

Webcast: Emerging Technologies For Finding And Mapping Discoverable Information.

Date and Time: Thursday, February 16, 2006 at 1:00 PM ET

Sponsored by Xerox Global Services

ASSOCIATION OF CORPORATE COUNSEL

Moderator: Glen Flowers

February 16, 2006

Operator: Go ahead please. (Glen) go ahead please.

(Glen Flowers): Good morning or good afternoon depending on what part of the country you're in.

And welcome to our session on emergent technologies and finding and mapping discoverable information.

We trust you'll find this to be a very interesting meeting. And we're very fortunate to have an excellent panel with us today for this discussion.

I'm (Glen Flowers) and I'm a principle with Xerox Global Services in the Corporate Litigation Management Practice. Today we're going to talk about the challenges that we all face in mapping the identity and location of information that we need for compliance or litigation purposes. We're going to get several perspectives from our panelists. You are free as participants in the audience to submit question. There should be a box in the lower right hand corner of your monitor where you can type in questions.

These questions will be reviewed throughout the presentation. And after each speaker has given their part, we will come back and have different people address the questions.

Finally we'll have concluding remarks and survey for you to fill in so that we can get useful feedback for future presentations.

Our panelists today include (Richard Wolf), who's the Chief Compliance Officer and Senior Vice-President of (Senden) Corporation; (Ann Curshaw), who's the Managing Principle of (A. Curshaw) PC, Attorneys and Consultants; (Robert Ballor), who's the recently retired Vice-President and Chief Technology Officer for Xerox Global Services; and finally (Nicholas Economu) who's the Chief Executive Officer, Founder, and Chairman of H5 Technologies.

When talking with people in the industry nowadays, whether they be with the Office of General Counsel or with law firms, there's a standard theme and a standard question about how do we handle preparing for the inevitable, whether they be compliance requests, administrative review requests, or discovery requests. How can we be more prepared to handle the inevitable requests that will come to us?

Operator: And (Glen) ((inaudible)) oh, this is the operator. If you can increase in the volume a little bit on your head set that would be great. Everybody else should be on mute. Thank you.

(Glen Flowers): Thank you. So in ((inaudible)) one of our speakers will address how he has handled being more prepared within his company.

And then there's the inevitable challenge of when you are engaged in a matter, how do you move from the universe of information through what is relevant? How do you set aside your privilege? And how do you come up with what is responsive in the most efficient manner, minimizing your risks and minimizing the costs of going through that procedure?

Our first speaker is (Richard Wolf). And (Richard) is as I mentioned the Chief Compliance Officer for (Sending). He's globally responsible for all business records and information management systems. He led the establishment of (Sending's) first records management and compliance program. He's the chair of the corporate counsel committee for the American Bar Association's section of administrative law and regulatory practice. And he's also the President of the greater New York chapter of the Association of the Corporate Counsel. Let me give him control and he can take over. Go ahead (Rick).

(Richard Wolf): Thank you (Glen). I only have a couple of minutes today. And forgive me if I go fast because there's much to cover. I'm not sure I could do anything but scratch the surface but I'll give it my best.

By just starting and saying that when it comes to information management in the electronic age what we're in there's a overriding need to control, not just the information that needs to be produced when you need to produce it and finding what you need when you need to find it. But also finding ways to management down the information that you don't need because the accumulation continues. And that's maybe an overriding theme that you'll hear through my piece and perhaps that of others.

We look to minimize risks and reduce information management costs by leveraging three main areas, people, policy, and technology. You can find the best technology available but if you don't have changed management techniques and buy in from the highest ranks in the organization and then ownership through out the organization, it's unimportant as to whether the technology works if nobody is going to use it.

Put differently, you can buy technology but you shouldn't use technology to do your old processes, the old way. So there needs to be change management that goes with any kind of system.

And then lastly of course you need rules of engagement and policy. And I'm going to try to address as best I can all three today.

In managing information from a policy perspective ((inaudible)) and I'll just give real quickly how we went through it. We've created a steering committee of cross functional representatives in the organization who help provide buy in support at the corporate level. And then on a micro basis the same is done in business units throughout the company.

But the creation of policy is an important document. But it's not hard for people or for companies to create policy. The implementation and enforcement of policy is really the critical part. And procedurally to follow the same steps regularly so that it becomes muscle memory of the organization. And there's some differences in the approach we've taken. And we've looked to try to simplify our policy and procedure so it's useable and manageable for employees. And to the point where they actually don't know that it's legal that's driving something. But it's actually part of the way we do business.

