IBM Extends ILG Suite and Big Data Governance with StoredIQ Acquisition
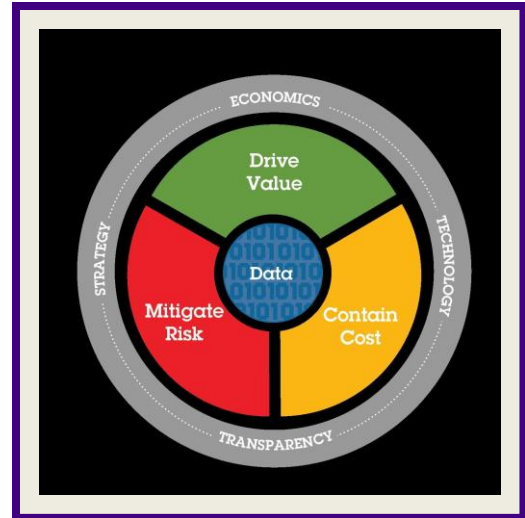
Principal Author
Gary MacFadden
Founder, CEO
Parity Research, LLC

IBM Extends ILG Suite with StoredIQ

On December 19th, 2012, IBM announced its intention to acquire unstructured data management solution provider StoredIQ. (The acquisition was completed on February 7th 2013).[1] This should come as welcome news to litigation support and compliance professionals, line of business execs looking to derive value from large volumes of organizational data and IT managers who are increasingly overwhelmed by a deluge of data, the bulk of which is duplicative, inaccessible, irrelevant, uncategorized and unprotected.

According to Deidre Paknad, Vice President of Information Lifecycle Governance at IBM, "CIOs and general counsels are overwhelmed by volumes of information that exceed their budgets and their capacity to meet legal requirements. With this acquisition, IBM adds to its unique strengths as a provider to help CIOs and attorneys rapidly drive out excess information cost and mitigate legal risks, while improving information utility for the business."

Over the last year or so, the synergy between IBM ILG and StoredIQ has strengthened, largely due to partnering on major client engagements, particularly in highly regulated and litigation heavy industries such as banking and financial services. As Paknad noted on a call with IT analysts, much of the StoredIQ functionality is complementary to IBM's existing ILG solution set, which includes PSS Systems, acquired by IBM in 2010 when Paknad was its CEO, as well as archiving, content collection, ediscovery, policy and process management, records retention and storage management capabilities.[2]
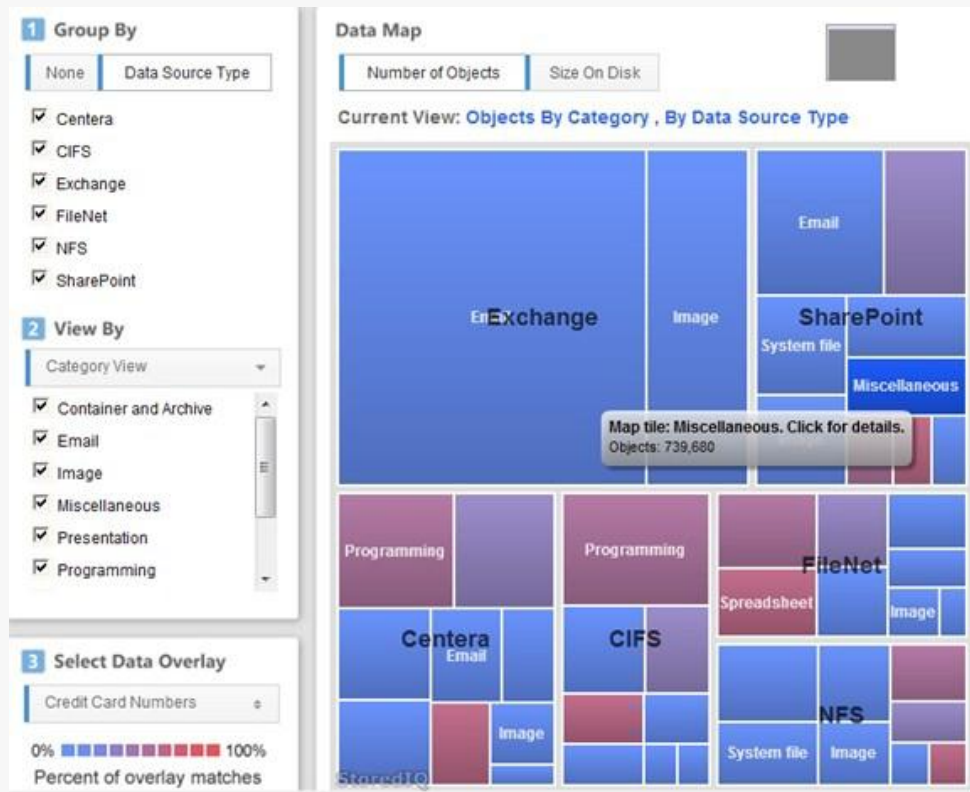
Manage Unstructured Big Data Where It Lives

