semele

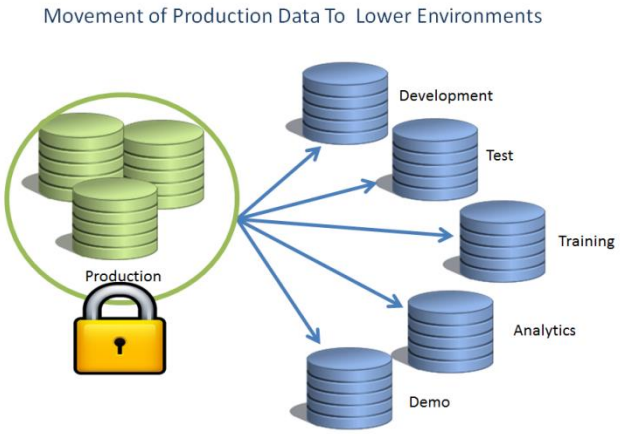
The Prevalence of Exposed Production Data in Lower Environments

A Complete Solution for Complex Data Testing Environments
semeledata.com

There is so much in the news lately about data breaches and the exposure of sensitive customer data; that it's enough to keep Chief Information Officers (CIOs) and Chief Information Security Officers (CISOs) up at night. Not to mention their boss who is ultimately responsible for the security of the company's data.

So where is your customer data? Shocking but true – most companies are not fully aware of their exposure. Sure, everyone knows where the "production" data is located. Most do a great job of securing that environment and keeping up with patches, Equifax notwithstanding. But almost all companies expose their production data and they do it multiple times, in their test and development environments. The lower environments are so called because they do not have the same protection and rigor as the production environment. The reasons for the presence of production data in the lower environments are myriad but mostly it is because software developers and testers need the complexity that can only appear in production data to properly test applications in development.

There are a number of technologies available to properly mask production data for use in the lower environments, and a good one should be implemented immediately. There is no excuse for exposing your organization to this sort of risk. The problem is that most companies have no idea where that data is located because they have allowed their staff to copy that data for so long at such frequency with little to no controls.



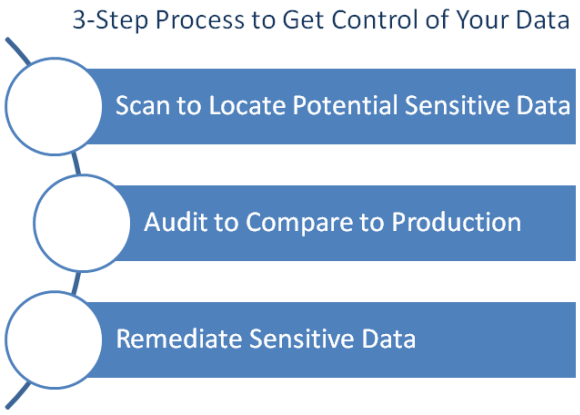
So how big is this problem? Well, the movement of sensitive data to the lower environments happens every day in nearly every company. There are few places that a company will allow their entire production data to be copied to, let alone allowing it to be copied multiple times. But that is precisely what they do when it comes to testing their software applications. They copy their production data multiple times and give wide access to their sensitive customer information to all development staff, vendors working with their company, offshore testing staff, etc. We have worked with companies that have over twenty copies of their production environment in lower environments to facilitate testing.

As if all this exposure of data is not bad enough, it gets worse. Usually, there is no one tasked to keep up with these lower environments and people forget that the data is there. This leads to an increasing risk of exposure over time.

Locating Your Data

The most effective process for solving this problem is to locate this data and then either delete it or mask it to obfuscate all the sensitive and identifying information.

So how do we find all this data and determine if there is exposure? To find the data, first a company must perform a scan of their file systems and databases and then audit the data in those files and databases. Both steps are necessary. A little definition of these two processes is in order.



Scanning

Scanning involves rapidly running through the file systems and databases to check if anything resembles real data. For example, a scan will tell you if something that appears in a file or database looks like a Social Security Number, meaning that it is a nine-digit number or follows the familiar pattern XXX-XX-XXXX. Scanning software may also interrogate the column names looking for things like "SSN" in the column name. Basically, the scan gives you a list of candidate places where you might be exposed. But well masked data or manufactured data will also look like it could be an exposure. The scanning software cannot tell the difference so you could have many false positives identified. The second step to the process, Audit, tells you if the data is real or not.

Auditing

An Audit will take the values that appear in your lower environment files and databases and compare them to the actual values in your production data to tell you if the data is an exposure. By definition, if a piece of identifying information in your lower environment matches a value in your production system, it is an exposure.

Semele's audit solution paired with our consulting services can help you find out if you are at risk with customer data in your lower environments. Semele can compare data across multiple platforms simply and quickly. And, it can be scheduled to run at a regular interval to proactively identify when toxic data has been moved to the lower environments.

The Semele Advantage

Semele is an enterprise solution for the fast, efficient and secure subsetting and masking of production data for testing. Semele was designed specifically to work within complex data environments. It is sophisticated yet easy-to-install, easy-to-use and supports all major databases.

- **Semele Audit Solution:** An automated solution that compares values in lower environments with those in production to identify the location of toxic data.

- **Semele Test Data Subsetting:** Simultaneously subset, transform and protect data for testing using an automated, repeatable solution.
- **Semele Test Data Comparison:** An innovative comparison utility that allows testers to compare data across any platform to quickly identify differences.

Semele was developed by Meridian Technologies, a leading consulting, staffing and technology firm, as a solution to a challenge facing many of its clients: the need for a complete enterprise solution for fast, efficient and secure sourcing and masking of production data for testing. Now you can leverage that expertise for your business's own test data challenges.

Contact:

Don Kiernan
Vice President, Sales
904-512-7917
don@semeledata.com

LOCATE YOUR EXPOSED DATA

- Easily compare values in lower environments to production
- Compare millions of records across multiple platforms simply and quickly
- Quickly identify differences in data sets through web-based reports.
- Set automated reports to run at regular intervals