One of the focus points and way to look to at information management at least from my perspective is you have at some point in time have to stop the bleeding. And that means you have to deal with legacy backlog of paper information that's accumulated in some companies over several decades. And find ways of managing that information including the electronic information that's accumulated on tape. And I don't know if any of the speakers will get to it later, but when you talk about the case in Florida with Morgan Stanley, and the accumulation of email back up tapes, there needs to be really two pieces of the program. One is to deal with the backlog or the back file. And the other one is to go post (spectively) in dealing with your information proactively going forward. Unfortunately you can't really stop one to do the other. Because when you're under pressure in a regulatory or litigation setting, you don't have the luxury of time. And as I like to say there's no mulligans in records management any more.

In terms of finding systems that work ((inaudible)) what we did and what I recommend to others is conducting a gaps analysis. And in the process of identifying which systems if any in the organization were scalable for use in terms of document management systems, we were unable to identify a system in the company that was useable in a mass way.

And they're identifying the technical that we ultimately up purchasing. We selected from among five vendors. And again the involvement of senior management in the process all along. But the investment in the technology really didn't come until after the policy and the process on the paper was in place for some time. But in terms of selecting a vendor, it's sometimes difficult to identify the most qualified when we're getting information in a new

technology area such as information management. So there's a substantial amount of testing a customization that's going to come with any new technology in this space.

And I'll just briefly try to discuss how we've dealt with at least email and other unstructured content.

When email was sent to receive in the organization ((inaudible)) and the figures are startling as to how many emails, or billions of emails are sent every day in business. The idea is to centralize as best a possible email in the organization so it's manageable by lawyers or by legal teams from one single platform or location.

So the information that passes through out servers or crosses our firewall, a copy of is placed onto a third platform or storage device. And in that place, although users will still have their inbox and their information as they are accustomed to seeing it every day. The information or the email actually sits in a single instance ((inaudible)) in other words there's no duplicates of that email in the other place, in the third place storage location. And the information that fits in there is then accessible by lawyers who can apply very basic search criteria to place a hold on the information so that the information is not destroyed or discarded. And we don't run afoul of any issues in the under the federal rules of procedure or with the rule afoul of subpoenas or obstruction of justice issues. By being able to identify and hold information centrally we can avoid those problems. But the issue ((inaudible)) and I think others will probably agree with this ((inaudible)) is that most companies and employees, especially at the senior levels save much information on hard drives in the form of PSC folds as it's known.

And unless there's restriction on the ability to save email on hard drives and migrate it off the central system, there's never an opportunity to say with certainty that all email has been held, locked down, and ultimately produce. So it's a problem that's ongoing. And it's not an easy one. And it requires much change management and buy in and support.

Centralizing a system ends my vast points. In terms of the law department is important as well. When subpoenas come into the organizations or processes served, I believe it's best to have it end up in one person or one location in the company. And that person or location or function should be the group that has capability to applied hold management to the email or to electronic information through the system or the application I just described.

So that as part of the process when notice is given of an action there's the ability to apply the electronic hold instantly or within a reasonable time. And then to past the notice or the process on to the people ultimately responsible for the litigation or inquiry. And at that point we then would go through the normal paper hold management process which is really notification to record managers across the organization that a matter is pending and that all information relevant to it should be preserved. But on the electronic side it should identify and end up to a central point in the organization.

I'm going to hand over the control now to end. And ...

(Glen Flowers): ... thank you (Richard).

(Richard Wolf): Let you go.

(Glen Flowers): Thanks very much (Richard). Our next speaker is (Ann Curshaw). She's the founder and managing principle of (Ann Curshaw) PC. She's an experienced trial attorney and consultant on electronic discovery matters. She's provided electronic discovery survey data and testimony before the Federal Civil Rules Advisory Committee. She's also an active member of the (Sedona) Conference working group. She's one of the principle drafters of navigating the vendor proposal process, best practices for the selection of electronic discovery vendors. And she's a contributing editor to the (Sedona) conference glossary free discovery and digital information management. Please take it away, (Ann). (Ann)? We may have had a technical difficulty with (Ann). (Jacqueline) could you take control of the slides please? (Jacqueline)? Sorry about this folks. We'll come back to (Ann) when she gets back on the call.

We're going to move next ...

Operator: ... I'm rebooting right now, (Glen).

(Glen Flowers): OK. Bob's having technical difficulties also.

(Nicholas) are you ready?

(Ann Curshaw): Hello? Hello? Yes, I lost connection there.

(Glen Flowers): Yes, and welcome back.

(Ann Curshaw): Hi.



Operator: (Ann) this is the operator. You want to turn down the speakers that you have on your computer.

(Ann Curshaw): They're completely down. OK. Do I have leader?

(Glen Flowers): You should have the leadership, (Ann). Thank you.

(Ann Curshaw): Yes I do. OK. Well, thank you. Sorry for the technical glitch there. Not sure what happened on the phone.

Operator: (Ann) we're going to need to deal with the audio coming through your speakers.

