Data Masking Made Simple With DataSunrise

DataSunrise, Inc.

http://www.datasunrise.com

Note: the latest copy of this document is available at https://www.datasunrise.com/documentation/resources/

Data masking enables database owners to protect sensitive data of any kind by “masking” it. It means, that the actual data is going to be replaced with some useless values. You could probably see the ATM receipts where credit card number is replaced with asterisks or Xs — that's the most obvious example of data masking.

ID	First Name	Last Name	Salary
1234	Justin	Doe	1500
1223	Kathy	Abrams	1400
1233	John	Bradley	1300
1244	Mary	Grant	1200
1253	Karen	Smith	1200
1352	Jack	Burns	1200
1454	Claire	Chow	1300
1341	Robert	Hill	1100

actual database entries

ID	First Name	Last Name	Salary
1234	Justin	Doe	MASKED
1223	Kathy	Abrams	MASKED
1233	John	Bradley	MASKED
1244	Mary	Grant	MASKED
1253	Karen	Smith	MASKED
1352	Jack	Burns	MASKED
1454	Claire	Chow	MASKED
1341	Robert	Hill	MASKED

masked entries

As you may guess from its name, DataSunrise Data Masking tool is used to mask data contained in the protected database. In this article we will highlight some data masking-related points.

What is data masking used for?

In simple words, data masking procedures used to protect sensitive data of any kind contained in a database.

Let's take a certain “Acme” company for example. Acme's database contains some clients' and employees' confidential info like national identification numbers or credit card numbers. Of course the database may contain sensitive data of other kind like production data, accounting info and so on.

Since the company's database could be used for reporting, analysis, software development or testing procedures, the sensitive data should be protected from being exposed. The point is that software developers as a rule don't need actual data the database contains — in most cases they need just a “dummy” database, that looks like a real one and works like a real one. And the best way to create such a dummy is to use data masking.

Some people confuse data masking with data encryption but it is not the same. Encryption makes data completely unreadable for an unauthorized person or application. Masked data, in turn, should remain readable and appear consistent while being, actually, a fake. Data masking does not prevent access to the data, it just hides the actual data like a mask.

Static and Dynamic data masking

There are two methods used to mask data in a database: static and dynamic. And here is the difference:

Static data masking procedures involve creating a copy of live database and replacing actual data with fake one. It is the only method of data masking used by companies sending their databases to outsourced software specialists for testing.

Static data masking method has some serious drawbacks. First, before masking applied, the real data should be extracted from database for evaluation and inspection, so this situation poses a potential threat of data exposure. Beyond that, database duplication requires some empty space on company's server or even a new server, so this method of data masking could be a pretty pricy. And last but not least: the “dummy” database lags behind the live database. Of course, it can be updated periodically but this process requires additional time and could be related with some issues.

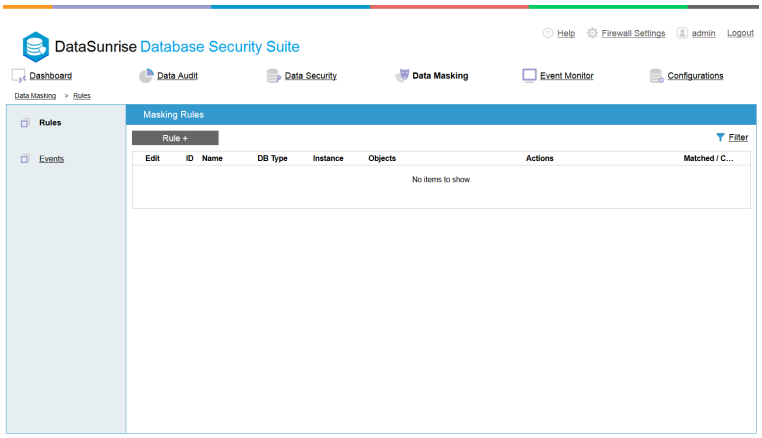
Dynamic data masking, in turn, involves replacing sensitive data with fake one on-the-fly, while the data is being transferred to a client. In other words, data masking software changes the way database responds a query, so it requires no interference to the database itself and the real database entries remain untouched. The data is masked before it exists the database so it is a very secure and reliable method.

How DataSunrise Data Masking works

As you can see, the dynamic data masking method is much more versatile and that's why we use it in our product.

