

# Test Data Management for Security and Compliance

Reducing Risk in the Era of Big Data

This document contains Confidential, Proprietary and Trade Secret Information ("Confidential Information") of Informatica Corporation and may not be copied, distributed, duplicated, or otherwise reproduced in any manner without the prior written consent of Informatica.

While every attempt has been made to ensure that the information in this document is accurate and complete, some typographical errors or technical inaccuracies may exist. Informatica does not accept responsibility for any kind of loss resulting from the use of information contained in this document. The information contained in this document is subject to change without notice.

The incorporation of the product attributes discussed in these materials into any release or upgrade of any Informatica software product—as well as the timing of any such release or upgrade—is at the sole discretion of Informatica.

Protected by one or more of the following U.S. Patents: 6,032,158; 5,794,246; 6,014,670; 6,339,775; 6,044,374; 6,208,990; 6,208,990; 6,850,947; 6,895,471; or by the following pending U.S. Patents: 09/644,280; 10/966,046; 10/727,700.

This edition published June 2013

## Table of Contents

<b>Executive Summary</b> . . . . .	<b>2</b>
<b>Big Data as a Security Risk.</b> . . . .	<b>3</b>
<b>The Informatica Advantage</b> . . . . .	<b>4</b>
Informatica Data Subset. . . . .	5
Informatica Persistent Data Masking . . . . .	5
<b>The Business Value</b> . . . . .	<b>6</b>
<b>The Informatica Solution in Action</b> . . . . .	<b>8</b>
<b>Conclusion</b> . . . . .	<b>9</b>

## Executive Summary

The amount of information managed in enterprise data centers is predicted to explode fiftyfold in the next decade,<sup>1</sup> with new types of data flooding in from previously unknown sources. Meanwhile, IT environments are increasingly complex, requiring multiple copies of data volumes for test, patch, development, and training purposes as well as full backups.

As big data gets bigger, your IT organization must develop scalable new strategies to mitigate the increased risk of losing control of it. This white paper examines how test data management can play a vital role in safeguarding data privacy and ensuring regulatory compliance by:

- Defining and classifying sensitive data
- Identifying where sensitive data lives across applications and databases
- Creating data subsets and applying consistent data masking rules across systems
- Measuring, monitoring, and proving data security

It discusses the benefits of the Informatica® solution for test data management. Built for scalability, flexibility, and ease of use, this solution enables your IT organization to protect private and sensitive data, decrease the risk of data breaches, and effectively meet compliance requirements on a timely basis—all while decreasing the cost of data, increasing its value, and maximizing its return.

<sup>1</sup> IDC, "The 2011 Digital Universe Study: Extracting Value from the Chaos," June 2011.

## Big Data as a Security Risk

The sheer volume, variety, and velocity of big data make it particularly prone to data breaches. A study by research firm IDC estimates the data in enterprise data centers will grow by a factor of 50 by 2021, driven in part by data generated by social media, mobile computing, and device-to-device interactions, all moving around the globe at high speeds. Your IT organization faces complex decisions about how to secure this data against external threats, including whether sensitive data should reside at the front, middle, or back of office applications.

At the same time, internal threats can prove just as dangerous. Forrester has reported that 70 percent of data breaches are caused by insiders.<sup>2</sup> In a May 2012 Ponemon Institute report, organizations surveyed said 50 percent of cases involve an insider such as a privileged user.<sup>3</sup>

The increasingly complex IT environment poses further challenges to data privacy. Most IT organizations need to develop and maintain multiple applications to support individual business units. Each production application may require multiple copies of data sets for test, development, and training purposes, as well as onsite or remote backups. Each copy, in turn, may have a number of resources with direct access to systems containing data that is potentially sensitive or subject to privacy regulations. To mitigate escalating costs, your IT organization may want to use offshore resource models or deploy software as a service (SaaS) or cloud-based offerings. To make the best use of these offshore, outsourced, or cloud models, you need to mask the data.

<sup>2</sup> Forrester, "Test Data Privacy Is Critical To Meet Compliances," October 2009.

<sup>3</sup> Ponemon, "Safeguarding Data in Production & Development: A Survey of IT Practitioners," May 2012.

