

Information Management Targets E-Discovery Compliance, Legal Risks While Delivering Long-term BI and Governance Benefits

Transcript of a sponsored BriefingsDirect podcast on the need to manage explosive growth of information within enterprises to head off legal risks and penalties.

Listen to the podcast. Find it on [iTunes/iPod](#) and [Podcast.com](#). Download the transcript. [Learn more](#). Sponsor: [Hewlett Packard](#).

Dana Gardner: Hi, this is [Dana Gardner](#), principal analyst at [Interarbor Solutions](#), and you're listening to BriefingsDirect.



Today, we present a sponsored podcast discussion on gaining control over [information sprawl at](#) . We'll take a look at the short-term and potentially massive savings from thwarting [legal discovery fulfillment](#) problems in advance [by controlling information](#). And we'll examine how management lifecycle approaches can bring long-term payoffs through better analytics, and regulatory compliance, while reducing the cost of data storage and archiving.

To help us better understand the perils and promise around the well-managed -- versus the haphazard -- information oversight approach, we're joined by two executives from [Hewlett-Packard \(HP\)](#).

Please join me in welcoming [Jonathan Martin](#), Vice President and General Manager for Information Management at HP. Welcome, Jonathan.

Jonathan Martin: Hi, Dana. Good to be here.

Gardner: We're also joined by [Gaynelle Jones](#), Discovery Counsel at HP. Thanks for joining, Gaynelle.

Gaynelle Jones: Hi, Dana. Good to be here.

Gardner: Let us start with you, Jonathan. We've seen quite a bit of change in the business issues around information, including risk, compliance, and oversight. Could you help set up our discussion by helping us understand how the world has changed in the past year or two, and why [information management issues have become more prominent](#)?

Martin: If you look at every organization over the last 20 years or so, fundamentally they've

built a majority of their business processes on top of information technology (IT). And, the information that flows through those IT systems tends to be one of the core assets of any business in today's economy.

Now, over the last five to 10 years, [we've become increasingly addicted to information](#), both at home and at work. At home, the idea, even three years ago, of walking around with 40,000 songs in your pocket would have been a little crazy. Today, we do it almost as habit. At work today, 85 percent of our business communications go by email.



What we've seen is a trend that's been going on for the last 15 to 20 years, and the size of it is beginning to [really impact businesses](#). This trend is that information tends to either double every year in developing countries, and tends to double about every 18 months in developed organizations. Today, we're creating more information than we have ever created before, and we tend to be using limited tools to manage that information. We're getting less business value from the information that we create.

Over the last few years, organizations have put in place what we might call an [ILM](#) strategy. Some vendors like to tell you, it stands for the "information lifecycle management." A lot of customers who have been through the pain of an ILM implementations will tell you, it's really about "investing lots more."

Throwing more capacity

That's really been the way that the problem is being solved -- just throwing more and more capacity at the problem and throwing more and more storage space, and more and more available space to store this information.

Unfortunately, in the last 18 months or so, the economy has begun to slow down, so that concept of just throwing more and more capacity at the problem is causing challenges for organizations. Today, organizations aren't looking at expanding the amount of information that's stored. They're actually looking at new ways to reduce the amount of information.



At the same time, we're going into the stage of the economic cycle where everyone is thinking beyond how to reduce cost and cash burn, and how to ensure that this never happens again. Eight years ago, we saw the economy go through a similar cycle.

As we began to bump along the bottom in 2002 and pull back into recovery in 2003, we saw the implementation of [Sarbanes-Oxley](#). Coming into 2010, both in the US and in Europe, there is going to be a new wave of a regulation that organizations are going to have to take on board about how they manage their business information.

Gardner: So the business risks have certainly gone up, but it looks like one particular type of risk is pressing, the legal risk. Gaynelle would you tell us a little bit about your background, what your role is at HP, and why the legal aspects of information management are so important?

Jones: Oh, certainly, Dana. I'm the Discovery Counsel at HP. I work with the litigation group in managing the discovery process for HP. I've been a litigation manager, as well as a prosecutor and a trial judge. Because we have black-letter law that computerized data is discoverable if relevant, and because of the enormous amount of electronic information that we are dealing with, the litigants have to be concerned with discovery, in identifying and producing it, and making sure it's admissible.



I'm charged here with developing and working with both the IT and the litigation teams around making sure that we are compliant, and that we respond quickly to identify our electronically stored information, and that we get it in a form that can be produced in the litigation.