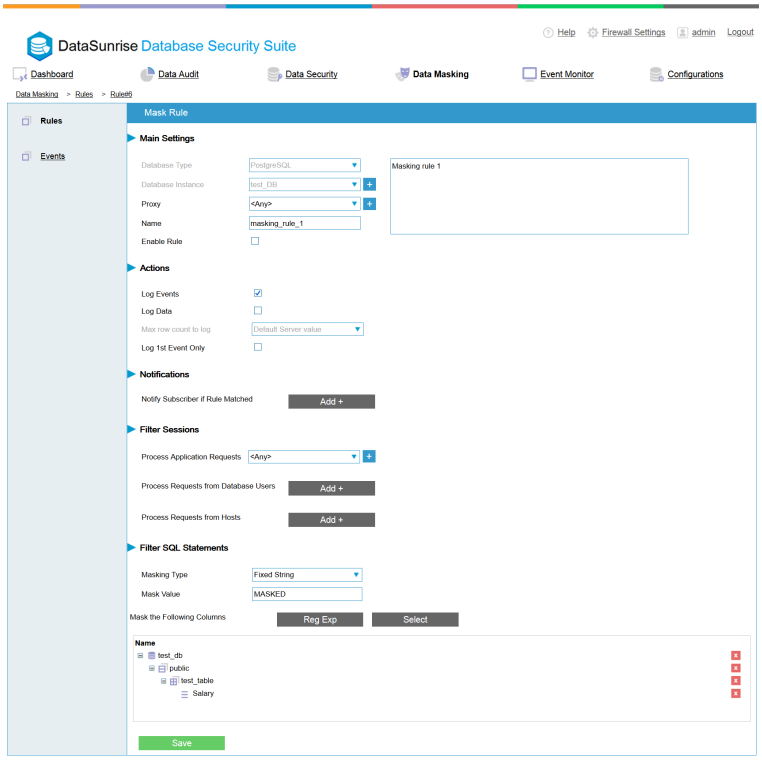
DataSunrise suite works as a proxy — it intercepts SQL-queries to the protected database and modifies these queries in such a way, that the database fetches not actual, but random or predefined data.

Before you use the DataSunrise Data Masking you need to determine which database entries need protection and where they are located. Note that DataSunrise can mask a complete database as well as data in separate columns only. DataSunrise logs all the actions, so you can check what is happening anytime.



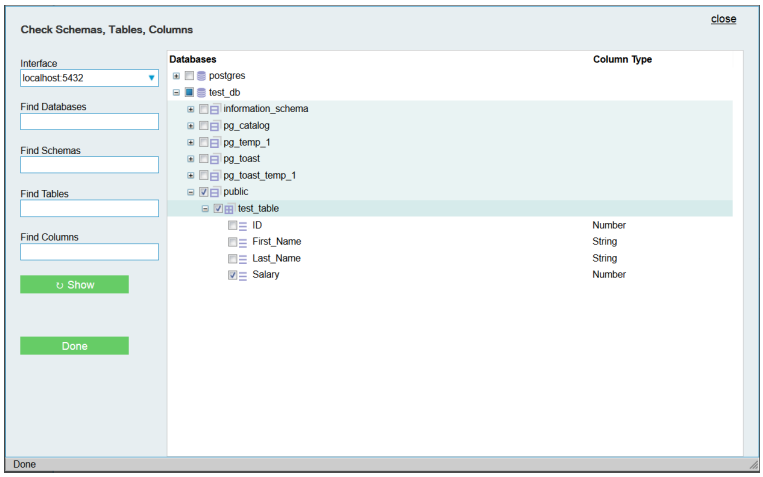
Data Masking user interface

Using DataSunrise data masking tool is very easy. All you need to do is to enter DataSunrise dashboard and create some safety rules.



Creating a data masking rule

Here you need to enter information required to create data masking rules. You can define application which requests will be processed by the firewall. Then you need to define SQL-statements to be filtered and select masking type to be implemented. It means, that you can select a method of fake entries generating.



Database elements explorer window.
"Salary" column of the "test_table" table will be masked

Then you should select the database elements (schemas, tables or columns) to be protected. It can be performed manually via handy database elements explorer or by using regular expressions.

And that's all. Quite simple.

Some conclusions

DataSunrise data masking provides you with another reliable tool for info protection. Along with the DataSunrise Database firewall and SQL injections prevention tool it can become an additional line of defence against digital threats.

For more information about DataSunrise Database Security capabilities please refer to DataSunrise user guide or contact us at info@datasunrise.com