10 Things

Scaling that mountain of data: 10 questions on records management with Kon Leong

By Jeff Cerny
February 28, 2011, 8:22 AM PST

Takeaway: Kon Leong, CEO of ZL Technologies, shares his insights on the biggest changes and trends in the rapidly escalating area of records management.

When you have 10 files, you have data. When you have a thousand files in different formats with dates, authorizations, priorities, and so on, you have data about your data. Creating new ways to keep it all sorted out and managed will yield data about the data about your data. Et cetera ad infinitum.

Managing metadata has been Kon Leong's consuming passion for more than a decade now. Kon has been successful at the helm of his data archival company, ZL Technologies, which is not surprising considering his accomplished background in computer engineering and finance. He earned an MBA with distinction from the Wharton School and received an undergraduate degree in computer science from Concordia (Loyola) University, after starting out with a year at the Indian Institute of Technology.

He spent eight years in various IT engineering and management positions at Burroughs, Philips, and Union Bank. He leveraged his finance experience at the General Motors Treasurer's Office in New York City, where he managed GM's venture capital investments in high tech. From there, he moved on to become first vice president of mergers and acquisitions at Deutsche Morgan Grenfell. In his first independent venture, he became co-founder and president of GigaLabs, a vendor of high speed networking switches.

Now his latest venture into enterprise software with ZL Technologies has streamlined things by not involving any external capital. ZL's customers include UBS, Wachovia, and Wells Fargo, to name a few. Kon likes to describe what his company does as the "manhandling of unstructured content."

Note: This article is also available as a PDF download.

1. Jeff: Records management is an area that seems to grow more complex all the time, developing from a glorified electronic filing cabinet into a process of overseeing the "lifecycle of a record." Is that accelerating complexity an accurate perception?

Kon: You're actually understating the trend. At first, you might be thinking it involves management of email and data files. But think of all the files you find in a large enterprise: XML files, SharePoint, OCS, fax data, legacy data, and so on. Even so, all of those others combined are a fraction of the amount that is email. That's where the real challenge lies. Most people underestimate the scalability required to manage email, and not by 50% or 100%, but by about 100- to 200-fold, typically, and that number is accelerating. Wherever you set the goalpost, it's going to come up short before very long. It's one of those things where there is what could be called a series of continuing afterthoughts. People are inclined to say, "Let's add on this feature or that functionality." There is no end to it — e-discovery, records, you name it.

2. Jeff: Do you see the policies that are in place around records management being developed more by companies internally or imposed from the outside?

Kon: That is an area where the emphasis has been changing, in my opinion. For 20 years, it has been the end users' judgment that tagged the records and made decisions regarding how records are maintained. Now the courts are weighing in and saying that may not be the best policy. Recently, there have been some major cases where end users making decisions about what is relevant have come under serious scrutiny. First, they don't have the legal training to assess what information is relevant to a case, and second, it's like asking the fox to guard the chickens. So the consensus is becoming that if end-user judgment isn't sufficient, you need to develop and require another way to classify using a more automated best practice approach and then have a manual override when it is necessary.

3. Jeff: As startup companies establish their own records-management policies with intentionality toward the current environment, do they set the pace for the legacy companies that may need to change their policies?

Kon: No, I don't think the smaller companies are leading the charge here at all. The cutting edge is really with the large adopters among enterprise businesses. They have the biggest headaches and the most complex set of requirements. The technology needs in the category of larger, more established companies are truly remarkable and very demanding. They are the ones who are setting the pace. Small companies don't really have the same kind of burden in this area.

4. Jeff: Which part/aspect of the records management lifecycle process do you see undergoing the greatest change right now?

Kon: For the digital track, all of it is really virgin territory. The sheer scale of it changes everything. There are some important changes and developments going on at each stage of the lifecycle, but I believe the last step is the one that poses the greatest challenge. It's one of the dirty little secrets right now and hasn't gotten sufficient scrutiny. At the scale of the Fortune 500, because of the complexity and the requirements, I see very few solutions out there being capable of actually eliminating a document. To go into that mountain of data on a regular basis and destroy the documents that have expired is something not many vendors can actually do. I don't think this area has been questioned sufficiently.