Go to control panel.

(Ann Curshaw): I'm there.

Operator: Then go to sounds, sounds and audio.

(Ann Curshaw): Yes.

Operator: And click mute. And click OK.

(Ann Curshaw): OK. Sorry about that. Hello, every one. Thank you (Rick) and thank you to the Association of Corporate Counsel, Xerox for giving me an opportunity to speak with you today.

(Rick) has talk about preservation issues principally. And I'm going now to going move into collection and review issues and discuss the needs for automation in the collection and the review process.

And of course the first question people ask is ((inaudible)) why do we have a need for automated collection and review? Well for large data producers current processes for reviewing documents for litigation simply costs too much and they're ineffective.

Why have we not gotten to the place where automated review and collection is routine? Well, from my work what I see is quite a bit of resistance to using new processes. People and lawyers in particular have always believed that people reading documents for at the original or initial assessment review as the gold standard. And people don't understand how technology works, or the differences between different technologies. There's also I perceive a culture clash. Lawyers move forward based on the principles of stare decisis. They are comfortable in doing again what the law has told them is the law from the past.

But technical processes are going forward business processes that require immense forward planning. And lawyers, particularly on the defense side, which is generally the large data producing side in most cases come from the reactive model. And it doesn't always work when you're trying to get new processes going.

And so people ask are there standards and solutions, moving forward on the slide, to assess the success of an automated approach?

Another piece of the resistance is ((inaudible)) does it work? How do we know it's working? Are there studies?

We are working on studies to validate technologies for search and review. I did a study a little more than a year ago. A very simple study where I had a set of 48,000 documents and three codes for relevance. And I engaged a software process provider who did very sophisticated initial relevance assessments to go through these documents and attempt to assign them to one of the three codes for relevance.

I then had six experienced document reviewers look at the same or a stratified random sample, ((inaudible)) almost half of the same set of documents for the same three codes. And both processes were working under the same relevance criteria. A very well developed writing as to what constituted a subject matter that was within the scope of relevance.

I then compared the results. And the results were startling. We found that while there was a large group of documents that the people in the software agreed were relevant. And as to those documents we assumed that they were correctly relevant. The results on a large group of documents that the software had coded as relevant that the people had not.

So went back to the same people. And this is important because people learn when they read. And people get a very strong understanding of the universe of documents they're

dealing when they review them, particularly if they do it together. And I said to the people was the software right or was the software wrong?

And I think in all but about 65 documents out some 2400 discrepancies between the software and the people, the people agreed that the software was right. That at the end of the day what we found was that the people were missing almost half of the relevant documents. The software was getting close to 98 percent of the relevant document. And when you crunched all the numbers at the end of the day what you found was that using the automated relevance tool reduced your chances of missing a relevant document by 90 percent.

Now my study was done in a private context without the benefit of an independent advisory board or academics. We are in the process of moving forward with further studies.

And these are the things that we need I believe. And what I hear from judges to gain acceptability among the legal community and the courts.

And in working on further studies we will have an independent advisory council of academics and experts in the field. And we're doing ((inaudible)) making these efforts to help large data producer's benchmark and to manage their exposure and risk.

The courts understand ((inaudible)) we know the law is that we have to be reasonable in how we go about identifying what documents might be potentially relevant.

And so the goal of the testing is to show that using these processes is at least as good as if not better than a traditional manual review. And then also perhaps they'll show that it's both cheaper and faster.

And I will now ((inaudible)) oh, I see I have one more question. Sorry. About how this is trending?

My personal view is that the trend can go only one way because the volumes are so staggering. And the costs are so high that there really is at a point now where we will have no choice. It's really building the argument for defensibility. And I think those arguments can be built. And I think it's just a matter of doing the comparison testing and the measuring so that people can accept or gain comfort in using these processes. And also understand the differences between different types of technologies or processes or how best to fit them in perhaps with a manual process and what that will do in terms benefiting your overall collection, review, and production.

(Glen Flowers): Thank you (Ann). Our next speaker is Dr. Robert Bower, known to his friends as Bob Bower, our recently retire Vice-President and Chief Technology Officer at global services. Bob's got a very interesting past as one of the founding team members of the Palo Alto Research Center. He served on the advisory groups for the National Academy of Sciences, the National Science Foundation, and the Departments of Commerce, Defense, and Homeland Security as well as (UNESCO). He's authored more than 100 technical papers and edited over 10 books in areas that include intelligent information access, user interfaces, software engineering, and support for collaborative ((inaudible)) sharing. So without further adieu, Bob Bower.

