Top Five US Bank Finds New Value from Metadata Application:
A CASE STUDY

by Charles Roe

ASG
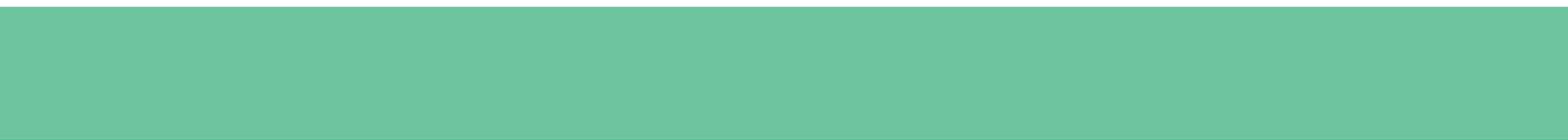
EXECUTIVE OVERVIEW

This study covers the early implementation and integration stages of a highly successful data-lineage project collaboration between a major financial institution and ASG Software Solutions.

The financial institution was looking to create a world class process and strategy to automate the data forensics and resolve regulatory requirements across the organization. In 2011, the institution purchased a portion of ASG's Enterprise Information Management Suite. The Initial drivers for the project were regulatory control and federal audit requirements. It is now growing into something much more comprehensive expanding across the enterprise delivering iterative business value. The following case study depicts a five year roadmap led from the Enterprise Information Architecture team and deployed by both business and IT leaders.

HIGHLIGHTS

- Early testing demonstrated that ASG's becubic could ensure regulatory compliance, provide different levels of change management, facilitate application modernization projects, and reduce operational risks. It also offered a vast range of technology support.

- ASG's Solution enabled visibility into all data assets, no matter whether the code existed in the distributed or the mainframe environment. Results show that ASG metadata suite increased efficiency 80-fold over manual data lineage and analysis processes.

- The cost savings were astonishing. Even after allowing for the initial purchase and installation costs, ASG's Solution reduced the actual expense of completing data lineage on 10 Key Business Elements (KBEs) in 100 applications, from $1,480,280 to $304,140. The costs would remain the same with ASG's Solution, no matter how many KBEs were included.

THE SITUATION

The financial crisis of 2007-2008 changed everything for banks. The status quo was completely disrupted: real estate markets were devastated, consumer confidence was at an all-time low, many financial institutions collapsed or were acquired by others, and a period of reorganization and more regulatory scrutiny had begun.

The bank needed to make sure it had control, insight, and ownership of its data. It needed a cohesive Enterprise Information Management (EIM) system that would allow deeper analysis of its data assets along its entire data streams, throughout all its applications. The system had to integrate various types of information about that data, so it could reliably:

- Guarantee regulatory compliance;

- Assess and ensure data quality;

- Effectively manage and utilize metadata resources;

- Ensure accountability and governance of all data assets, starting with key data elements and their technical instances;

- Provide a better system for both technical and non-technical users to access, identify, and use the data; and

- Set the foundation for future integration and growth.

The project's primary stakeholders knew they couldn't fix everything at once. The institution was too large to undertake a massive, enterprise-wide solution from day one.

They also knew any "one tool will fix everything" model would be unrealistic and naïve; so the project's key stakeholder and his Information Architecture team in Retail Banking started planning.

Their solution had to focus on mitigating the risk of inaccurate data, while also providing a single portal for data discovery. According to the key stakeholder:

"Our first task was to make sure we could satisfy federal and state regulators, by ensuring the institution could demonstrate a measure of data quality, by at least understanding how a given number came to be. We had to demonstrate reliably that the Bank has a handle on where data comes from, where it goes, and how the data changes along the way. We needed an effective data lineage system that could be implemented incrementally, that we could build a single data interaction portal around, and that had room for eventual enterprise-wide expansion."

The Retail Banking Division accounts for approximately 70% of the institution's data, so it was the logical starting point. Once the Information Architecture (IA) team could get that under control, the rest would fall into place. The IA team knew that, going forward, it would want to expand the program from data lineage, toward more proactively maintaining and measuring code quality.

But first, the team needed a working solution.

THE APPROACH

The IA team needed a data lineage solution that would work at the mainframe and distributed-environment levels. The team looked at a range of tools, and eventually decided to do Proof of Concept trials with three vendors, including a portion of the ASG EIM Solution suite.