## The Informatica Advantage

Informatica offers a solution for test data management that substantially reduces the risk of data breach and data volumes while simultaneously improving compliance with data privacy policies, regulations, and mandates. This unique solution is based on the industry-leading Informatica Data Integration Platform™, a comprehensive, open, unified, and economical platform that supports a centralized data management approach so your IT organization can leverage the solution across multiple business lines to conduct audits and comply with data privacy policies and regulations enterprise-wide. The Informatica solution for test data management supports your organization's data governance program and includes built-in best practices and templates to accelerate implementation.

This comprehensive solution helps ensure that reducing the volume of and mitigating risk around your test environment is not just a one-time initiative, but part of an overall, ongoing program by:

- Addressing the most comprehensive set of databases and applications, on- or off- premises
- Providing a centralized management and control center for consistent enterprise-wide data privacy protection and test data management
- Masking to support a variety of custom and packaged applications, databases, and data center policies
- Handling data volume growth—either organic growth or as new applications are deployed in the data center

Leveraging the Informatica Platform, the Informatica solution for test data management addresses each part of the data lifecycle:

1. Defining and classifying sensitive data, including data and metadata patterns.
2. Discovering where that sensitive data lives across databases and applications.
3. Applying policies to create subsets of production data for testing and training purposes.
4. Masking data consistently across the systems of an organization to meet various compliance standards.
5. Measuring and monitoring to provide ongoing proof that data has been protected.

The Informatica solution for test data management is comprised of two products: Informatica Data Subset and Informatica Persistent Data Masking. Working together, they seamlessly protect test data in any format—unstructured, semistructured, or in industry data such as SWIFT, EDI, and HIPAA.

## Informatica Data Subset

Informatica Data Subset is flexible, scalable software for creating, updating, and securing data subsets—smaller, targeted databases—from large, complex databases. These referentially intact subsets of production data from interconnected systems dramatically reduce the time, effort, and disk space needed to support nonproduction systems.

Informatica Data Subset quickly replicates and refreshes production data with only the most relevant, high-quality application data. In the era of big data, Informatica Data Subset can substantially reduce the data required for testing and QA.

## Informatica Persistent Data Masking

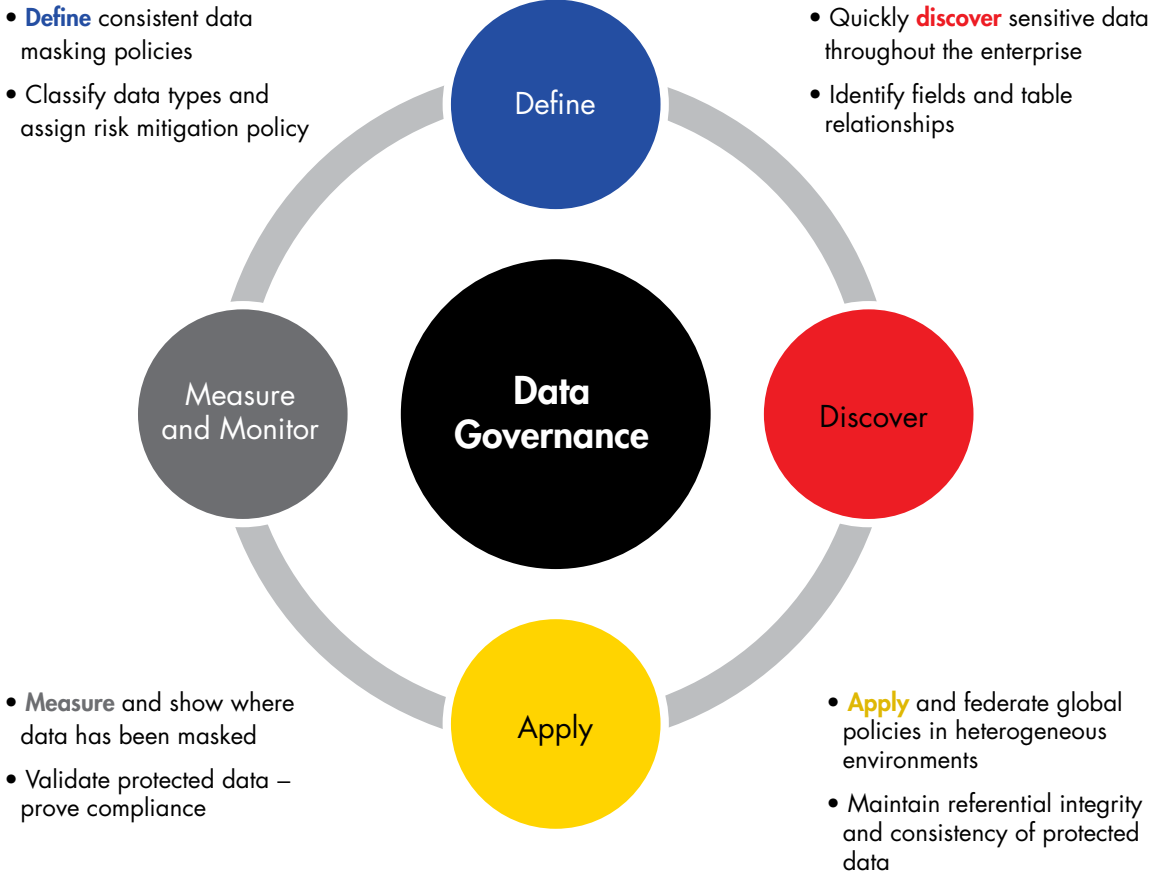
Informatica Persistent Data Masking allows your IT team to create, maintain, and apply data masking policies to secure the sensitive data in your test and production environments and shield it from unintended exposure. This scalable data masking software provides unparalleled enterprise-wide scalability, robustness, and connectivity to a vast array of databases, masking test and development environments created from production data regardless of database, platform, or location.