Bob Bower: Thanks, (Glen). Well, I appreciate this opportunity and I've been watching some of the other questions that are coming through. A question came after (Rick's) talk about which enterprise records management system and email archive system ((inaudible)) I'm going to be talking at a level slightly different than that to built on (Ann's) point which is that it actually ((inaudible)) there's a whole issue of policies and how do you retain things? And how do you make sure that nothing's been deleted? But then the question is ((inaudible)) how actually do you call and determine things?

And the latest trends in technology in the legal industry and risk management is around concept analysis. So I want to start by defining for you, at least the way I think about what is a concept? And the reason it doesn't matter at some level what enterprise management records management system you have or what email system or how you archive it, is because at the end of the day all of these systems deal with documents, be they email documents or word documents, or images from scanned paper documents as containers. That's what I mean by container. It's sort of the bucket of information shown on the left schematically that represents some content that is locked inside of it.

And the first step to get to what is a concept? Or what's the information that you're really going to be using in order to get higher level understanding for determining relevance and privilege has to do with breaking that into the individual content that's inside those containers, those documents, those emails.

So here's just a quick little example from an accounting manual that ((inaudible)) where a page from an accounting manual might have a procedure and a description and a date, a

task, and ((inaudible)) and some reference links. Each of those labels that I just mentioned that are shown highlighted on the right of the slide represent a descriptor of content. And the variable information about that content is what is included in the container. And the first thing that is necessary and where technology is moving, is no matter what system you're storing your documents in is to rip that out, rip all the content out of the container and store as first class citizens ((inaudible)) for example all the approvers and all the tasks and the dates that that happened. Because that's actually for the accounting procedure which you may want to actually determine what of this was approved at the right date for the right kind of thing in order to determine the litigation that's in process.

So the first thing that occurs when you try to get to what is a concept is to deal with ripping apart the content elements. But then the next element is actually ((inaudible)) that doesn't get you to a concept ((inaudible)) that just gets you to pieces and bits of information at sort of a low level of grain size, some of which is still unstructured.

And what you have to do is add context to that. Where context represents as shown on the very bottom of this slide by this little comment I've made here ((inaudible)) it's about how do attorney's , how do people, not just attorney's, but how do people who are using these documents, who are sending these emails ((inaudible)) how do they actually use these documents? What is the way they describe what's in them? What's the way they actually take action? What kind of things do they keep and they not keep? Those are not content per se. But they are in fact the context of use and the way in which people access things and how they send them around.

And understanding this ((inaudible)) creating a user interface design as shown on the right here that will enhance the ability ((inaudible)) to actually understand how those content elements are actually put together and use is of course what's going on in all legal procedures any way. I mean it's not just that people had access to a documents. It's whether they used it for good or ill. Or whether they even didn't pay attention to it.

So they're two pieces of a content analysis that are shown here on the left. Now I'm going to quickly mention because I was asked to cover what's the linguistic content analysis and statistical content analysis. There's also image content analysis which we won't be discussing today. But I just wanted to point to which has to do with the fact that there's a lot of information just on the way that documents are annotated. The way that fields appear in letterheads. All of that represents the visual kind of information that human review and discovery will be very adept at. And to which technology is moving in order to create better understanding of context in which those content elements appear.

So this is a matter of ((inaudible)) the technology that I'm talking about is trying to take things that start in fairly locked containers, either as images or as text on the left and doing detailed analysis using statistical methods, natural language messages, image analysis, give users interfaces such that it can be created in a tab manner which is what the XML stands for on the right.

Just quickly ((inaudible)) and I don't want to go too much into technical detail. It's just that you'll hear these terms when you go to buy systems. Statistical techniques for the most part are what are used to first of all do the culling. Figuring out what categories of content



are even interesting to a case. What are the cluster of documents? And how do I categorize them in order to create a taxonomy of terms that would describe the content?

So I'm not going to go through the detail here. This will be up on the web site for your review later. And of course there can be further interactions.

But shown on the right slide is just sort of a clustering technique which is strictly based on statistics. How often does a particular term occur? And how many of those documents fit within that type of cluster of information?

And this shows four clusters of documents. And the thing that's important in terms of clustering and categorizing statistically is that there's never one right answer. You'll be able to see when you look more closely on the web site that down in the lower left hand of this user interface, which is what I was going to emphasize, is a work bench that tells you that this particular document which is a patent ((inaudible)) that a patent attorney is trying categorize properly ((inaudible)) actually has a 70 percent chance of being in one category. And less than 10 percent chance in two other categories. So that's based on simply statistical probabilities. And the amount of certainty and uncertainty is of course where the human intervention occurs in the way that these concepts search tools are used today.

Finally there is ((inaudible)) I'm going to talk about linguistic, which is to understand the text itself. This is a different user interface, a different workbench. We're now on the left hand panel here. You can see the, you can see where there are lists of the taxonomies that are probably determined both from what is the standard way that you discover and you do

discovery. Plus things that are developed by clustering and statistical categorization techniques in order to taxonomy.