The team explored a number of use cases before making a final decision:

- **Use Case One: Mainframe Testing.**
The first use case dealt with loading the platforms in question into the highly complex IBM MDM mainframe system, for initial testing on how information is produced, difficultly of use, and complexity. ASG proved quite successful at this. The IA team's growing positive relationship with ASG helped to drive the final decision. The team decided to continue forward with more use cases, to see how ASG's technology would perform.

- **Use Case Two: A Distributed Environment.**
Part of the IA team evaluated the possibility of leveraging Java code within an Oracle DB environment. Team members tried to analyze approximately 10,000 Java class files, in 25 DBs, with over 700 stored

procedures and millions of lines of code. The IA team spoke with ASG about uncovering and analyzing this code. Within four days after installing ASG's becubic, all of the code was extracted, loaded, and analyzed. The IA team learned a number of significant insights about the code; these facts ultimately helped drive the team's decisions about the entire Java code-analysis project. The ROI of the use case was astounding: what would usually have taken them many months and an average of 25,000 to 30,000 hours to complete manually took four days, and was almost completely automated, with the ASG Solution.

- **Use Case Three: Migrations and Conversions.**
The third use case explored a mainframe migration of DB2 from version 8 to version 10. A developer release note had identified a specific SQL function in one line of code. This function needed conversion, where a default argument was coming back as nulls instead of blanks. The release note was accidentally overlooked at some point during conversion; and an issue was already happening within a smaller, in-production, customer assistance application. The support team had worked for a few days to find all the places the function was being used, and had resolved the issue. The IA team then considered trying such a task on the much larger and more complex IBM MDM platform. To successfully resolve the same issue in the larger application would have taken weeks, according to the support team's manager. The ASG's Solution did it in under a minute.

The team was amazed, not just at the speed of the results but also at their overarching ramifications. The SQL function issue could cause considerable outages and other problems for in-production systems. But, ASG's becubic supplied a triage tool that could easily alleviate such issues.

Such a ramification was not easily quantifiable. The IA team needed a tool that could do data lineage and provide reliable quantitative metrics; ASG's Solution could also give endless triage possibilities and valuable predictive project-planning options.

The team conducted other use cases; but it was already becoming clear that the case was won.

THE SOLUTION

The IA team decided to purchase an initial license for a component of ASG's Solution (becubic for ADU) in 2011. It was installed and ready for its first production cases in early 2012. The IA team named its project "Forensic Data Analysis."

The project was set up in a unique fashion, differing from the application's typical use cases. Most customers use ASG metadata suite to understand the application code, while comprehension of the actual data is often incidental. The IA team decided to use the ASG Solution to actually understand the data itself:

1. Where does the data originate?

2. Where does it go?

3. What is impacting the data stream?

4. What applications are using it?

The Bank's initial concern was to satisfy federal and state regulators in terms of data lineage. The Bank had to be fully compliant with TARP and Basel II; it needed a firm grasp, with clear reporting mechanisms, of exactly what happened to a data asset from end to end along its data stream.

The ASG Solution supported the capability to drill into the source application, to understand the source code information and dependencies. Its benefits were well-defined [Figure 1]:

ISSUE	BENEFIT
Ensures Regulatory Compliance	Helps facilitate proper verification of data asset changes and potential impacts through internal control mechanisms. Understands the source of the information and where it is distributed throughout the organization. Ensures government regulations (such as Basel II, HIPAA, and others) are accurately adhered to.
Reduces Maintenance Expenditures	Quickly assesses application complexity. Identifies occasions to remove, fix, or reuse various areas of code. Gives recommendations about impacts when making such changes within applications.
Alleviates Operational Risks	Identifies code base problems before and/or during production phases, minimizing downtime. Provides strategic methodology to assess risk impacts.
Provides Multiple Levels of Change Management Facilitation	Enables more efficient impact analysis, dependency analysis, complexity analysis, sequence analysis, CRUD analysis, root cause analysis, and defect resolution, across all areas where a data attribute resides.
Aligns Enterprise/Business Objectives with Application Portfolio	Improves comprehension of how an enterprise's entire application portfolio can be directed, renewed, and transformed to better fit with overall objectives of the business.
Supports Broad Range of Technologies	Supports more than 100 technologies, including 17 mainframe languages, 20 distributed environment languages, 7 mid-range AS/400 ecosystem languages, 17 different databases, 12 scheduling tools, 14 SCM tools, and much more.

