which it is stored, in the form in which it is maintained? Do any of those excite you?

MR. BUCHANAN: What excites me is getting the data in an accessible — no, those words don't excite me, frankly.

JUDGE SCHEINDLIN: Okay.

MR. BUCHANAN: It is in a form that is as accessible or as usable as the form in which the defendant maintains their data.

JUDGE SCHEINDLIN: So maybe "maintains" is the one that does.

Okay, Paul?

MR. ROBERTSON: I hate to disappoint once again, and I think it is a function of how reasonably David approaches most of these issues, but I don't substantially depart from what he is saying.

I think that, again, the first question, "Is there a problem?" — as he said, "Look, in some cases I want the data in its native format," and I think that is absolutely right. In some cases, there are issues where the data in its native format is relevant. I think that in other cases you don't want that.

The Sedona Principles took the approach of: Look, in most cases production in paper or TIFF images is acceptable. That draws gasps from a lot of plaintiffs, and rightfully so in these mega-document cases. But I think that sometimes we forget, with all of these numbers of terabytes and petabytes up on the screen, that in most cases the typical sides are not looking at that kind of volume of documents, they're looking at a smaller volume of documents.

So when you create a default position that says things like "you have to express how you want electronic documents to be delivered to you," often it is only 1,000 pages or 2,000 pages, and so getting electronic documents isn't necessarily useful.

I think there are two questions here. One is: is there a presumption that a party should have to produce things in its native format? I think that the answer is there should not be such a presumption because it is a very fact-driven issue.

The second issue I think is a little more conducive to having something done in the Rules, and that is it is a communication problem. The three cases that you see collected on page 8 of the materials are all situations where a producing party gave paper and the requesting party

said, “I don’t want this. I wanted something in electronic format.”

JUDGE SCHEINDLIN: Should the requester have to ask up-front and specify?

MR. ROBERTSON: I think that the answer is this. A caveat, though, is of course if you specify electronic documents — that doesn’t get you to where you want to go, by the way, because you will sometimes get a TIFF image, and that still is the equivalent of getting a hard-copy document. So I think that, as Dave was saying, you want something that is both electronic but then searchable in the same way that the defendant had it.

I think that there are three ways to go about this: education is one; two is putting something in the Rule 26(f) conference; and three is putting it in Rule 34. I think that the first two answers are the way to go. I think that this is something where education is needed, where places like the *Manual on Complex Litigation (Fourth)* has some language to encourage parties to talk about this. I think it is important to put it in the Rule 26 checklist, to make both parties talk about these kind of things, or suggest that they do, so that they can avoid these situations in cases where they are relevant.

I don't think it is appropriate to put it in Rule 34 for a couple of reasons. One is that there are good reasons that defendants do not to produce things in their native format, and it is not simply to hide things. It is because, for example, you can't Bates stamp things; they are manipulatable by the discovering party; and they can be changed so you go show up in court and something that you produced in one format looks much different than it otherwise did.

I think the other problem is that if you set up in Rule 34 the suggestion that one party "must" or "may" specify and the other party has the right to object, you create a sort of presumption that there is this right to get things in a native format. And I don't think that anybody is going there. I think people are saying it is a communication problem, which I think is best handled with 26(f).

JUDGE SCHEINDLIN: How does that play into our discussion of metadata and embedded data, though? I mean, if you are doing a TIFF image, you are presumptively not getting it. If you are doing paper, you are presumptively not getting it. I think Dave said he thinks he presumptively should get it.

MR. ROBERTSON: I think that there — and I'd be willing and eager to hear David's comments on this, of course — but I think that in many cases the metadata is not relevant. When you say metadata, I think that you want to have it searchable.

You want to get an email that even if it is produced in TIFF, you have a concomitant list of searchable data that allows you to organize it by sender, by recipient, by date. That is important. But when I think of metadata, I think about pages and pages of code about how the email got from Tallahassee to Gainesville via some server out in the western part of the country. I don't think anybody wants that and it's very rare that it is needed.

JUDGE SCHEINDLIN: Given our time constraints, we are going to the judge.

JUDGE FRANCIS: This is a series of issues where I think I am firmly ambivalent. I think that I disagree with Dave and believe that it probably would be helpful in avoiding conflict to require the requesting party to identify the form of production. Now, I do not think that that needs to be done in the Rule very specifically, but the Advisory Committee notes might point out that the greater the specificity, the more likely that we will be to avoid

future problems.