And then there's sample documents that are shown highlighted here in the panel in the lower left corner which represent different pieces of content for discovery that are representative of what the legal process of discovery is going after. And up in the upper corner, they're actually relationships among that content so that you can then get a measurement of independent of whether the words are singular or plural, or whether they're gerund? It takes all of the uncertainty of language away and creates a measure of how often a particular name or date or project occurred.

The thing that's important here is that in the terms of concept search which is what I was asked to mention and give you an overview on ((inaudible)) is that this kind of intelligent data mining is still very focused on human interpretation. And I think as (Ann) said which is really critical here ((inaudible)) at the end of the day, humans can only learn and balance a few criteria that they are looking for in these documents. Technologies can actually build these criteria's better in order to automate the tasks of actually finding the valuable information and providing where the relevance is. And by using these kinds of workbenches to build statistical models, one can actually create very high level conceptual understandings that can be determined and shown in the very bottom of this slide.

Where the example I'm giving is actually from a case study of a construction contract where there was a dispute over the commercial building that was built. Where the general contractor claimed that they didn't have to in fact take care of certain obligations. And so there was a model built for what is an obligation? Through understanding the way that

obligation is used in context. The way the synonyms for obligation are. But not only the synonyms. To create different types of understandings of what constitutes an obligation in this case. And these three are actual sentences that were removed from the contract automatically by the system.

Notice from the trivial that a contractor will keep an organized office. That the contractor will be responsible for taxes. To the one that are most important, ((inaudible)) taxes of course is something that every one would be able to find in a document because you could search for that. But a contractor needing to report discrepancies or mistakes on a drawing immediately was something that was signed in the contract. And that they're two procedures. One involves email and one involves voice messaging to this individual, (Matt Blankenship). That's something that you would not be able to find under any synonym for obligation. You wouldn't be able to normally characterize a discrepancy or mistake or a drawing as being something that you would search for from a terminology point of view.

So instead by being able to create models for concepts like obligations you could actually much better than that. And so at the end of the day if you were ((inaudible)) if I had to summarize where the technology is today and where it's likely to move over the next couple of years. Almost all of the companies that are providing search support are really looking at humans on the left side of a process who are creating workbenches to model discovery.

But at the end of the day what's coming out of the computers after they've created this domain model that you see here and then they call from an existing collection of documents ((inaudible)) after that's all be processed statistically and linguistically, ((inaudible)) at the end of the day there's just another collection of content that another individual has to have

the human features and foibles of actually reviewing. And I think what's most exciting is techniques that are occurring now that will really represent a change in the whole matter of how discovery as well as risk management occurs from the point of view of eliminating this human being as being the only mechanism for doing the search of this new labeled conceptual document collection. But rather to have systems that can automatically track new documents that are coming in all the time. New emails, like (Rick) said that actually can be caught before something bad happens for an internal process to do something wrong. Those are the kinds of systems that will use these technology underpinnings in ways that will really revolutionize the whole process that is becoming more and more labor intensive. And depending more and more on pieces of technology.

So with that I want to hand it off to (Nicholas Economu) who's going to speak to you about some of these issues in terms of what constitutes risks. And what does it really mean to be able to determine relevance and privilege?

(Glen Flowers): Thank you Bob. Our next speaker as Bob just indicated is (Nicholas Economu) who's the Chairman and CEO of H5 Technologies. (Nicholas) has spent the last 15 years providing advisory and consulting services to governments, government agencies, corporations, law firms, and financial services firms in areas ranging from information and intelligence monitoring to document review process design and information management. That sounds pretty complicated to me. He's also a member of the Vendor Advisory Panel to the (Sedona) Conference's Electronic Document Retention and Production Working group. And take it away (Nicholas).

(Nicholas Economu): Thank you very much. I would like to first thank ((inaudible)) for letting me participate.

I'll start with really the question that I think most of the people that we work with ((inaudible)) the general counsels office really have most trouble answering even after investing very significant amounts of resources and money in trying to manage the risks of really identifying correctly what information matters in a case or in the context of retention or out of compliance efforts.

And really the question that should, I think should be answered in the end, is really after we invested everything that we have in our systems and we've done all the work in the review. We've done, really of all the documents that we should have retained or produced or found, how many did we actually retain, produce, or find?

And that's a question that unfortunately when you look at the available academy data the answer is you're likely to have missed a lot of the documents that you should have retained, found, or produced.

And I'm showing you here the sort of philosophy of where the academic data ((inaudible)) the data that's actually study of the of cost conducted by a major law firm. And really the idea was to look at how, how well humans do and then see how technologies or new processes can do. And the adjudicator, the law firm in this case, chose a sample of documents. And they determined that within these documents after a very thorough and slow review, et cetera, et cetera. That really there were for two very well articulated issues

((inaudible)) is exactly the same protocol as a study I mentioned ((inaudible)) this is really what see in the green circle was actual the relevant set.

