



Axis Technology, LLC's Data Masking Precludes the Risk of Sensitive Data

Analyst: Anne MacFarland

Management Summary

Data leakage makes front-page news. Many organizations are still stuck in denial about the extent of their vulnerability. Due to the prevalence of sharing information with partners, the benefits of harvesting customer experience, and the need to leverage the knowledge of the crowds, many more sources of data, including data that is sensitive or may become sensitive, are captured. Enterprises use data in a variety of ways, beyond the process that captured or generated it. It is used, in the aggregate, for research, analysis, trending and as the basis for operations finesse. Developers and testers use data for application and disaster recovery testing. In almost all of these situations, exposure of sensitive particulars adds nothing of value – only a great deal of risk. **A consistent, properly formatted but altered value of sensitive information – a masked version – will do just fine.** The mostly-occluded account numbers on restaurant receipts are a familiar example of this approach.

Quite apart from regulatory demands, good business practices dictate that all sensitive business data that migrates beyond the controls of your production environment be protected from any exposure. If your best efforts are palliative, not preventative, that is not good enough. Palliative efforts include partial solutions, like securing information in a particular database and email, but not the multiple copies of that information that reside elsewhere. They include manual, inconsistent efforts. They include the once-and-done approaches that do not address the constant changes in the roles of individuals, in regulations, and in the data itself.

Deciding how to protect sensitive data is made more difficult by the variety of solutions that pertain to the problem, but may not fully address it. Encryption/decryption has great value – but as the sensitive information is decrypted, the data leakage risk re-emerges. The ability to save information decrypted – or to print – may make an encryption solution inadequate. Full-blown digital rights management can curb the ability to print, but, often, is not politically or practically supportable. **Data masking is an alternative that addresses the risk of sensitive information permanently.** It transforms sensitive data elements with different values in the sale format. You still have validity and consistency – but without enabling identity theft or invasion of privacy. This approach takes reliance on the integrity of others out of the equation. **Done right, data masking can replace uncertainty with a provable way to secure, permanently, all sensitive data elements even as the nature of these elements changes.**

Axis Technologies, based in Boston, Massachusetts, has been providing custom data masking solutions to large financial services companies for some years. It recently has brought to market *DMSuite* data masking software to address smaller, younger, less complex companies who need a data leakage solution that is complete, affordable, and doable. The technology is agnostic to topography (centralized, distributed) and platforms. For more details about Axis Technologies' approach, please read on.

IN THIS ISSUE

- **Masking 21st Century Business Data 2**
- **Axis *DMSuite* Quick Start..... 2**
- **Conclusion 3**

Masking 21st Century Business Data

The increasingly multi-sourced nature of 21st Century business operations and the financial attractiveness of software-as-a-service (SaaS) both require a comprehensive, provably-effective approach to protecting confidential information. Internal operational environments usually have good, granular access controls. However, if data is going to an untrusted environment, masking sensitive elements will be a good precaution. Today, in many business ecosystems, data travels extensively.

For data masking, you want to solve the problem – but not to cause unnecessary disruption. Masking is a simple transformation. It has to scale to handle cumulative data growth. It has to cross platforms. It has to address applications in language they expect. However, in all but these essentials, it must be lightweight – otherwise, companies may postpone implementation of a solution.

While technology is part of the solution, so is process, cautions Mike Logan, President of Axis Technology, LLC. How the masking process is organized and implemented turns the necessary evil of data masking into a feasible-and cost-effective initiative. Once data masking has been set up, (the few weeks of services needed to set up the process and transfer the skills is built into the price), checking for the intrusion of unmasked data and implementing new masking algorithms can be done episodically.

Axis has found that its many large customers have tended to use a *Center of Excellence* approach. This ensures that the definitions stay consistent across business units and that the whole of a domain is addressed. For smaller companies, the term *virtual center of excellence* might be more appropriate, as the *Axis DMSuite Quick Start* data masking software can usually be run in a virtual environment. DMSuite was designed to be usable by a data administrator. It is driven off a discovered inventory or a list of sensitive data elements. It performs a search of data sources – including those in disaster recovery environments – to discover where data can be exposed. It has no need to duplicate business logic. It manages limited metadata – only that needed for masking and for provisioning. Typically, 5% of data will need masking. The rest will flow through the process.

Masking Process Requirements

For data masking to be effective, enforceable,

scalable, and provable, the following characteristics must be present.

- **It must be done consistently across the entire relevant environment.** Otherwise, the ongoing task of identifying unmasked elements and new elements that must be masked is much more difficult. It should be done using a set of standard algorithms that mask all the data elements of a given domain (e.g., First Names) the same way. Otherwise, consistency will be hard to maintain.
- **It should be independent of a particular database.** While major databases provide masking tools, they are specific to a particular database. Any organization that has to share data with others will find a database-specific approach probably does not suit its situation.
- **To be lightweight and easy to deploy, it should be independent of other database and data warehouse tools** such as ETL (extract, transform, and load) tools. These tools are overkill for the problem of protecting sensitive data and come with an overhead that makes the deployment process more difficult. As President Mike Logan puts it, “Data masking works by generating code – not by invoking another, more complex process.”
- **It should not impede performance.** Databases are growing, and anything that slows their performance is becoming intolerable as a “solution.” Masking data should be as easy as putting on every-day make-up. It should not require a crew of three and special lighting.
- **It should not be considered a once-and-done-forever process.** As data sources change, regulations change, and business relationships change, definitions of what must be protected evolve. This is just another reason that the technology and process that support data masking must be comprehensively scalable, capable, self-documenting, and cost-effective.