Gardner: Now, the stakes here are quite high. Not being able to fulfill the quest for discovery and for information in its various electronics forms can come at a high penalty. Do you have any examples of how that can work?

Horror stories in the news

Jones: Oh, absolutely. There are horror stories that have been in the news in recent years around major companies such as [Morgan Stanley](#), Enron, Qualcomm and a host of others being sanctioned for not following and complying properly with the discovery rules. In Morgan Stanley and [Zubulake](#), the court issued adverse inference instructions, because data was lost. Morgan Stanley had a jury return of verdict around \$1.4 billion, and in Zubulake, the jury returned \$29-million verdict.

In each case, companies failed to properly implement litigation rules, directly pointing to their failure to properly manage their electronic information. So the sanctions and the penalties can be enormous if we don't get a hold of this and comply.

Gardner: And these types of legal requests or at least legal issues are not uncommon at large organizations. Do you have any sense of what the typical legal caseload is typically?

Jones: It depends on the enterprise. At HP, we have everything, which puts us sort of on the cutting edge to develop and really come up with some best practices. But the typical enterprise would have data around employment matters. You would be dealing with the human-resource databases.

It might have contracts, and have to deal with keeping up with the contracts, emails, and correspondence. Emails, by themselves, have tremendous issues in terms of identifying and preserving, as well as voice mail and instant messages.

At HP, we have hundreds, if not thousands, of database applications that contain our business records, our sales records, our revenue, our marketing, and so forth. So, we have dynamic databases, and all of these things can come into play in litigation, if they have relevant information.

Gardner: Jonathan, you mentioned earlier the issues around compliance and the fact that regulations are bound to creep up in a number of industries and in different countries as well. Between the regulatory issues and the legal issues, it seems like there is an awful lot of money to be saved by doing this correctly.

Martin: Absolutely. We've seen, over the last few years, organizations move from taking a very reactive approach on these kinds of issues to a more proactive or more of a robust approach. We heard from Gaynelle earlier some of the examples of the reactive approach. Organizations that are in this mode tend to take one of three ways as strategies for solving their problems.

They may trust their employees to do the right things. Obviously, everybody knows that translating policy into day-to-day behaviors for employees is not always easy. Employees, by their very nature, in the main don't tend to be particularly okay with legal or regulatory requirements of the organization.

Trusting the lawyers

The second one is that they trust their lawyers. When they run into an issue and they're subpoenaed for some information, or required to present information from an audit, they pull in a couple of bus-loads of lawyers, and get the lawyers to dig or troll through miles and miles of content to try and find the relevant information. It tends to be a very, very expensive approach to just finding information.

The third one is that they trust that their IT consultants. When the subpoena for a piece of information to come in, to find that one email in hundreds of millions of emails that the organization sends. So, as you see, we've got lots of examples in the industry of why taking a reactive approach to a litigation readiness or the ability to respond to an audit is a bad one.

Over the last two to three years, organizations began to take a more proactive approach. They're gathering the content that's most likely to be used in an audit, or that's most likely to be used in a legal matter, and consolidating that into one location. They're indexing it in the same way and setting a retention schedule for it, so that when they're required to respond to litigation or are required to respond to an audit or a governance request, all the information is in one place. They can search it very quickly.

Gardner: Of course, we've also seen a great deal of additional types of content. We're talking about all sorts of electronic. We're talking about social media, where folks are blogging, and they are using sources that are offsite and that are in someone else's servers, perhaps in a [cloud](#) environment, and we're also, I suppose, thinking about paper. How do we think about approaching this proactively, when it's such a mish-mash of content types?

Martin: At first, the problem statement may look absolutely enormous. What we see is that organizations begin to chunk that problem statement down into the areas that a subpoena is most likely to target or the area an audit is most likely to target. Typically, if you think of the things that subpoenas look for in an audit, they tend to look for business conversations.

We've already identified that 85 percent of business conversations today go through email. So, as organizations begin to take a more proactive approach to electronic discovery -- as it is called in the US, and electronic disclosure, as it is called in the rest of the world -- they really focus on email first of all.

So they're gathering all emails that the organization sends, consolidating them into one place, indexing them, setting retention schedule for them, and getting them ready, should they be required to respond.