So no matter what you're trying , what method you're using, really the goal would be to find amongst all these documents characterized by the blue squares you see here, was to say that the goal was to find really the documents in the green circles.

And the test that was done showed that full manual review conducted under really good circumstances resulted in the following ((inaudible)) which you can see on both sides here that really the human full manual review missed in both instances in excess of 50 percent of what actually was important. And what the senior in this case litigator had determined they would have wanted to find or they would have found it if they had looked very carefully at each document themselves.

So the result of risks of course involving missing so much information. And the question is really why that happens? But the academics ((inaudible)) what I showed right now that there is an enormous amount of information typically that is missed by human review is very well supported by the broad academic evidence. So I'm not going to go through the results of the studies that you see here. This will be on the web as well.

But the academic consensus overwhelming is that no matter what search technology you use, no matter whether you have humans review that fully, your collections fully ((inaudible)) when you have large scale review to conduct really there is a lot missed.

And two interesting studies was ((inaudible)) that they were generally well known is the Churchill study of '94 mentions here the second one, which really showed ?looked at how well paralegals and attorneys would do at identifying relevant documents in a large collection. And the result was that really ((inaudible)) most of the documents were missed. (Blair and Marren) in '85 conducted another study that allowed attorneys to review a collection until the attorneys were satisfied that they had found 75 percent of the relevant documents were correct. Documents they should have found. And it turns out that in the control setting it was found that they had only identified correctly about 20 percent of the relevant documents.

So the academic consensus is really that you really achieve only very low accuracy or likely to achieve very low accuracy by using other manual reviewers or search tools.

Now why is that a problem? And most of you probably deal with that every day. But from a cost perspective there is really the notion that if you really inaccurate and you're bring in a lot junk for review with your keyword or research systems, there is an enormous cost in analyzing the relevant documents. Sometimes ((inaudible)) we conducted a number of cost benchmarks with our clients ((inaudible)) it literally runs sometimes in the 10's of millions and as a proportion of the over review effort it can run into easily 60 or 70 percent of the overall costs just to review irrelevant documents returned by your search or concept systems.

The second risk of course of low accuracy is the queue ((inaudible)) you have significant financial exposure from just the failure to retain or produce the right collections or to find the right collections if you're in the middle of litigation.

And lastly if you have a lot of junk that you are bring in ((inaudible)) documents you've not really looking at, but are returned by your search systems, you end up having tremendous delays in completing the review you need to do. And that as well has of course an impact on your ability to prosecute the case or to complete your whatever other compliance effort you're trying to complete.

And the question is really why is the performance given those risks and the stakes, why is the performance so low? And the key message I'd like to convey is that really the biggest challenges you face are not technological. The usual answer to I have a big search problem, is well let me buy a search tool. And it turns out that the fundamental problems that you're facing are not technological. They're really the ((inaudible)) I've tried to list here. It is the way we've identified them.

The one is really the knowledge transfer. Making certain that all the people who serve as your proxy if you want to identify the right documents. The humans helped by search tools really understand what the lead litigator or the senior counsel wants.

The second challenge is ((inaudible)) think of the challenge of the consistency. Even once you've found 50,000 documents to review with your keyword tools, you still need human structure to review those and make an assessment. Is it relevant or not? That is Bob Bowers point really. That in the end no matter how you search a human will make an assessment. Humans are inconsistent. That has nothing to do with the technology.

And lastly as far as we see, they're very few companies or law firms who really understand how to quality assure their processes. So once you've searched, well how do know how well



you've done. And the knowledge transfer, the inconsistency in the quality assurance which I've describe as teaching challenges and human nature challenges and processes design challenges ((inaudible)) I think are quite self evidently are not addressed by buying near search tool.

And the question is well, if you solve, if you try to address these challenges how well can you actually do? And (Ann Curshaw) mentioned in her studies which actually were the study of a system addressed these challenges. And the result is that you can achieve results of this kind. So the protocol here again is a protocol that's generally accepted in the academic consensus. And I can't go through the detail right now. But the protocol , this particular study which is not the one that I mentioned was that you can really achieve results ((inaudible)) the blue circle here is a study of H5. You can achieve results with ((inaudible)) the results are practically indistinguishable. The blue circle here from the underlying green circle I showed you earlier. And you can measure that. You can determine whether or not you were able to identify ((inaudible)) well or better the correct result set than if an expert human had reviewed the whole collection at great, great, great, length.

So the truth is that if you place technology in proper context and solve these additional non-technological challenges, you can actually achieve very significant levels of accuracy or improvements in the accuracy that you achieve.