Figure 1 – Defined Benefits

The IA team had found its data-lineage solution. And, as stated by the key stakeholder, the team had found so much more:

"Our initial project was for data lineage and the meeting of Federal compliance guidelines. But the real sell was the quality of the code and the metrics that demonstrated the ROI of using the tool. We can now also look at potential impacts of code quality, simplify data center applications, modernize platforms, and be strategically set to deliver for the changing needs of our customer base, all with less risk and more responsibility, at lower costs."

THE RESULTS

At the project's start, the IA team had set out a 36-month roadmap into 2014. The team designated several Key Performance Indicators (KPIs) for success:

- Show regulatory compliance of data assets – primary KPI

- Decrease Time to Market Windows

- Decrease Cost per System/Cost per Element

- Demonstrate decreased costs, decreased risks, increased efficiency, and increased transparency

The initial project installation launched in early 2012 on a smaller scale, with 133 Key Business Elements (KBEs) on ten in-production systems, principally focused on credit card applications. The preliminary results were astounding – especially when tallied with the results of the use cases mentioned above. Compared to the cost of manual data lineage analysis of just one element, automated analysis using ASG's Solution gave an 80-fold difference in terms of work hours alone [Figures 2&3]:

WORK EFFORT ASSUMPTIONS	DESCRIPTION	MANUAL (W/O TOOL)	AUTOMATIC (W/ TOOL)
Simple Application	Straight I/O	80 Hours per app	1 Hour per app
Medium Application	Crosses Platforms & Jobs, Straight I/O, Single DB	160 Hours per app	2 Hours per app
Complex Application	Crosses Platforms & Jobs, Multiple I/O, VAPS and Multiple DBs	240 Hours per app	3 Hours per app

Figure 2 - Work Effort Assessment – Manual versus Automated Processes

	MANUAL PROCESS	TOOL BASED PROCESS	MANUAL PROCESS	TOOL BASED PROCESS
Number of Elements	1		10	
Number of Applications to Search	100 (40 Simple, 40 Medium, 20 Complex)		100 (40 Simple, 40 Medium, 20 Complex)	
Amount of time to require to search each application	14,400 Hours	180 Hours	64,360 Hours (see page 15 for details)	180 Hours
Resource Cost (@ $23/hr.)	$331,200	$4,140	$1,970,893	$4,140
Total Cost	$331,200 (no software cost)	$304,140 (Capital for Software = $300K)	$1,480,280 (no software cost)	$304,140 (Capital for Software = $300K)

Figure 3 – Cost Analysis - Manual versus Automated Processes

The numbers presented in Figures 2 and 3 deserve deeper consideration. The fact that the IA team could address one data element in 100 systems (40 simple, 40 medium, and 20 composite) in 180 hours, versus 14,400 hours when performed manually, is surprising but accurate.

(The IA team's estimate of $23/hour in manual-side resource costs is on the low end of the spectrum. So, while the figure of $331,200 for manual operation vs. $4,140 for a tool-based process is an effective measurement, actual costs on the manual side could be significantly higher than shown here.)

The project's legitimacy was further demonstrated when the team scaled up to 10 elements. Such an undertaking would cost $1,970,893 for manual processing, but only $4,140 with the ASG Solution. Once the initial install (with setup of application code, testing, debugging, and related activities) is complete, the processes are automatic; these costs do not change when new elements are added.

In terms of risk avoidance, comparing manual to automated processes also provided profound results [Figure 4]:

SME MANUAL APPROACH	ASG SOLUTION
Missed entire sets of programs and functions areas within the programs.	Uncovered all programs and functions.
Spent months gathering incorrect data.	Gave instantaneous results, once adapters were installed.
Missed hops in the lineage.	Continued to produce more uncovered hops during incremental loads.
Difficult to navigate IIS environments without scanners.	Could go between mainframe and distributed environments.

Figure 4 - Risk Analysis - Manual versus Automated Processes

The initial run of 133 KBEs gave the IA team enough confidence to move onto the next stage in early 2013. The team set a goal to implement 1000 KBEs and, subsequently, 40-50 in-production systems. These were expanded to include home loans, credit cards, deposits, and customer information systems.

After the project's initial success, the IA team began developing and testing a Metadata Portal. This enables technical and non-technical data consumers to efficiently drill down into the transformation detail of any data asset currently in the system. They can move through a data asset's stream (from consumption to origination point), view Data Quality metrics at any point in the stream, analyze file controls and data changes, and follow exactly what happens to a particular element throughout the stream of provisioned applications.

