

Information Governance & the Cloud

Take advantage of the move to the Cloud by integrating Information Governance.

The proliferation of the Cloud is not relenting. Your firm will (and should) make the move. Since it's not a question of "if," but "when," best to be prepared. Integrating Information Governance technology as part of your cloud transition will provide both quick wins (e.g., defensible disposition) and long term strategic advantages (e.g., data security).

The Inevitability of Cloud

The storm is no longer brewing in the distance; the market for shared IT infrastructure and services delivered over a network (a.k.a. Cloud) is established and growing fast. The old guard of Dell, EMC, HP, Oracle, Cisco, and other vendors of expensive infrastructure for private data centers are giving way to new players like Amazon, Google, and Microsoft, which are growing rapidly serving businesses with the much needed cloud utility. Amazon is by far the largest player; in 2014, [Amazon Web Services](#) (AWS) earned \$4.6 billion in revenue, an increase of 49% over the previous year, and is on track for \$6.23 billion in 2015.¹

Besides obvious value propositions like efficiency and agility (in terms of pricing, scalability, and redundancy), cloud providers offer simple access to innovative technologies, easily administered over their network. Given the rapid commoditization of storage (memory)

and compute (CPU) resources, it is not surprising that application-as-a-service offerings have emerged as a key competitive advantage. Amazon, in particular, has built a successful business model around leveraging its EC2 platform to offer a variety of integrated, cloud-based applications, as well as various other applications delivered through their AWS Marketplace.

The Inherent and Comparative Security of Cloud

The industry's growth is evidence of many organizations embracing the move, however, some remain concerned about data ownership/rights, accessibility, and management in the cloud, as well as its (perceived lack of) data security. While it is true that cloud networks are vulnerable to security threats, it is also important to remember that traditional networks are similarly insecure. The difference, however, is that large and established cloud providers like AWS have the scale and resources to offer true security best practices. Ensuring security for their most security conscious clients means that other customers benefit as much from those economies of scale as they do in paying for storage and processing.

While some organizations can provide the same high level of security because they are built with almost unlimited budgets (e.g., those of large financial institutions) the vast majority of organizations cannot. Small- and medium-sized enterprises (SMEs), state and local government agencies, and many other organizations simply lack the budget and staff to match the rapidly evolving landscape of cyber threats. As attackers evolve from individuals to large groups backed by governments and organized crime, it will become increasingly difficult for smaller organizations to keep pace without leveraging the economies of scale offered by the cloud. Thus, cloud adoption for the vast majority of organizations will be a boon to security. Furthermore, if a breach does occur, agreements with cloud services providers may offer some contractual protections.

The Rise of Information Governance Technology

Concurrent with the growth of cloud, Information Governance (IG) technology arose to enable more efficient, automated management of unstructured enterprise data. IG tools provide direct visibility into and control over information in place, and significantly enhance e-Discovery, records retention, compliance, and other programs aimed at regulatory and legal risk mitigation. However compelling the value proposition for IG is (particularly when faced with litigation and the burden of e-Discovery), proactive preparedness and improved response capabilities have not been the driver for wholesale IG technology adoption that the [Electronic Discovery Reference Model](#) (EDRM) or the [Information](#)

[Governance Reference Model](#) (IGRM) would seem to suggest.

Going forward, growing concerns over data security and privacy (newly included in the IGRM), coupled with the simplicity of cloud-based IG, will be the tipping point for widespread adoption of IG tools and practices. It is already irresponsible for leaders of organizations entrusted with sensitive data not to be doing all that is possible to protect that data. As true cloud-based Information Governance tools emerge, adoption of IG to identify and secure sensitive data will quickly become as standard as network monitoring and anti-virus tools.

Information Governance through and within the Cloud

In response to a trend towards cloud-based applications and cloud-delivered software, [Rational Enterprise](#) has begun an effort to offer its Information Governance and Review products through the AWS Marketplace. Hosted review products are not new, but a review product delivered through a large cloud provider like Amazon is, and as such will bring a level of commoditization to the industry beyond that brought by the Relativity reseller program. Moreover, an IG solution spanning on-premises and cloud data, delivered through the cloud, is a first.

Rational Enterprise is also heavily focused on managing data that is stored in the cloud. In addition to existing endpoint support for Google email and Docs, Rational has recently announced support for Office 365 (including Exchange, SharePoint, and OneDrive). Rational

will soon add endpoint support for all three tiers of Amazon's S3 storage (i.e., Standard, Infrequent Access, and Glacier), as well.

By integrating Information Governance directly at the AWS storage layer (i.e., Information-Governance-as-a Service (IGaaS)), Rational will transform information management in the cloud. It will soon be possible to add the full range of IG and downstream discovery and review capabilities to any unstructured cloud data with a single click. Capabilities including search; analytics; machine-learning-based classification; automated, policy-based retention, disposition, and preservation; early case assessment (ECA), in-place legal hold, notification, collection, review, redaction, and production – all become built-in functionality for unstructured EC2 data. Such a comprehensive IG solution for cloud data will certainly allay concerns about cloud services not offering adequate data management and security.

Information Governance: Enabling Migration to the Cloud through Legacy Data Cleanup

While cloud storage Information Governance tools provide value once data has been moved to the cloud, those same IG tools can also be used on premises to facilitate a strategic foray into the cloud by ensuring a more efficient and cost effective cloud migration. IG tools can alleviate the costs and risks associated with the massive over-retention problem facing most organizations; when deployed for legacy data cleanup, IG decreases the universe of vulnerable data on insecure storage and

simultaneously protects the most important data.