If there is such a requirement, should there be a default mode identified? I think that is important. If there is no default mode, then judges are left with the question: "Well, he didn't identify the form of production; that means I should impose a default mode, or it means the request should be stricken?" I don't think that provides enough guidance, so I think there needs to be a default mode.

Which brings us to the $64,000 question, which is: what is that default mode? There I am truly at sea. If I were to write a Rule for today, I think I would say paper production, because that is what everybody is capable of doing, everybody who receives it is capable of analyzing it. It is cumbersome, it is burdensome, but everybody can deal with it.

But as I hear our technological people tell us that paper is going to disappear, that I think would be a Rule that would be quickly archaic. So I am looking for something that would be a reasonable default position. But I think that there needs to be a default position.

And finally, in terms of whether there should be an identification of the responding party's right to object

because it is inaccessible or hard to produce, I think that is in the Rules. I don't think there is a necessity for electronic or other kinds of information to specify the opportunity to object.

Rule 26(b)(2) sets out terrific guidelines for weighing factors to determine whether a document production is too burdensome, too costly, and so forth, and those Rules encompass the question of accessibility. I don't think there is anything necessary to be done there.

JUDGE SCHEINDLIN: Given our time constraints, our last topic, responding to interrogatories, which is found on pages 21 and 22 of your materials, we are going to give that the shortest treatment so we can get to your questions. I'm just going to call on Adam to quickly cover the question and leave it at that.

MR. COHEN: Right.

[Slide] As we race through this, there is the current Rule. I hope that was enough time to read it.

[Slide] And we have some suggestion as to what we could do to sort of tweak this to deal with electronic information. That would include producing the electronic information and, I suppose, identifying it as well, and possibly giving computer software so that you could derive

the answer to your interrogatory from that electronic information.

[Slide] Questions raised in the materials are:

- Whether we need to include this option of giving the "computer software," or whether we stay with, I guess, the more general solution of giving sufficient information to find what you are looking for.

- And a question as to whether parties are employing 33(d) with regard to hard-copy and computerized files. In my experience, they are. This is one of those situations where you might invite somebody over to come run queries on your database.

- And how does the fact that in many cases data produced is prepared for the purpose of responding to an interrogatory — how does that mesh with the obligation imposed under Rule 34?

JUDGE SCHEINDLIN: I think the real question here is: if a producing party takes the option of producing in this way, you would think the requesting party wants to be able to use it; so if you are going to produce it, do you have to produce enough to make it usable, which may mean the software or other material that goes with it?

Do you want to say just one thing and then we're going to go to questions?

MR. BUCHANAN: The premise of this provision is that it is as easy for the receiving party to access the data as it is for the defendant or for the producing party. If you don't have the software tools to access the data, you don't have the same ease to access the data that the defendant does. So I think any production of electronic data pursuant to an interrogatory request has to be accompanied by the tools to access the data.

JUDGE SCHEINDLIN: Okay.

Now, we have left a full fifteen minutes for you folks, so please use it. I see a hand way back there.

QUESTION [Paul J. Pennock, Esq., Weitz & Luxenberg]: Paul Pennock from Weitz & Luxenberg.

It seems to me that if — well, let me ask you. If we want to specify what we want from a particular company, don't we first need to know how they are doing it, what they have stored, how they have stored it, where they have stored it? And, to touch on the balancing test, the utility against the cost, what are they saying is the burden of producing it? So in order to get to that specification stage, don't we need some type of mechanism to get some very

quick and early depositions of the people in the know? I mean something even a little easier to do that 30(b)(6), something very focused on IT issues, where we can get in — Dave Buchanan has done these depositions many times — and figure out exactly what it is that we are trying to get, and then we can specify what we need?

JUDGE SCHEINDLIN: Let me make a point to the whole audience. You don't have to ask a question. Unlike the usual panel where they say "ask a question, not a statement," not true. You are perfectly welcome to make a statement. So if you want to answer your own question there, what would be your suggestion, and is it Rules-based? Whatever your suggestion is, is it Rules-based; is it something we should do in the Rules?