When you have billions of documents, you need a way to identify a document, including the cc's, bcc's and group lists. Ideally, you would have one document, and then the duplicates are represented by five to 10 pointers. Here's one that is pointing to a legal document, and another is pointing to an environmental issue. Those can't be deleted. And this daily burdensome cycle of selecting each document for purging, checking all its pointers and related retention periods, may be repeated for millions of documents a day, scattered across billions of documents in the enterprise. It's a herculean task, but it's not impossible. It just needs to be carefully architected.

5. Jeff: We see an increase in routinely created documents like HIPAA acknowledgments every time we visit a doctor. When does it hit a critical point?

Kon: As you know, the storage part of the equation continues to become less expensive. The hidden cost, which is becoming more significant, is the management and tracking of the content. For example, let's just think of the smaller example of a laptop. Most people could not guess how many files are on their own laptop. It used to be that ten or twenty thousand was a lot. Now it is routinely in the hundreds of thousands. Here's the difference: With our own personal computers, we're used to thinking of it in terms of megabytes or gigabytes. But when it comes to tracking, what matters is number of documents.

6. Jeff: In the area of security, what is the next big development you see?

Kon: Security is a many-faceted area — data threats, authentication, authorization. Within the area of authentication, which is a key area, is the topic of electronic signatures. One aspect of security that has not been addressed is a universal certificate of authority. I see this being very important but I don't believe enough people see the value for it to happen in the near future. I'd like to see more emphasis on it, but right now there are some hurdles that seem insurmountable. One group wants their certificate of authority, another wants theirs, and the state wants another. It's going to be difficult to agree on one universal authority, and most people don't want to see it be controlled by Big Brother.

7. Jeff: What is your opinion of the recent Department of Justice proposal to require businesses to retain more customer information and mandating Internet data retention by ISPs for its potential use in criminal investigations?

Kon: I have mixed feelings about that. There is already much more going on in that area than most people are aware of. You hear about some things only when you are confronted, and then it's a bit of a shock.

8. Jeff: Should we be comfortable with the standard level of security for doing online transactions?

Kon: It's a question of the size of the risk in relation to the severity. I am concerned with the volume and frequency of business that is taking place online. Online transactions are directly connected with your security and privacy and the whole gestalt of who you are. The main thing to keep in mind about making transactions online is that they relate to a collection of independent factors about you as a person. In the same innocuous way, each step of Facebook or Google is getting a little more of the information about you together in one place, but with a transaction, it includes even more sensitive data. Once the pieces are out there and confirmed and cross-checked, you really don't have any way to have privacy anymore.

9. Jeff: What new developments in the marketplace over the next couple of years do you see affecting records, compliance, and risk?

Kon: First is the expansion of scope, and then convergence and integration. More formats and functionality are going to be added, hopefully de-duplicated into one data instance, rather than putting up with all these clones proliferating. Policies are going to be more granular; you can expect to check and enforce them all from one place. Real convergence will depend to some degree on reaching an agreement on standards and finding consensus, but the technology will be a lot more challenging, primarily due to the sheer scale increase. For an organization of 25,000 people with a retention timeframe of seven years, for example, you're going to need the capacity to manage 4.5 billion documents. That's a number like the national debt; it's a lot, but I don't really know what it means. So for comparison, a similar size collection of documents at 4.3 billion would be Google at the time of its IPO in 2004. And by the way, Google doesn't have enterprise requirements like tracking every document and not losing any of them.

10. Jeff: Speaking of Google, it initiated a rewards program offering between $500 and $3,000 for finding security vulnerabilities last year, and there are hackers who have made up to $20,000 already. What do you think of this kind of open source approach to security?

Kon: It's a worthwhile experiment. I could give you many examples of the proprietary method of solving problems not working very well so far. Google seems to have success with a lot of its experiments. The open source method is certainly worth a try, and we'll have to see how well it works over time.

Kon Leong is the CEO of ZL Technologies and is based in San Jose, CA. When he's not busy running the company, he enjoys roaming around the globe, learning new languages and cultures. A lifelong goal is to start a knowledge compendium tentatively titled Integration of the Humanities*, to be completed over a decade or two with the collaboration of hundreds of experts across related disciplines. You can get in touch with him at kleong@zlti.com or follow him @zltechnologies.*

You May Also Like

- Best practices for software license managementTechRepublic
- 10 IT risk management issues that are often overlookedTechRepublic
- Windows 7 finally more popular than XP, OS X advances past VistaDigital Trends
- DNA is the new cloudBizJournals.com

about these links