Subsequently, the area that is very much in focus now is Microsoft [SharePoint](#) content. As you know, SharePoint is a business that is exploding for Microsoft right now. It's growing like wildfire through many organizations, and is being used in a very different way than traditional content management applications.

It's a very collaborative, free-form, and very easy-to-use set of tools. Typically, we see lots of projects spinning up within organizations. As the project begins, they will spin up a SharePoint along with it, as a repository, where they can put all the content relating to the project, as well as the meeting minutes for the projects, the collaboration, etc.

New wave of solutions

This really becomes the central point within the project team for them to collaborate. Typically, the things that are going into those are meeting minutes, statements of work, draft contracts, submissions, etc. -- by anybody's definition, true business records. We're beginning to see a shift from the auditors and the litigators, away from just focusing on business conversations on email, to begin to target this new wave of content-management solutions, particularly around SharePoint.

Gardner: I see. We're looking at structured content, unstructured content, communications, and even minutes for meetings. I want to go back to Gaynelle, if I could. We're not just talking here about crisis intervention. It seems to me that, over time, this is going to be something that will

pay back in significant ways, when it comes to managing intellectual property, protecting rights, and the use of very important information within the organization.

Jones: Absolutely. You have to be concerned with all sorts of issues and litigation, including the ones you've mentioned, as well as privacy issues with the data that you're dealing with, and other regulatory areas, issues that might impact upon the information that you have.

You have to be able to identify and manage the information and think ahead about where you're likely to have to pull it in and produce it, and make a plan for addressing these issues before you have to actually respond. When you're managing a lot of litigation, you have to respond in a quick timeframe, by law. You don't have time to then sit down and draw up your plan.

When you are doing it then, you are paying outsiders -- legal fees to the outside counsel, their associates, and so forth. This makes the process at least twice as expensive, than if you've planned ahead, strategized, and know where your information was, so that when the time comes, you could find it and preserve it and produce it.

Gardner: Jonathan, the economic payback here can be very large and impactful, because of the prevention of these discovery problem awards. Certainly, you can react more quickly to issues around security and risk, but it strikes me that there is a long-term benefit as well. The [return on investment \(ROI\)](#) isn't immediate and impactful at a crisis level, but perhaps at an analytics level. We hear similar rationale around why we should invest in [business intelligence \(BI\)](#), for example?

Martin: Today, eyeballs are very focused on information governance around risk litigation. What we're seeing, though, is that organizations that went through this shift from reactive to proactive two to three years ago have actually generated a new asset within the organization.

If you logically think through the process, as an organization, you are [taking a more proactive stance](#). You're capturing all of those emails, you're capturing content from file systems and your SharePoint systems. You're pulling sales orders. You get purchase request from your database environment. You're consolidating maybe miles and miles of paper into a digital form and bringing all of this content into one compliance archive.

This information is in one place. If you're able to add better classification of the content, a better way of a layer of meaning to the content, suddenly you have a tool in place that allows you to analyze the information in the organization, model information in the organization, and use this information to make better business decisions.

Unstructured data

Traditionally, as you've mentioned, BI is focused solely on structured content, content that sits in databases. Today, 80 percent of the information an organization creates is actually in an

unstructured form. If any of you went to business school, you'll know all about that 80-20 rule. You're supposed to focus on that 80. In BI, we tend to focus on the 20.

Organizations are finding, by going through an [e-discovery](#) initiative and by going through a more proactive approach to this, they ultimately end up with a brand-new repository in the organization that can help them make better business decisions, leveraging the majority of the content that the organization creates.

Gardner: Now that we understand the dimension of the problem and that there are significant short-term and long-term payoffs, how do you start approaching the solution? You did talk about chunking it up a little earlier. That made a great deal of sense, but is there an overarching vision of how to think about information differently that perhaps sets the stage for beginning this process?

Martin: There are probably a [couple of stages that we see organizations go through](#). The first one is just to catalog the information that you have out there. Use some kind of stored user-management technology to find where all the information resides in the organization.

An example of that might be something like the [Storage Essential Suite](#) from HP. This really allows you to identify where all the applications are in the organization, where all the content sits, where the databases are, and where all the storage arrays are. That gives you the ability to find all the information.

The second step is to de-duplicate the content in the organization. There are really two ways that we can do this. First, may be to take a huge swage of information by retiring legacy applications, through applications that fit into data center that may be required for regulatory or reporting reasons, but consume power, heating, lighting, support, service, license requirements etc.