QUESTIONER [Mr. Pennock]: Yes, I think it would be very helpful to have a Rule that guides magistrates and judges to say, "We're not being unreasonable if we're stepping in within weeks of commencing an action and asking for a series of depositions of particular IT people in order to identify answers to a series of questions," a Rule that would be very specific in order to make the depositions happen quickly and efficiently.

JUDGE SCHEINDLIN: But you can do that at the Rule

16 conference with the court.

QUESTIONER [Mr. Pennock]: You could, but it just seems to me it would be better if the court had some guidance that this is not only a good idea but a necessary predicate to pursuing the electronic discovery that we may be pursuing under these new Amendments.

JUDGE SCHEINDLIN: Okay. We've got to go all over, but here. This gentleman was next. You will be third, Jonathan, I promise.

QUESTION [Michael P. Zweig, Esq., Loeb & Loeb]: I'm Michael Zweig from Loeb & Loeb.

A practical question with respect to the form of production. What method or methods are being used if you are producing material in so-called native format to replace the traditional Bates numbering system so that we could have some evidence of when it was produced, who produced it, where it was produced from, etc.?

JUDGE SCHEINDLIN: Okay. I think we can turn to the former panelist George Socha to maybe address that, because I think there is the equivalent of Bates stamping. I think I have seen it. George?

MR. SOCHA: Sort of.

JUDGE SCHEINDLIN: Sort of, okay.

MR. SOCHA: I think this is one the knottiest issues of producing information in native form: how do you track what you got, who you got it from, who you've given it to, and so on? We're used to Bates numbers with paper. It's easy to do. We do not yet today have any generally accepted counterpart to that. The best you can do is the type of thing that Joan I know has done as well: you keep a log of where the file came from, a path of where it came from if necessary. It's a cumbersome process. We need to figure out a way of doing that, and as far as I know there is no generally accepted approach.

JUDGE SCHEINDLIN: Okay.

MS. FELDMAN: If I could reply now?

JUDGE SCHEINDLIN: Yes, Joan, right. That makes sense.

MS. FELDMAN: You can basically put a wrapper around these documents that allows you to affix any information. Most of the software out there today allows you to easily switch back and forth between native format and this wrapper format.

I have one more comment about embedded versus metadata, if I may.

JUDGE SCHEINDLIN: I promised Jonathan.

MS. FELDMAN: Okay, then I'd be happy to do it.

JUDGE SCHEINDLIN: Okay. Jonathan Redgrave?

QUESTION [Jonathan M. Redgrave, Esq., Jones Day]: Jonathan Redgrave. I'm with Jones Day.

I've got a number of observations on this panel and the presentations here, because in many ways this is the nub of are the Rules going to change.

First — and it is a point that I think Paul made very well — we've got to remember, for the Rules Committee, that the vast bulk of cases don't deal with the mega-documents, the mega-issues, and any Rule change really has to take account of the fact that it shouldn't force onto those cases a world in which it just doesn't make sense for the economy. We've already got enough issues about the price of litigation putting people out of the reach of district courts, and we do not want to impose something that will go further.

Second, the perfect is often said to be the enemy of the good. Perhaps the definition of a "document" could be tweaked or added upon, but is it really necessary? Is it something that we need to change or is it good enough as is? I think my response to that is it probably is if it was standing by itself, *Ceteris Paribus*.

But if you really want to get down into making maybe some tier presumptions about whether certain data should be produced or not produced in the case presumptively — for instance, if you are going to take out inaccessible data, or if you want to take out metadata — if you are going to create those kinds of presumptions in the Rules somewhere else, in Rule 34, then perhaps you do need a separate definition so you can refer back and make it a better process. So I think for changing the definition that is the only reason you want to do that, if it is going to be part of a bigger picture scheme of changing the Rule to build on presumptions.

From the Sedona work, obviously I think there are some good places for building in presumptions — and remember, presumptions are things that can be overcome — but to deal with the vast bulk of cases they could be done this way, but then there is a place for those exceptions to be dealt with.

I think the Rule 16/Rule 26 conference is a great place to get that initial information that helps you understand where that case is going to break on the presumptions one way or another. In the IPO cases I know up

here in New York, part of that Rule 16 conference was an enormous survey of information of the various kinds. They had to come forth and say what their processes were, where their tapes were, where their data were, where the key employees were. That was something that was done up-front, and I think it tremendously guided those parties both on finding the responsive discoverable information and also keeping costs down, so it helped both sides.