Axis Technology’s DMSuite

The goal of DMSuite QuickStart is to make data masking more feasible and pervasive in an organization, in a way that is less intrusive, quickly deployable, and more moderate in price than other alternatives. While Axis’ experience gives them the competencies to do extensive customization across the full range of servers including *IBM System z*, the sweet spot for DMSuite QuickStart is companies that are,

inevitably, complex – but not downright peculiar. These often-youngish companies see data masking as a necessary evil that must be done competently and simply.

DMSuite QuickStart is modeled on a five-step process.

Step One – Identification

This is the step where ease of deployment and completeness make data masking feasible for smaller organizations.

DMPProfiler identifies sensitive data. It does not require DBAs to present it with sets of primary and foreign keys, but instead discovers elements (usually columns) that contain sensitive data. It presents results in a sensitive data inventory where users can review and make changes as needed. This may be a case of the inferential sensitivity described in Exhibit 1, to the right.

DMSuite is not limited to structured data, but is limited in what unstructured elements it can mask. It requires a small amount of customization to do things such as check images, text in power point presentations or PDFs, or logos.

Going through this process will produce an inventory, in *DM Generator*, of sensitive data element occurrences.

Step Two – Selection

This is a step where senior business management expertise is key – and buy-in to the process occurs.

With the elements identified by *DMPProfiler*, a company has three choices.

- It can delete the information. When faced with the wealth of information they have amassed, many companies find a great deal of information that they are not legally obliged to keep – but are obliged to protect if they do keep it. The lowest cost and lowest risk option is to delete it. This has the side effect of recovering storage capacity and simplifying the data landscape.
- It can isolate the information in systems that are physically secure and not networked. This is the proper solution for the family jewels (formulas, trade secrets, etc.) that must be retained in the clear.
- It can classify the information (confidential, highly confidential, internal), and mask the information appropriately.

Step Three – Validation

This is the step that directly addresses IT

Exhibit 1 — Kinds of Sensitive Information

Sensitive information is not as cut and dried as it used to be. Customer information (both institutional and individual) has always been sensitive. Information about products yet to be released has always been considered confidential, but exceptions to this rule abound. With customer information into product development, the grey areas of potential data leakage grow.

Website visitor information is also useful to an organization – but as visitors become customers, information linked to them may become confidential.

Some information – inferential information – may be sensitive in the aggregate. Masking the particular holdings of a stock portfolio will not foil a spy of the values, and their fluctuations, are in the clear. The more detail we gather about the health history of individuals, the more precisely their history identifies them even if traditional sensitive information is unavailable. The cadence of information events (regardless of the information itself) can also reveal an acquisition or spin-off.

These are all considerations that were not so visible when systems were more isolated. Some of them – many of them – can be addressed by data masking. Others must be mitigated by legal contract.

requirements.

Before a company proceeds further, it is prudent to examine the full effect of masking. Many developers are accustomed to using their personal data to make sure things work. Using the masked form of that data requires more memorization. The ability to validate the effects of masking assuages many qualms.

Step Four – Implementation

This is where the rubber meets the road – and where Axis' experience with complex situations has allowed them to craft a simpler masking solution.

DMApplicator is an XML-based tool that interfaces with the information stored in the sensitive data inventory created by the *DMPProfiler*. It has the necessary plug-ins for specific platforms that allow it to mask data in a cost-effective way. For Oracle, it generates SQL

Loader code. For the mainframe, it can generate COBOL copybooks and generate masking code. Smaller businesses will need a smaller variety of plug-ins – but completeness is essential.

Over time, there always changes to information and the kinds of information an organization generates and uses. DM Applicator works in two modes. For the initial pass, it works in a batch-like mode. For subsequent use, it works in an update mode that is less performant than batch mode but non-intrusive.

Step 5 – Certification

This is what lets both business and IT rest easy.

DMAuditor supports the certification that lets all parties sleep at night. This tool rescans the environment to discover if unmasked data has polluted by data that has not been masked. It does this by running algorithms to find (but not process) exceptions. This may happen with manually input data, specialized testing, etc. With these exceptions flagged, it is easy to make the delete-mask-isolate decision and restore the completeness of the solution.

Conclusion

The business risk of data exposure is high enough to make a data leakage solution mandatory. Data masking is a permanent solution that, when done right, does not affect operations. It can be done cost-effectively. It is a process that, to be effective, has to permeate assumptions of both IT and the business. By doing so, it can ease challenges in both realms.

In summary, DMSuite QuickStart creates a measurable, documented, and repeatable process to protect sensitive data elements throughout a business environment. It also allows a business to use data flagrantly in support of its business without risk of a data breach. Both are essential to doing business with confidence.

DMSuite Quick Start closes back doors that cause worry and risk. Consider how its use would benefit your organization.



About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

Anne MacFarland is Director of Information Solutions for The Clipper Group. Ms. MacFarland specializes in strategic business solutions offered by enterprise systems, software, and storage vendors, in trends in enterprise systems and networks, and in explaining these trends and the underlying technologies in simple business terms. She joined The Clipper Group after a long career in library systems, business archives, consulting, research, and freelance writing. Ms. MacFarland earned a Bachelor of Arts degree from Cornell University, where she was a College Scholar, and a Masters of Library Science from Southern Connecticut State University.

- ***Reach Anne MacFarland via e-mail at Anne.MacFarland@clipper.com or at 781-235-0085 Ext. 128. (Please dial “128” when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, The Clipper Group Explorer, The Clipper Group Observer, The Clipper Group Captain's Log, The Clipper Group Voyager, Clipper Notes, and “*clipper.com*” are trademarks of The Clipper Group, Inc., and the clipper ship drawings, “*Navigating Information Technology Horizons*”, and “*teraproductivity*” are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.