One of the key differentiators for StoredIQ has been its philosophy of leaving unstructured data such as documents, emails, images, and other text-based files where they reside rather than moving data to a central repository in order to index and manage it. For large organizations trying to manage petabytes of unstructured digital content contained in hundreds or perhaps even thousands of local and geographically dispersed locations across their enterprise, attempting to consolidate data to a few physical or logical locations is time consuming, expensive and often counter productive – assuming data aggregation is even a viable option.

While content or data federation to link disparate repositories is not a new or unique concept, the StoredIQ approach has proven to be relatively straightforward and quick to implement – completion of the task can be measured in weeks instead of months or years. This past summer, StoredIQ launched DataIQ,[3] a web-enabled visualization tool that leverages the StoredIQ Platform. DataIQ gives organizations the ability to view very large data sets, then identify what data is relevant to different business stakeholders. Users are able to more easily analyze data before deciding how to act appropriately, for instance: place legal holds on documents, archive emails, move content into a records management repository, deduplicate or delete content according to a company or an industry policy.

The DataIQ application centralizes access to key metadata elements. However, data stays behind the customer's firewall or can be accessed if it resides in a cloud archive or repository.

StoredIQ's grid-like architecture allows additional data servers to be deployed across the enterprise but is centrally managed through a "single pane of glass." StoredIQ's version 7, also released earlier this year, includes an open API that enables integration to IBM products along with many other content management solutions in the market, including SharePoint.

DataIQ Visualizes Unstructured Data

StoredIQ's platform provides distributed indexing, searching (with a natural language processing (NLP) capability), tagging, classification and policy execution, allowing for all data servers to be used in a distributed, massively parallel fashion. Each StoredIQ data server is isolated so that users are not impacted by other ongoing activities within the distributed solution. With over 250 out-of-the-box attributes available for capturing and acting on unstructured data, StoredIQ offers customers perhaps the largest set of globally federated metadata for creating reports and metrics across the enterprise.

Defensible Disposal

One of the more salient solutions coming from IBM's ILG group this year has been the Defensible Disposal application set.[4] The concept is simple: Help organizations reduce storage costs and legal expenses while reducing risk. As defined on IBM's website, defensible disposal "enables IT, legal, records and information management (RIM), privacy, and the business to work in a shared transparency model to understand the obligations and value for information and automatically syndicate policies to systems with custody of the information."

> *The concept is simple: Help organizations reduce storage costs and legal expenses while reducing risk.*

IBM internal figures suggest that as much as 68% of enterprise data has no current business value. Consisting of both structured and unstructured data, some percentage of that amount can be defensibly disposed of, and 80% or more of all data in the enterprise is typically unstructured. "Together, IBM and StoredIQ can empower organizations to more efficiently use and govern their unstructured data to increase its value and eliminate unnecessary cost and risk," said Phil Myers, StoredIQ's CEO.

Paknad added that "StoredIQ extends IBM's ability to reach, govern, discover and dispose of data more broadly across the enterprise." Reaching beyond enterprise storage, cloud repositories and distributed file shares, StoredIQ can also "discover" files and documents on desktops and "crack" image files (utilizing OCR/ICR components), such as TIFF or unsearchable PDFs that may contain sensitive company or customer information, allowing stakeholders to act on that data as appropriate.

Enterprise-Ready Apps with Additional Benefits

The appeal of IBM's Defensible Disposal solution is undeniable – especially for enterprise legal teams tasked with protecting their organization from litigation risk and IT executives who have to support them. Paknad and the ILG group have smartly focused most of their efforts on these two constituencies and their pain points because that's where highly regulated industries, including banking, energy, financial services, healthcare, insurance, and pharmaceuticals, are spending billions. As a result, IBM ILG customers are also collectively saving billions.

Meanwhile, Myers, who joined StoredIQ as CEO in 2010, and his team undertook a comprehensive rewrite of their solution, including a new, more adaptive and massively scalable architecture, as well as a new branding and go-to-market strategy encompassing information management and governance, ediscovery support and Big Data enablement.

Expected benefits of implementing ILG solutions include:

- Faster access to critical information
- Higher probability that all relevant data is accessed
- Lower probability that data resides in unsecured locations
- Reduction of storage costs and related expenses (power, cooling, admin)
- Information sharing across business units and departments
- Access to and aggregation of both unstructured and structured data
- Lowered cost of litigation and reduced risk

The challenges that data growth and Big Data pose for organizations (IT industry consensus estimates corporate data doubling every 15 to 18 months) is a huge opportunity for ILG-related vendors. A well defined and implemented information governance program and strategy includes cultural change, process improvement, policy management, enabling technology and services. Vendors such as IBM have the ability to support this transformation beyond simply implementing a set of applications.

> **The biggest opportunity for extending ILG's value for both enterprises and vendors is in the realm of analytics. You can't analyze what you can't find.**

Thus far, legal and IT have the most glaring fires to put out, accompanied by the biggest operational budgets for ILG-related solutions and services. However, several positive, perhaps unintended benefits can be realized by organizations that implement ILG and have the vision to exploit their investment beyond its initial purpose. The biggest opportunity for extending ILG's

value for both enterprises and vendors is in the realm of analytics. You can't analyze what you can't find. ILG promises the ability to find and connect, contextually, with data wherever it resides, regardless of the format.

Unexpected benefits of implementing ILG solutions include:

- Enterprise-wide, analytics-readiness
- Communication between data stakeholders across the enterprise
- Opportunities for additional savings beyond IT, legal, RIM, security
- Adding value to customer service, finance, marketing, product dev, sales efforts
- Reducing backup windows and need for additional compute resources

Competition

IBM and its closest competitors in the ILG and adjacent spaces, including EMC, HP and Symantec, have all invested billions to round out their content, data and information management capabilities. Not all of those investments have appeared to pay off.

EMC, subsequent to its purchases of Documentum, Captiva and Kazeon, has created the Information Intelligence Group, which also includes its rebranding of its SourceOne email archive, acquired in the Legato acquisition, as the linchpin of its "Intelligent Information Governance" offering. While case management, content management, ediscovery, email archiving, information management, intelligent capture and storage management are all critical components of an information governance strategy, EMC has struggled to demonstrate to the market the viability of these solutions as an integrated, enterprise-ready suite. EMC also lacks a viable search engine and has had little success growing its Documentum and SourceOne market share.