Target those applications. Some applications you may never be able to retire. Other applications, you might have a duplication of capability within the data center. So, begin to de-duplicate the systems or retire the systems that are no longer required. Equally, get focused on de-duplicating the actual content that the organization creates.

Take an HP example, if [HP Chairman and CEO] [Mark Hurd](#) sends out an email to everybody with 2009 goals, everybody realize this is an important email. It's got nice attachments and a PowerPoint associated with it. So, everybody in the organization says, "I need to save this." Mark Hurd sends out one email, and it ends up getting saved 300,000 times.

That's an extreme example of duplication. On average, every piece of information an organization creates gets duplicated somewhere between 5 and 20 times by the time it has been backed up, sent to other people in emails, etc.

The second step, once you've discovered all of this content, is to begin to de-duplicate or cull down the amount of content. Once you've done that, the third step tends to be to take the content that's most likely to be used in a discovery exercise and put it into a system of record.

Consolidating content

There are a [series of products from HP](#), products like [HP TRIM](#) for records management, and [HP Integrated Archive Platform](#) for storing, archiving and retrieving content, that allow you to take all of these different types of content, consolidate them into one place, index them, set the retention schedule, and store them for long term preservation.

The final step, once you've got all that content in one place, is to add a layer of analytic or modeling capability to allow you to manipulate that content and respond quickly to a subpoena or an audit request.

Gardner: Gaynelle, listening to Jonathan explain the overarching vision for this, could you, as a consumer of this, help us understand, when you do this, what the net results are?

Jones: Absolutely. We've been really fortunate to be able to jump up and get first in line, shall we say, for the benefits of [these products](#). We're working right now on putting an evidence repository in place, so that we can collect information that's been identified, bring it over, and then search on it. So, you can do early electronic searches, do some of the de-duping that Jonathan has talked about, and get some early case assessment.

Our counsel can find out quickly what kind of emails we've got and get a sense of what the case is worth long before we have to collect it and turn it over to our outside vendors for processing. That's where we're moving at this point.

We think it's going to have tremendous benefit for us in terms of getting on top of our litigation early on, reducing the cost of the data that we end up sending outside for processing, and of course, saving cost across the board, because we can do so much of it through our own management systems, when they're in place. We're really anxious and excited to see how this is going to help us in our overall litigation strategy and in our cost.

Gardner: Now, of course, now a days, we can't look for much discretionary spending. Any requests for spending are highly scrutinized, but I'm curious. When it comes to this legally mandated and enforcement approach with information, where does the [PO](#) come from? Who signs on the bottom line to say that we need this? Is this an IT expenditure, a business expenditure, or a legal expenditure. Gaynelle, do you have any sense of that -- or you, Jonathan, as well?

Jones: I'll start, and Jonathan deals with it at a broader stage. Ideally, we get involved, if we don't do it ahead of time in terms of planning. This is what we're doing with our evidence

repository, and it becomes a part that will be shared across the business. If you wait until you are actually in the litigation, then it generally ends up being paid by the business, or the group that owns that litigation. So the earlier you do it, and more that it's planned, the more it's a shared expense. At least, that's the way we do it at HP.

From reactive to proactive

Martin: Absolutely. If you move from that kind of reactive to proactive approach, you commonly see the creation of what we would call "the committee." Typically, the committee is a combination of representatives from the legal side of an organization, as well as the IT side of an organization.

Typically today, these e-discovery projects don't get funded at the start of the year. They're one of those things that IT typically doesn't have a line item for. When they get subpoenaed to do something, then it suddenly become a priority, as we've already heard.

What you tend to find is that the committee gets together, looks at what the legal operating budget looks like, and where they're spending money on doing these requests, and by bringing these capability in-house, are they able to shift money from the legal operating budget to an IT budget to be able to gain some efficiency.

Just from a purely IT perspective, you can see almost immediate return by going down this path of retiring applications that are no longer required in the organization. Obviously, every application you have up and running creates a footprint, requires cooling, lighting etc. You can decommission these applications, while maintaining access to the content that they use to create. There's an immediate return for the organization.

Gardner: I'm curious about what is to come in terms of technology. We've certainly seen lot of interesting advances around consolidation and [warehousing of data](#). Again, perhaps mostly on the structured side, we've seen them around BI and analytics, but are there any activities at, say, HP labs for example, that point to the opportunity be doing yet more with this problem set?