The software provides sophisticated, flexible masking rules that allow your IT team to apply different types of masking techniques to various data used in testing, training, and other nonproduction environments. With Informatica Persistent Data Masking, IT organizations can create enterprise-wide data privacy policies while maintaining segregation of duties. Auditors and security officers can define policies while developers, testers, and trainers retain access to contextually rich, functionally intact, and realistic-looking data without impacting application functionality.

# The Business Value

Minimizing the risk of data breaches is only one of the business benefits of a test data management solution. Real-world results and industry benchmarks quantify the business value across multiple criteria.

The business value of a test data management solution can be mapped to each phase of the data privacy lifecycle (see Figure 1).





PHASE	BENEFITS
Define	<p>Increase Quality - Define realistic data in QA and development, reducing development, rework and production downtime</p> <p>Increase Testing Productivity - Reduce time it takes to identify optimal test case data, reducing overall testing time</p>
Discover	<p>Mitigate Risk - Avoid breaches, reducing victim notification costs, fines and other costs by identifying sensitive data</p> <p>Accelerate Sensitive Data Discovery - Rapidly identify sensitive data across all legacy and packaged applications and systems, reducing time and costs</p>
Apply	<p>Increase Development Productivity - Develop global masking rules more efficiently through accelerators, pre-built masking techniques, reducing development costs</p> <p>Increase Testing Productivity - Reduce time it takes to identify optimal test case data, reducing overall testing time</p> <p>Hardware and Infrastructure Cost Savings - Subset (create smaller copies of production for test purposes), lowering overall cost of storage. Reduce costs of maintaining network security and other software to secure environments</p> <p>Outsourcing Savings - Because data is masked, companies can then outsource application development or support</p>
Measure and Monitor	<p>Increase Compliance Reporting Productivity - Provide audit team with reports that show what masking policies have been executed, when data was masked, and what it was masked to</p>

Figure 1: The business value of a test data management solution can be mapped to each phase of the data privacy lifecycle.

Validating the approach above with customers and industry analysts, Informatica has created a business value assessment that quantifies the cost savings and avoidance of data breach by using the Informatica solution for test data management. This business value assessment is based on customer testimonials of their cost savings and industry benchmarks such as average cost per record breached.

The business value assessment compares what it would cost for an employee to manually implement a test data management solution to what it would cost to purchase and implement the Informatica solution, including all the time savings achieved by using the Informatica solution.

## The Informatica Solution in Action

Let's take a look at the Informatica solution for test data management in action.

Ochsner Health System relies on the Informatica Platform to streamline patient care, improve patient outcomes, and increase the value of data at lower cost. Ochsner is southeast Louisiana's largest healthcare delivery system, with eight hospitals and more than 38 health centers across the state.

The health system needed to integrate data from more than 38 clinical, scheduling, and billing systems into its new Epic electronic medical record (EMR) system. It needed to support all types of data-intensive projects, which included masking sensitive patient information. As part of rolling out its new EMR system, Ochsner needs to mask medical record information from the production environment in support of secure, compliant testing and development.