That goes also to think ((inaudible)) the point that (Rich) started with in saying the best technology on its own does not solve your problem. That is absolutely true.

And so I went quite quickly through that because of course there are probably some questions for panelists but the message I want to leave you with is ((inaudible)) if you think about measuring and managing risks, you do not need to be an expert in search technologies. You need to know though how to assess them. And I think (Rich) as well talked about the difficulty in assessing new technologies and vendors and promises that are made. Well, I ((inaudible)) protocols exists. (Ann Curshaw) is working at the study that she mentioned I think now or design ((inaudible)) the involvement of academics that she talked about was in this direction of saying in academia there are academic protocols along with what I showed with the same diagrams that can help you design processes to assess ((inaudible)). Well use them. You don't need to be an expert in search techniques. You need to know the people who are experts in the protocols that then allow to assess the performance of your own review system, retention system whether it's driven by humans, humans and technology, technology alone.

Once you know how accurate or inaccurate your systems is overall you have very good data to progress. And you can measure and really benchmark very accurately the financial and time to completion costs that result.

Again very few ((inaudible)) it is very difficult to measure your cost and exposure until you know how good you are at finding the documents that you should retain for use or review.

And the last point I'd like to make is really think of this challenge of identifying the right documents and the right collections to retain or produce or review in litigation as a system. A fairly complex system that included IT and processes, but also information retrieval expertise. Very often it's a task that's assigned to the IT groups. And it's important piece of

it. But there is a tremendous piece of fact which has to do with information retrieval expertise and the design of protocols to assess how well you're doing.

And all of these issues together, all of these aspects of IT, IR, and process design and quality assure has to be brought together. And you do not need again to be an expert at all of those. But you need to have the ability to find the people who can design these protocols and processes for you.

So that's the end of my participation. I'll turn it over to Glen.

(Glen Flowers): Thank you (Nicholas). Well folks for the remainder of this hour we are going to attempt to answer your questions. And you can keep those coming by entering them in lower right hand corner of your screen.

Additionally you'll notice just above that lower box there's a box with links. And link number six will take you to a survey 'cause that helps us assess what we ((inaudible)) how we did in this presentation? What any suggestions you have for future presentation. And that feedback is critical to us to enable us to continue to bring these programs to you.

So why don't we go first to (Richard Wolf). There's a question from the audience on how did you go about determining which systems to get for your records management and your email archive?

(Richard Wolf): The question is about the process we followed or the product that was selected?

(Glen Flowers): Well, why don't you start by talking about the process you went through and some of the considerations in determining what you wanted to do, to choose.

(Richard Wolf): Well the process we undertook involved using a consultant or advisor with no vested interest in any of the technologies that we ultimately might purchase to help give us some level of objectivity.

We also because we're a public company but we're not a highly regulated company, we wanted to have technology that met if possible the highest standards, the Department of Defense standard, and SEC standards. But we didn't need to have all of the bells and whistles to capture ((inaudible)) although I've learned something on this call with taxonomy in capturing different information to sorting in to buckets. I know that's required in many of the broker dealer context. But for us we didn't need that.

So ultimately we're looking for a good document management system with capabilities that helped us in the email area. And ultimately we took a product that's used by Documentum call per se. Have record services for email. And really strip that down so that it had the component parts that we needed to sort the email once it was journaled which is the process of moving something out of Microsoft outlook into a storage device to sort the information there. And at the highest levels, I'll add ((inaudible)) for email the custodian base search will be done. And then the more robust search could be then addressed in the manner that Bob was describing earlier.

I'm just not sure at this point kind of what the standard is for how much searching needs to be done in producing information for discovery.

(Glen Flowers): OK, thank you. Thank you (Rick). The next question is directed to (Ann). And that is where can people learn more about the study that you have already authored? And you reference also some new studies that were in process. So maybe you could tell the folks where they could find out more.

(Ann Curshaw): Sure, sure. Of course I'm happy to take anyone's phone call or send me an email address and I can tell you all kinds of ((inaudible)) the detail about the private study I did ((inaudible)) I wrote an article for (Pike and Fischer) which I can also forward to folks. But the study that I did, because it was in a private was confidential and privilege data. And without independent validation of an oversight committee.

There's no white paper on it. I ((inaudible)) the further studies that I intend to do ((inaudible)) we'll have those criteria. And I've started a research organization called Reasonable Discovery. And I'm very fortunate to have had a fortune toy company donate a really excellent data set to touch. Because that's been one of the challenges is pulling in a data set that can be eligible for review. Because there's a number of issues there.

