INTERVIEW WITH CRAIG RHINEHART

Eric Green: Hello and welcome to a new podcast series from IBM software that explores the challenges IT managers and business professionals are facing today. I'm Eric Green and I'll be talking with a range of experts to discover new perspectives, approaches and examples that can help meet these challenges and introduce you to the capabilities of smarter software from IBM. So let's get started.

Welcome back to the final half of our enterprise content management podcast with Craig Rhinehart, Director of ECM Strategy and Market Development for IBM. Craig, I believe we left off last time discussing advanced case management, so how about we move on to information life cycle governance and your thoughts on that.

Craig Rhinehart: Yeah, happy to … happy to cover. This is one of my favorite areas, because most people think that, you know, I sometimes joke, I say nobody wants to do governance, everyone sort of has to do this governance thing. But the fact of the matter is that there's now such an opportunity for cost savings in addition to the risk savings that we're seeing that organizations are really driving towards this now and have shifted their mentality from, you know, having to do it to wanting to do it. And here's why. You know, I mentioned this earlier, 90% of this information is formed digitally and 95% of it has an expiration date. The problem is that most organizations ignore the expiration date and are keeping everything forever, and that model has failed. The keep everything forever model has completely failed. Why? Because information is growing and budgets are not.

Information is going to grow 44 times between now and the year 2020. Most IT budgets are flat and many are declining. That's a lose/lose proposition. It's a math problem. Most organizations you spend most of their budget, many north of 80%, on existing projects, and they don't – today even – have the dollars available to invest in projects that could be used for business advantage. If information keeps growing, that 80% is going to continue to grow with it, and eventually all of your money will be spent on, you know, just keeping stuff around or just keeping the systems that you have operational.

So what's the answer? Well I can tell you one of the answers is information life cycle governance. Because if we understand what information we have, and we understand its lifespan, and we know

when it's supposed to go away, and we actually dispose of it in a defensible manner, so that no one calls our behavior into question. Then we can reclaim that infrastructure. Then we curb that losing math problem associated with all of this information growth. So it's a compelling opportunity. There's four areas where we're helping customers today. One is something called Smart Archive, which basically helps us in that lifespan identify information that's not being used but we still have to keep, and let's move that inactive information to a more affordable infrastructure. We don't have to keep all of the information in our most expensive infrastructure. We can save costs by moving the inactive stuff to more affordable environments. So this is done with e-mail, files, the collaborative environments, SAP, structured data systems, any form of information.

Second area is E-Discovery Management, which really helps our colleagues out in legal. When we're sued or when we suspect that we may be sued, we have an obligation to preserve information. This triggers a set of processes and costs that we need to manage with more rigor and discipline than we've ever had to in the past. Electronic discovery is the largest uncontrolled cost in corporate America today, and that's increasing, as is the frequency of litigation. So having better control of those processes, and how we manage our information when we're under threat of litigation or when under actual litigation is a great opportunity to reduce cost in legal. Especially when you consider that 70% or more of the information we have today – this is from a recent CGOC survey, is already past its retention date, which brings me to the last two areas – retention and disposal. If we know when it's supposed to go away and we can put policies on it, why the heck are we still keeping it? No reason to do that. So the, you know, the third and fourth areas where we're helping customers today is through records management and retention management where we make sure that all information has the right retention policy on it so when it is time to dispose of it, we're disposing of it in a defensible manner. It's a compelling, compelling opportunity to reduce risk, particularly the legal risks associated with this as well as cost opportunity.

An example is the Thomas Miller Group, which is an insurance organization who has reduced their storage costs by 60%, and I'm talking about e-mail storage costs. They specifically focused on the retention and archiving of e-mail and they've been able to, you know, save 60% on their storage costs, reduce their storage space. They're also, you know, archiving about 30,000 e-mails daily, and,

you know, that's not a lot. So the savings, regardless of the size of their organization – the larger you are, the more opportunity to have to save on this storage equation. And that's just one example of how information life cycle governance can save from a cost and risk perspective.

Eric Green: And this is really where the regulatory environment lives, I'm assuming too, the HIPAAs of this world and all of these other regulations, correct?