In terms of the form, the last thing on the form, I think the proposed Rule change is a good idea to force people into saying, "I really want it this way," and the responding party saying, "I can do it or not." It builds on the Rule 16/26 concept, but I think the proposed Rule does make some sense in that area.

JUDGE SCHEINDLIN: Ms. Tadler? I promised her next. He is absolutely next.

QUESTION [Ariana J. Tadler, Esq., Milberg Weiss Bershad Hynes & Lerach]: Ariana Tadler from Milberg Weiss.

I think both David and Paul commented that they do not think there should necessarily be a presumption of native form being required. But I guess an observation, or perhaps just an issue for thought, is whether there then has to be a corresponding counter-rule, which is that the

expected producing party should not have the ability to set the form prior to any conference and then have the ability later to say, “We’ve already set it in a form and now it’s a burden to go back.”

That is something that certainly I think we as plaintiffs’ counsel have encountered at times, where we understand that on the defense side you may be wanting to set things up so that they are available and preserved, but simultaneously we want you to have an understanding that we may very well want it in a different form and we don’t want to hear “burden” later.

JUDGE SCHEINDLIN: I promised, so right there.

QUESTION [William A. Fenwick, Esq., Fenwick and West]: Bill Fenwick, Fenwick and West, from Silicon Valley.

I have three points to make really. The first one I would consider to be more global. One of the problems that we are struggling with here is that we are trying to impose what is an obsolete paradigm on what is the current reality. What I mean there is we should try to get rid of this word “document” and we should go to “information.”

If you look at Rule 34, the word “document” appears four times in Rule 34. If you changed it to “information,” I think you would enlighten rather than

confuse.

The second point that I wanted to make had to do with the business about you can't tell if there has been any change in data if you produce it in native format. That is not true. You can put a lock that will tell you immediately if one bit of it has been changed. So that is not an obstacle to native format. That doesn't mean that I necessarily endorse native format in every occasion.

The other issue — I'm sorry to say I take issue with the Judge because I thought he did an excellent job — is I think your default is wrong. It should not be paper. If you want paper equivalent, go TIFF. TIFF is pretty easy to access and it greatly increases the efficiency of the review.

I happen to think that images are not the way to go because they do not become machine-searchable, and if you start looking at this from the standpoint of its utility to what you talked about, justly resolving disputes, you ought to be looking to the cost of reviewing, identifying, and specifying the presentation that you are going to make.

JUDGE SCHEINDLIN: Thank you.

QUESTION [David M. Bernick, Esq., Kirkland & Ellis, Standing Rules Committee]: My name is David Bernick.

I'm with Kirkland & Ellis in Chicago. I'm a member of the Standing Committee.

I think we are finding here kind of a reinvention of the same discussion that the prior panel ultimately came to, which is focus and relevance. The issue in Rule 34 is not whether it is relevant, it is not whether it is costly; it is whether something is producible. It is not whether production is required; it is simply what is producible. As a consequence, the definitions or the items in Rule 34 are incredibly broad. I suppose we could make them even broader if in fact what is then included is producible.

But what we end up talking about in the context of this discussion concerning Rule 34 is the ultimate issue, which is then managing the production requirement. That doesn't have to take place through Rule 34. That should take place using all of the other means that are available under the Civil Rules.

For example, the conferences at the beginning of the case — there don't even need to be depositions. You could have an informal conference with the court that helps the parties identify what the landscape of information is and what kind of production requirements make sense. Rather than creating an obligation on the producing party to

specify or on the requesting party to specify, have a discussion at the beginning, find out what the available information is, and begin at the very first part of the case to focus on relevance and need. That really should be the benchmark for then figuring out these other questions.

So I don't think it is really a Rule 34 issue. I think it really focuses on the other Rules.

MR. BUCHANAN: May I respond?

JUDGE SCHEINDLIN: Yes. One of our panelists asked for a quick response to that, and then we'll go to Mark. You've had your hand up for a while, Mark.

MR. BUCHANAN: Just one observation, and I think I echoed this in my comments. That is that the conferences are really where I would expect that the parties will be able to frame whether native is a good idea, whether embedded information or metadata is a problem in the context of a given case.