Martin: Yes, there are probably two big areas on the horizon. The organizations that have been through the fundamentals like the capture process, the collection process, and the preservation process are beginning to think about.

The first is the scope of content that they capture. Increasingly, we're seeing more and more content move into the cloud. This is may be coming from a top-down initiative, or from a cost or capability perspective. Organizations are saying, "Maybe it's no longer cost effective for us to run an email environment internally. What we'd like to do is put that into the cloud, be able to manage email in the cloud, or have our email managed in the cloud."

Or, it may come from the grassroots, bottom up, where employees, when they come to work, are beginning to act more and more like consumers. They bring consumer-type technology with them, something like Facebook or social networking sites. They're coming to the organization to set up a project team and to set up a Facebook community, and they collaborate using that.

New implications

So we're seeing either top-down or grassroots-up content moving into the cloud. From a regulatory perspective, a governance perspective, or a legal perspective, this has new implications for the organizations. A lot of organizations are struggling a little bit on how do they respond to that.

Gardner: So, in this case, the source data might not be in your control, but you would have access to the [metadata](#) about that data, and that becomes yet another aspect of your systems of record in your index.

Martin: Yes, potentially, and how do you discover this content, how are you required to capture this content, or are they the same, legal obligations, the content that's inside your data center of this various IT [data centers](#)? How do you address applications, maybe [mashups](#), where content may be spread across 20 to 30 different data centers. It's a whole new vista of issues that are beginning to appear as content moves into the cloud.

Jones: We're seeing some of that now with situations in our litigation, where we have our third-party set managing our data. We have an obligation to make sure that that gets preserved.

Even smaller enterprises that perhaps may think that they don't have to deal with some of these issues, if they're providing services to companies like ours, will need to be able to have management or preservation programs in place, because we have to reach out. We're seeing in litigation where you have to deal with telephone company providers, the cable company providers, and other providers. So, it's not only managing your information, but getting access and preserving for litigation that information that others maybe managing.

Gardner: So, from your perspective, Gaynelle, we've already got a difficult situation that perhaps is going to become more difficult with the advent of a cloud era.

Jones: Right. The courts haven't yet addressed the cloud era, but it's going to definitely be one for which we're going to have to have a plan in place. The sooner you start being aware of it, asking the questions, and developing a strategy, the better. Once again, you're not being reactive and, hopefully, you're saving money in the process.

Gardner: I appreciate the discussion. It's been very interesting. To finish up, how do folks start to get a handle on this? Are there some steps or some places for information? Where do you begin on this journey?

Martin: Probably one of the best ways to learn is from the experience of others. We've invested quite heavily over the last year in building a community for the uses of our products, as well as the potential use of our products, to share best practices and ideas around this concept of information and governance that we've been talking about today, as well as just broader information management issues.

There is a website, www.hp.com/go/imhub. If you go there, you'll see lots of information from former users about how they're using their technology. If you're interested in going beyond education and getting an understanding of how HP might be able to help you in your environment with these kind of issues, we run something called the Information Management Transformation Experience Workshop, which is quite a mouthful.

If you search for IM Transformation Workshop on the HP site, you'll find that, and from there you'll be able to engage with us. Typically, it's a half-day workshop experience where we come in and brainstorm on what the issues might be and best practices that we might have for getting them solved. It's a kickoff to a broader engagement.

Gardner: Very good. We've been learning about the perils and promise of mismanaging, and then perhaps getting a proactive handle over, information and content and providing a governance approach, and a life cycle approach. We really appreciate the input from Jonathan Martin, Vice President and General Manager for Information Management at HP. Thank you, Jonathan.

Martin: Thanks, Dana.

Gardner: And also, Gaynelle Jones, Discovery Counsel at HP. Thanks so much, Gaynelle.

Jones: My pleasure.

Gardner: This is Dana Gardner, principal analyst at Interarbor Solutions. You have been listening to a sponsored BriefingsDirect podcast. Thanks for listening, and come back next time.

Listen to the podcast. Find it on [iTunes/iPod](#) and [Podcast.com](#). Download the transcript. [Learn more](#). Sponsor: [Hewlett Packard](#).

Transcript of a sponsored BriefingsDirect podcast on the need to manage explosive growth of information within enterprises to head off legal risks and penalties. Copyright Interarbor Solutions, LLC, 2005-2009. All rights reserved.