And we're now embarking on putting together the methodology of the study. Having assembled the ?an advisory panel of academics and experts in the field. And we are looking for support, financial support. We're looking for feedback. We're looking for data. We're very excited to get this off the ground. And if anyone wants to send me their email I'd be happy to forward you any information that could help get you involved. Should I give my email address, (Glen)? Or is ...

(Glen Flowers): I believe those are all posted in the bios which are also available on the ACC web

site. So everyone can contact. And our emails are up there. Thanks (Ann).

OK, another question I have for the whole panel ((inaudible)) depending on where different companies are in their evolution and understanding of preparing for litigation and managing their records and their documents ((inaudible)) how does one proceed in evaluating and moving forward to a more technological approach? To working with this information proactively for compliance purposes or to prepare for litigation?

Who would like to jump in on that?

(Richard Wolf): Well, this is (Rick Wolf), it might be more in my area than the others perhaps.

And I think the answer is ((inaudible)) has very little to do with technology. The use ?and I'm going to go back to email again just because it creates such volume and exposure.

If users in companies use email for archiving documents and there's no central repository for them to find their records otherwise. Then email is going to be use for that purpose. And email will also be used for saving information for personal historical reference.

Those types of behaviors and habits need to be changed as part of the process. And the technologies will do their work. But until information is centralized and behaviors are modified in the way we do business, I'm not sure that you really even scratch the surface on getting your arms around information overload.

(Glen Flowers): Yes, (Ann) could you add to that in terms of some of the processes you've gone through with the companies you've worked with in terms of moving towards more advanced methods of being prepared and responding?

(Ann Curshaw): Oh, sure. Well of course what you're going to do is very, very, specific to the particular company's processes, type of litigation. How much litigation they have? What you generally want to do of course is eliminate the ramp up, ramp down phenomenon that occurs from a reactive mode of handling litigation. Because that's very expensive and also it usually results in the classic fire drill. And fire drills in technology environments are very, very, dangerous. Because that's when you're going to make all of your mistakes.

These processes have to be very carefully planned and thought out and tested and piloted. And so, outside planning and being proactive in managing all this date and being ready for litigation or being ready for the need before the need is on the table is really key.

And if I was working with a company a company that had a constant stream of ongoing litigation. And I think mostly like products companies. They know what people want. They get sued all the time. And so I would like to generally implement a proactive method for correcting that date. Of course that requires getting caught up.

And there's ways to do that. But basically you can do things where people write emails or create documents if they're one of your significant interests' employees for litigation or ((inaudible)) litigation hold notice you would get a pop up box that says do we need to save this. And if you say yes, that particular email would go off or be blind copied to a separate server.

Not a perfect solution. An even better solution is if you've got a document management program and you would go then to your document management program, some sort of litigation hold designation through a click of a box or something of that nature. You could also do it with privilege. I find that a lot of companies don't have document management programs at least consistently throughout their departments. The companies that I work with tend to be very compartmentalized.

I don't know if that answers the question. I mean I could go on for probably an hour or so. You have cut me off.

(Glen Flowers): Well thanks (Ann). We are running near the end of our program. If you have any other quick questions in the audience, please send them in.

Take a moment again to remind you of the survey on item number six in the box to the right in the middle. Please take moment to fill that out.

And would it be ((inaudible)) panel ((inaudible)) would anyone on the panel like to add anything before we make our closing remarks? I guess we're all pretty thorough weren't we?

Male: Well, one thing I could add to (Ann)'s comment which I think is relevant sort from the technical underpinning point of view. In order not to go through fire drills all the time, one of the advantages not just thinking about having a tool to do a fire drill like a search engine or buying a tool. Is that as you build and have the ability to characterize the emails and the collections and have the taxonomies and have these models, it takes more investment in time



up front. But then the ability to adapt as issues may change during a case as new documents collections maybe discovered on other ((inaudible)) by the way enterprise records management systems are wonderful, but almost every large company I've worked with in the Fortune 500 have multiple document management systems that have to be all culled together.

You want to be able to have sort of a central model of the kinds of content and the kind of context of the case so that you can rapidly adapt it over time rather than just a tool set that someone can run in and do. So I think that from a point of view both a technical and the process point of view there are trends in this field that really allow for having a more orderly way of going about this very difficult problem that all this electronic information is causing us.

(Glen Flowers): Thank you Bob. Well, this will conclude our program. Thank you again for you attendance. And thank you panelists for participating. The slide on your screen right now gives you a few links. Certainly our web site a Xerox for corporate litigation, H5 Technologies, and (Ann Curshaw's) web site.

Additionally my and (Kim Gripps) contact information are there if we can provide any additional information or connect you with any of our speakers.

Thanks again for your attendance and please fill out that survey and have a great day.

Female: Thank you.

ASSOCIATION OF CORPORATE COUNSEL

Moderator: Glen Flowers

02-16-06

Confirmation # 1425957

Page 34

END