The problem that litigants have today — and Ms. Tadler highlighted this — is that getting the information today from the defense prior to taking an MIS deposition is virtually impossible. And I'm not sure that the meet-and-confer structure will flesh out in enough detail the type of information that is necessary to develop a meaningful

request under Rule 34. It is my hope that the Rules will be interpreted broadly to require the litigants to really share their IT information at an early phase so that we can accomplish those ends.

And just to respond to Mr. Redgrave on the IPO data questionnaire, what precipitated that questionnaire, as you know, was a spoliation challenge raised by some reports in the news media. Absent that, I'm not sure the litigants would have felt comfortable making a motion to compel for certain discovery that led to that questionnaire.

So I think the current Rules system as it relates to the meet-and-confer process is broken in the context of 2004.

JUDGE SCHEINDLIN: Sadly, we have only three minutes. We are taking until 12:05 because we didn't start until five minutes late, so we have three minutes left. Mark?

QUESTION [Mark O. Kasanin, Esq., McCutchen, Doyle, Brown]: Mark Kasanin from Bingham McCutchen.

As a former member of the Advisory Committee, I have a comment on the rule-making process itself as it relates to Rule 34, and perhaps to other Rules. We already know that technology may very well overtake whatever we do

here because of the three-four-year lag that it is going to take to get a Rule in place even if we move on it right now. By then, some of the things that we are talking about now may well be obsolete or soon become obsolete after that effective date. Other technology will come into play.

If we start with very precise and more specific definitions, I think what we are doing is embarking on a course where we need to keep amending the Rule as we go along to take account of future developments. I'm not sure that that's something that the bench and bar really want to see, and we know that from past experience, that the more times we amend the rules, the more criticism we get. So I think that is just an overall consideration to be kept in mind.

JUDGE SCHEINDLIN: Okay. We've got to go to this side.

QUESTION [Patricia A. Martone, Esq., Fish & Neave]: Patricia Martone, Fish & Neave.

I handle very complex, high-stake, patent infringement litigation. My experience is I think it would be excellent to make sure that electronic discovery is discussed in Rule 26 conferences and Rule 16 conferences. But one of the challenges that we always face is how not to

drown in so much information and documents that we lose sight of the big picture of the case, and keeping the big picture of the case is very important in representing both plaintiffs and defendants.

My experience is not about arguing about metadata. My experience is trying to find out as a plaintiff, for example, what kind of databases the company keeps that allow me to track processes used that are infringing processes to products. And so in the beginning of a case I don't need to discuss the IT aspects of every document in the company, and so I would rather not be — while I think we have to discuss the initial format of production, I would like the flexibility to take depositions from engineers, for example, find out how they access this information, and then perhaps ask focused inquiries.

So I guess I would say I am not in favor of amending the definition of "document." I am very much in favor of early discussion of electronic discovery.

JUDGE SCHEINDLIN: Joan, last word. Did you want to talk now about the metadata and embedded data you mentioned earlier you were about to address? We'll close with that.

MS. FELDMAN: I will say that you may have to

discuss form of production as it relates to preservation. Let me give you an example of a document, a PowerPoint presentation. You have metadata associated with it. That includes the true author's name and the company that it originated from. That's in your metadata, that's in your file properties.

Within your PowerPoint presentation you may have an animated slide from somebody standing like this [gesturing] with the animation going like this [gesturing]. If you print that document out, you don't get the metadata, you don't get the true authorship information, and you don't get anything that might have been embedded in there, say through animation.

Why would that be helpful to you? What about a case involving a biotech firm that had spent $6 million in developing their product, the PowerPoint slides were acquired from someone who had left to start another company? If they had simply printed out that information, no one would ever have known that the true authorship and ownership of that slide was actually the originating company; they wouldn't have known what those edits were.

So when you are making this decision about "here's how we're going to produce it in a static format," you have

to make some provisions at the very least for preservation of the native and some way of linking back because it is part of the document.

JUDGE SCHEINDLIN: Thank you.

I want to take a minute to thank these panelists. Believe it or not, it looks so easy, but they put in a lot of time to prepare. So thank you very much.

DR. CAPRA: We will take a lunch break. We will reconvene here at 1:15.

[Session adjourned: 12:05 p.m.]