Craig Rhinehart: Well throughout that lifespan of information, there's all sorts of, you know, ways that you have to manage it, and whether it's to protect privacy or to produce it for a legal situation, if you don't have control over the lifespan then you know, of the information, it makes any of these things harder. Like HIPAA or you know, Sarbanes-Oxley, or you know – they're all punitive and a pain in the neck. So it starts with, you know, managing the information correctly in the first place, which makes all of the downstream stuff like compliance regulations, you know, just that much easier. It's the ounce of prevention is worth a pound of cure strategy.

Eric Green: Exactly true. So let's move onto content analytics, how about that?

Craig Rhinehart: Yeah, this is a pretty exciting area, and it touches on some really interesting innovation that the IBM Company did recently with the Watson project that I hope many are familiar with where we built a specialized question and answer computing system to test its mettle, if you will, or prove the technology on the Jeopardy game show. Now we're not going on the game show circuit, but what we are trying to do with that if you think about it – all of this information is unstructured to begin with, and all of the expected growth is going to be in this unstructured area, all of this social content, all of this social media.

We need better tools and better ways to understand all this unstructured stuff. There are plenty of good tools today to do structured data analytics. There's plenty of, you know, business intelligence and business analytics capabilities out there, but there's a need for a better understanding of the unstructured information. Why? Well, if you think about it, a five is always a five, in structured data terms. We know what it means, we don't have to think about it, we know it's half the value of ten. It's completely defined, it's unambiguous. Unstructured information is a completely different problem. If I use the term major, some of

you will think of a musical reference, some of you thinking will think of a military title, and some of you may think of it as an adjective, you know, that's a major improvement we made on our house, as an example. My point being is, unstructured is much more about the conversation, it's about our intuition, because as human beings, we speak and communicate in riddles, in slang, in pop culture references, in idioms, acronyms, abbreviations, and on and on and on. And computers have a real difficulty understanding all of this unstructured information.

So that's why IBM has, you know, invested in and has made this great breakthrough with Watson. And what's really understanding from an ECM perspective, this fifth and last area, is part of the technology that's inside of Watson is available today in the form of our IBM Content Analytics product. This is a capability that we're marketing and selling today for customers who have all of this unstructured information and they want a better understanding of what's in that unstructured information. In simple terms, you could think of this as BI for content or business intelligence for content. If you think about a call center where, you know, people are taking notes about a conversation and there are recorded conversations and you want to understand what the voice of the customer is. Well, the real nuggets are found in the unstructured call notes, or they're found in what a customer says, or they're found in what a customer writes on social media, on the web. That's that part we don't understand and need to understand.

If you think about in crime fighting, we have police departments who are looking through thousands of investigative reports or inquiries or written up and looking for trends and patterns and correlations so that they can identify patterns of crime and detect, you know, where thing are likely to happen. Think about insurance, you know, lots of unstructured information in a claims process. Let's identify the fraudulent claims as they're coming through the process before they're paid out, not after they're paid out.

So this IBM Content Analytics product in this fifth area is a very exciting area because now we're really able to unlock the value of all this content that's been trapped. We can now bring insight to this stuff. And I mentioned a call center earlier. We have a customer, a large telecommunications customer in Japan, who was one of our earlier customers on this several years ago. They have reduced their customer call hand offs by 92%, which is an enormous – call transfer, rather, which is an enormous cost savings

for them. But more importantly, what they've learned through these many interactions in customers across all these different channels is they've learned what their customers want from them. They've introduced new packaging plans, new pricing plans. They introduced an entire new customer loyalty points program, all based on the information that they were able to insight, if you will, that they were glean from this content analytics capability. So for me personally, I'm pretty excited about this area, because not only is the Watson stuff neat, and that's a clear direction for where we're going, but customers are going to be able to apply that really advanced analytics technology today on all of this, you know, unstructured information, that same unstructured information that makes up 4/5 of everything that they have in their enterprises today.

Eric Green: Well that is a great amount of information for our audience, and all the time we have for today, but Craig thanks so much for joining us.

Craig Rhinehart: Well, you're welcome, Eric. I hope I didn't go too fast or maybe provide too much information, but you know, it's a big topic, it's an exciting topic. I mean after all, it does represent, you know, 4/5 of this stuff, and you know, we're pretty excited about what we're doing in this space. We're even more excited about what our customers are doing with our solutions.

Eric Green: Thanks for listening to this two part podcast on enterprise content management.

Eric Green: Thanks for listening. Please do visit IBM.com/software to connect with our experts, continue the conversation, and to learn more about smarter software from IBM. Let's build a smarter planet.