Unfortunately for HP, the most talked-about aspect of its move to acquire Autonomy last year is the inflated $10 billion price tag it paid for the company and its recent decision to pursue criminal charges against Autonomy's former management, all of whom are long gone from HP. Autonomy's "meaning-based computing" IDOL search engine and a relatively large install base were justifications for the purchase which, on paper, would help HP move to a more solutions-based company. Autonomy acquired many companies over the previous 10 years, including WCM vendor Interwoven, the assets of Iron Mountain Digital and Zantaz, leaving them with several ediscovery and archive platforms. Autonomy had yet to solve existing integration problems prior to the acquisition.

Earlier this year, Symantec announced its Intelligent Information Governance solution that "allows organizations to bridge the gap between business, legal and IT, reduce their risks and costs, and empower employees to work freely in a connected world." This announcement followed its acquisition of Clearwell, an ediscovery appliance that had already been integrated at the behest of customers utilizing Symantec's industry-leading message archive, Enterprise Vault, acquired in 2004 when it purchased Veritas for $13.5 billion. Last year, Symantec also purchased cloud-based archiving solution LiveOffice. In addition, Symantec offers data protection, storage management and security solutions. However, Symantec lacks a viable enterprise-class content analytics and search capability and its business analytics or business intelligence capabilities are also limited compared with IBM and others.

Futures

In its StoredIQ announcement, IBM mentions "building on prior acquisitions," including Vivisimo,[5] an enterprise search vendor that "helps organizations unlock and optimize the true business value of all their information, regardless of application or source, in order to drive innovation, real-time decisions and actionable insight." IBM's plans for Vivisimo, which is part of the Information Management group that also includes prior acquisitions such as Cognos, for business intelligence, and SPSS, for predictive analytics, are not yet clear.

Nevertheless, SPSS modeler has become an integral predictive component of *Watson-Ready* industry solutions for finance and healthcare, and Cognos is used by customers to analyze unstructured and structured data sources within the same ETL process or operation. Since ILG belongs to the IBM ECM group which, in turn, is part of the Industry Solutions Group headed by Craig Hayman, expect IBM to continue to cross pollinate solutions throughout its portfolio. Also, expect IBM to retain most if not all of the 50-plus StoredIQ employees. This would be consistent with IBM's recent strategy of retaining talent and resources to scale smaller acquired businesses. Examples include PSS Systems and the intelligent capture solution provider Datacap.

Conclusion

With the acquisition of StoredIQ, IBM has an additional solution set to provide its clients with both tactical and strategic benefits. Given, the momentum of IBM's ILG solutions group and the fact that StoredIQ has already been an ILG partner for several years, the integration should be relatively seamless, and StoredIQ should be a much-valued addition to the ILG portfolio.

StoredIQ provides IBM clients with both tactical and strategic benefits in their efforts to tackle Big Data.

End Notes

REFERENCES AND LINKS

[1] Link to IBM press release for StoredIQ acquisition
[2] Link to IBM ILG Solutions
[3] Link to DataIQ
[4] Link to IBM Defensible Disposal
[5] Link to Vivisimo press release

CONTACT INFORMATION

This paper, which originally appeared as a blog posting on the Parity Research blog site under the same title, was sponsored by StoredIQ, an IBM Company.

For more information regarding this document and other **Parity Research** products and services, please contact Gary MacFadden at gary@parityresearch.com, 1-203-313-0443 or http://www.parityresearch.com.

About IBM's Information Lifecycle Governance Solutions

Information is an organization's greatest asset, and the source of its biggest cost and legal exposure. As data growth exceeds governance response, an organizations' ability to gain insight and control over information rapidly diminishes. IBM's Information Lifecycle Governance business gives companies the ability to define and improve their information economics strategy and provide intelligent solutions that benefit all corporate stakeholders including IT, Legal, Records, as well as the business. With IBM, organizations can increase the value of their information assets while driving down cost and risk. For more information visit IBM Information Lifecycle Governance.

IBM Contact Jacqui Galow | Director, StoredIQ Marketing | IBM Information Lifecycle Governance (ILG) | 512.334.3156 | jgalow@us.ibm.com