Wholesale migration of massive volumes of low-value (e.g., redundant, outdated or trivial (ROT)) data along with unknown volumes of high-risk and sensitive data (e.g., personally identifiable information (PII), protected health information (PHI), intellectual property (IP), etc.) to the cloud increases costs, does little to mitigate risks, and wastes an ideal opportunity to focus on cleaning up unstructured data stores. Indeed, Gartner's 2015 research note entitled "Organizations Will Need to Tackle Three Challenges to Curb Unstructured Data Glut and Neglect" states:

Across the enterprise, employees are blindly building a bottomless lake of "dark data," and, in many cases, a corporate mantra of "save everything, just in case" is encouraging the behavior.

... No one really knows the true scope of enterprise data glut. This is because the IT organization has, until recently, been in the dark. It has been restricted to storage resource management (SRM) and search tools, which are often not deployed and provide little to no functionality for determining if data – in particular, unstructured data – has any real business value or if it is sheer waste.

So what do we do? And, perhaps more importantly: given dropping storage costs, does uncontrolled data growth even matter?

Uncontrolled data growth does matter. Client inquiries suggest that, for many organizations, around 30% of data is redundant, outdated or trivial (ROT). Inquiries also suggest that around 50% of data has an indeterminate value, while the remaining data is mission-critical.

... Keeping everything not only can lead to extremely costly and damaging issues of noncompliance, but also creates a bigger pool of sensitive and personally identifiable information (PII), vulnerable to improper access.²

Just as you would have a garage sale before moving to a new house, so should you leverage an Information Governance tool for defensible disposition as a core component of your cloud migration strategy. Once a target data source has been selected, [Rational Governance](#) (RG) can index the text and metadata of every document existing on that store. In conjunction with [Rational Analytics](#)' (RA) machine-learning-based automated classification tool, RG can identify both low-value and high-risk data. RG's Policy Manager can then automatically enforce the lifespan and location of each piece of content. In this way, low-value data can be defensibly destroyed, while high-risk data can be migrated to defined high-security cloud storage locations; inactive records requiring low cost long-term storage can be migrated to AWS's Glacier storage, while active data can be moved to Standard S3. Rational Governance will continue to manage the lifecycle of all unstructured data once it is moved to the cloud.

A comprehensive Defensible Disposition and Legacy Data Analysis Program moves organizations towards the cloud:

- ✓ *Target high-impact legacy data systems*
- ✓ *Employ advanced analytics and machine learning to categorize, understand, and cull legacy data*
- ✓ *Establish retention, hold, and disposition policies consistent with regulatory and business needs*
- ✓ *Enforce policies, defensibly deleting unnecessary data and migrating retained data to cloud*
- ✓ *Govern retained data in the cloud*

Information Governance: The Convergence of Security & Cloud

Even if all backend organizational data has been moved to the cloud, a substantial volume likely remains on local end-points such as desktops, laptops, tablets, and smart phones. Easily lost, stolen, or signed on to insecure networks, these devices remain at the mercy of inadvertent or devious actions taken by employees. Moreover, data residing on these devices is some of the most prone to human error, as sensitive data can be mistakenly emailed to the wrong recipients, and viruses/malware can compromise the system, device, and attached network. An Information Governance platform that bridges cloud and on-premises data can dramatically enhance security on these vulnerable data stores by providing control of over data residing either in the cloud or on end-point systems.

Similar to the legacy data cleanup use case, Rational Governance allows an organization to identify and remove sensitive and high-risk content existing on insecure data stores continuously. With RG running on Windows or Mac PCs, the text and metadata of every document, including location information, is indexed, tracked, and made available for

analysis in real time. Sophisticated analytical tools can identify any document containing sensitive information; RG can then enforce explicit control over those documents in place, and move them to a more secure (cloud) repository, where it will continue management. Protecting the enterprise's most sensitive data in this way is a quick win in the move to cloud.

Content-centric data visibility and classification of on-premises end points significantly improves baseline security protections and supplements cloud security.

- ✓ *Track and control data at its source, helping to mitigate inevitable human error*
- ✓ *Use document classification to inform existing security systems to enforce smarter rules*
- ✓ *Search indexed content on business devices so that the enterprise can quickly assess risk if a device is lost or stolen*
- ✓ *Proactively remove the most sensitive information from devices that are not properly secure to house it and store it in the cloud*

End-to-end IG running in the cloud, controlling both cloud and local data, will also dramatically improve data security during e-Discovery. Law firms, consultancies, and other litigation service providers that deliver litigation repositories over their own (hosted) "cloud" service often have vastly inadequate security protections and processes requiring numerous handoffs between staff and systems during each phase. Unfortunately, e-Discovery is where data security risk is at its greatest, as litigation data often includes some of the most sensitive information an enterprise owns. However, with Rational Governance and **Rational Review** (RR) running in the cloud and managing both local and cloud-based data, an organization can accomplish seamless end-to-end discovery all within a single connected

environment, greatly diminishing security vulnerability.

Conclusions

Information Governance and Cloud are symbiotic. IG is a pathway to the cloud, a way to improve data management within it, and a bridge between the cloud and on premises data. The cloud benefits IG as its most dynamic delivery method. Cloud-based Information-Governance-as-a-service will allow organizations to adopt IG with storage and processing economies of scale and with little to no administrative costs. And finally, both IG and cloud programs independently, but more so collectively, improve enterprise data security. Adopting one program without the other would be to the detriment of both.

¹ Metz, Cade. "Amazon Reveals Just How Huge the Cloud Is for Its Business." *Wired.com*. Conde Nast Digital, 23 Apr. 2015. Web. 13 Nov. 2015.

² Gartner, Inc., Organizations Will Need to Tackle Three Challenges to Curb Unstructured Data Glut and Neglect, 2015, Alan Dayley, Debra Logan, June 17, 2015, <http://www.gartner.com/document/3077117>. *Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.*