Recently, the IA team presented its Forensic Data Analysis tool to a C-Level executive interested in learning more about the system. Following the presentation, the executive said, "Either you're lying to me, or this is the most underutilized or undervalued tool at the entire bank." The presenter smiled and told the executive that the project is only just beginning, that the team's already building on top of this foundation.

MOVING FORWARD

The project is still in its infancy; the IA team has only just started to gain a greater awareness of what ASG can do for the institution. The second implementation stage of 1000 KBEs in 40-50 systems is moving along smoothly. The team has a roadmap for 2014 and beyond.

The VP of Technology in charge of managing the execution of data lineage remarked:

"We've just begun this journey, really. We are nearing our first decisive maturity plateau, but we are not there yet. We are excited now that we've reached a point where people are really starting to take notice of the project; they are sitting up in the seats and contacting us for more information. We have a major opportunity over the next 3 to 5 years. The chess pieces are set; and our strategy is underway to move our entire game forward within the organization as a whole."

The Metadata Portal has averaged about 800 user hits per month. This is expected to increase considerably as the team's marketing campaigns take off. The intentionality built into the entire project from its infancy has proven successful. Now, with so many more Lines of Business waiting to come onboard, the team believes the project will keep delivering greater, continuous value.

The team continues to gain traction within the institution. While the application discovery initiative initially started as a project, it has proven itself advantageous in so many other areas; it's becoming a central aspect to the institution's entire EIM program.

CRITICAL ELEMENTS OF SUCCESS

Choosing the right technology is obviously critical, and the ASG Enterprise Information Management Solution certainly fit the bill. However, the bank also got several other critical things right. The IA team:

- Focused from the beginning on the entire process, not the outcome (team members knew the desired outcome, but stayed focused on step-by-step processes).

- Had support networks and procedures in place, and clearly documented the project from the beginning.

- Set correct levels of expectations with key business partners and stakeholders, around the amount of effort involved in doing this level of analysis.

- Evaluated three vendor trials supported by several key use cases, to discover the products' full potential before buying.

- Planned advertising and marketing campaigns from the project's earliest stages.

- Clearly defined a 36-month roadmap, whose iterations were carefully detailed and followed.

- Made executive sponsorship evident from the project's early stages.

- Focused on KPIs that could be quantitatively demonstrated as the project moved forward.

CONCLUSION AND DIRECTION

The ASG Solution helps to align various corporate information management efforts, by enabling more informed conversations on how the data is used by business and IT. The collection and analysis of such data brings all data users (business and technical) onto the same page. It helps an enterprise deal more strategically regarding its applications and subsequent authoritative source systems. It helps users understand which factors lead to redundant data, how to resolve those redundancies, and how to prevent them in the future. It helps an enterprise leverage all its information assets more successfully.

When IT plays a significant role toward a company's overall business strategy, strategic IT evaluations lead to a better business outcome. Forensic Data Analysis efficiently finds and characterizes data structures and elements, including your many non-linear elements. It helps you understand the state of information at all levels within your organization. It helps to ensure that business and IT are aligned together, to promote marketplace success and future progress.

THE CASE STUDY PARTICIPANTS

A recognized innovator in enterprise IT software solutions, ASG Software Solutions has optimized 85 percent of the world's most complex IT organizations for over 25 years.

ASG creates and deploys unique software solutions that reduce cost, mitigate risk, and improve service delivery throughout the IT lifecycle. ASG's comprehensive solutions help to solve today's challenges, such as Cloud Computing and Big Data. They drive business forward with insight and control across Cloud, distributed, and mainframe environments.

The ASG Metadata Solution provides the foundation for optimizing the business value of diverse application portfolios. Through its advanced application discovery and understanding capabilities, it presents an end-to-end view of an organization's applications assets across the enterprise, from mainframe to distributed technologies.

The bank discussed in this case study is one of the world's largest financial institutions. It serves a wide range of consumers, from individuals to large corporations, with a full range of banking, investing, asset management, risk management, and other financial products and services. It serves more than 50 million consumer and small-business customers from more than 5,000 retail banking offices and 15,000 ATMs. Its award-winning online banking service has more than 12 million mobile users. It is among the world's leading wealth-management companies, and is a global leader in corporate and investment banking. It offers industry-leading support to more than 3 million small business owners though many innovative, easy-to-use online products and services. It operates in more than 40 countries.