The Informatica solution also provides a single, shared view of critical data for business intelligence across the organization, finds and fixes data quality issues, and supports data masking for tests and development. When the integration journey is complete, Ochsner will have every piece of health information available instantly and in real time in one system.

Other Informatica customers are using the solution to mitigate risk in their test environments for cloud-based applications.

In one case, a company transitioning from on-premise to cloud-based HR systems deployed the test data management solution to mask sensitive HR and payroll data in the test environments (see Figure 2). Masking the data allowed this customer to de-identify sensitive data in just two weeks, half as long as planned; as a result, the new hosted HR model launched three weeks ahead of schedule. In addition, masking and protecting sensitive data has allowed this customer and others to realize additional savings by outsourcing development and application support without fear that unauthorized eyes can see the unmasked information.

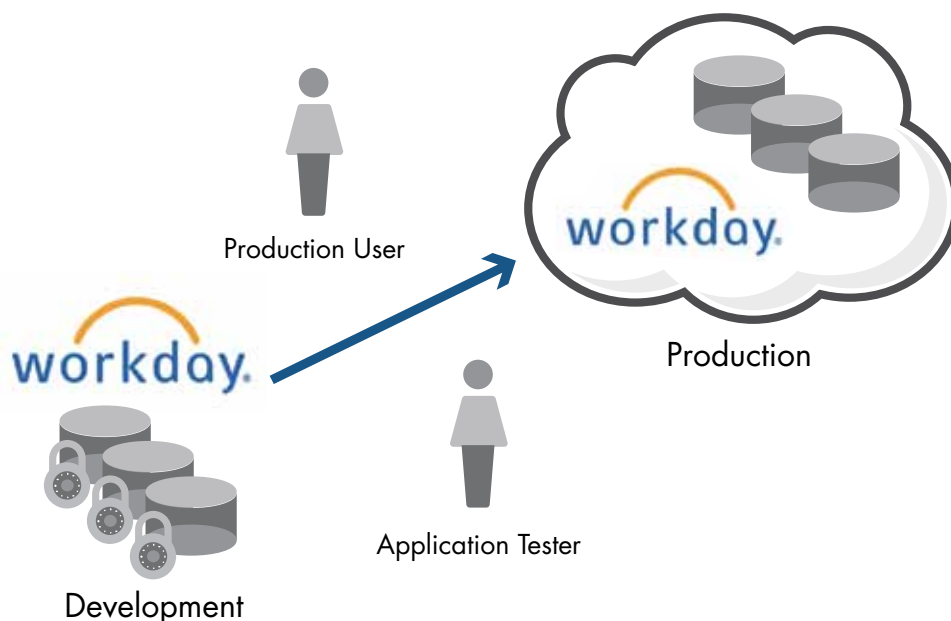


Figure 2: The Informatica solution for test data management created a fully masked data subset of production data, allowing application testers to ensure a smooth transition to a cloud-based HR application without giving them access to sensitive employee information.

## Conclusion

A test data management solution that identifies sensitive data, and creates functional and secure data subsets for testing purposes allows organizations to continue reaping the strategic benefits of big data while minimizing the risk of losing control over it.

The Informatica solution for test data management fulfills those requirements with high performance and comprehensive connectivity while providing an audit trail to support regulatory compliance. Built for scalability, flexibility, and ease of use, the solution enables your IT organization to protect private and sensitive data, decrease the risk of data breaches, and effectively meet compliance requirements on a timely basis—all while decreasing the cost of data, increasing its value, and maximizing its return.

Contact Informatica about performing a test data management business value assessment for your organization today.

## About Informatica

Informatica Corporation (NASDAQ: INFA) is the world's number one independent provider of data integration software.

Organizations around the world rely on Informatica for maximizing return on data to drive their top business imperatives. Worldwide, over 4,630 enterprises depend on Informatica to fully leverage their information assets residing on-premise, in the Cloud and across social networks.



Worldwide Headquarters, 100 Cardinal Way, Redwood City, CA 94063, USA Phone: 650.385.5000 Fax: 650.385.5500  
Toll-free in the US: 1.800.653.3871 [informatica.com](http://informatica.com) [linkedin.com/company/informatica](https://www.linkedin.com/company/informatica) [twitter.com/InformaticaCorp](https://twitter.com/InformaticaCorp)

© 2013 Informatica Corporation. All rights reserved. Informatica® and Put potential to work™ are trademarks or registered trademarks of Informatica Corporation in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks.