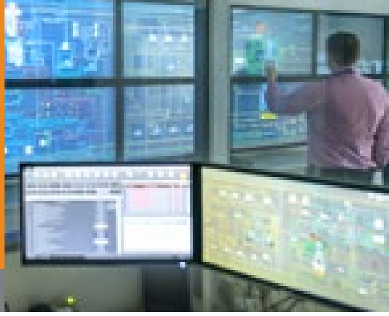
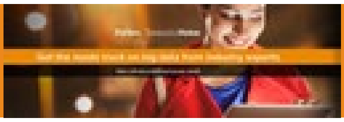
Is Your Big Data Safe?
Beware The Siren's Data Song

CYBER SECURITY



As published in TeradataVoice on Forbes.com



It's no surprise that big data is all the rage. After all, organizations now have within their reach the power to identify and leverage unique customer and operational insights that yield benefits never before possible.

To that end, businesses use a variety of tools and technologies—such as MapReduce functions and Hadoop—for data storage to derive valuable business insights. These tools originated as open source technologies designed for use by an emerging group of analysts, called data scientists. In contrast with traditional analysts, who begin with questions and use data to arrive at answers, data scientists explore data to learn what new questions can be answered.

It's All in the Data, Risk Included

Increasingly available through new data management and analytic technologies, big data is often thought of as large data volumes created from ever-emerging new sources. However, big data is broader than its volume, and can be further characterized by its variety, velocity and complexity. Its complexity, in particular, is derived from non-traditional sources, such as image, text and voice data, in addition to machine-generated, non-relational and semi or unstructured data from Web servers, sensors, and telematics or interactional data from social networks.

That's a lot of data.

And while having access to its breadth and depth can be empowering, it can also expose your organization to potential breaches and lost revenue if not properly safeguarded.

Data Sirens: Beautiful But Dangerous

The prospect of pooling disparate data from a broad range of sources and formats, and then discovering new relationships or patterns in the data, is very appealing. But while the seemingly endless possibilities are alluring, big data projects can have unintended results.

One shortcoming of aggregating data into a single location is the risk of creating a super high-value target for adversaries—both external and internal—whose goal is to steal and disseminate data for financial gain, political purposes, or revenge.

> One shortcoming of aggregating data into a single location is the risk of creating a super high-value target for adversaries—both external and internal—whose goal is to steal and disseminate data for financial gain, political purposes, or revenge.

Another consideration are the risks associated with transforming otherwise innocuous, safe or low-value data into high-value sensitive data that may now be classified as personally identifiable information, healthcare information, or other sensitive data subject to security and privacy regulations.

Start-ups and mature organizations alike sometimes err on a strategy to build the business and create value fast, then implement a conventional security infrastructure later. While intentions may be good, the rush to monetize discoveries or secure cash flow can blind one to what can go wrong along the way—like a massive, costly data breach.

Big data projects without proper security are not only critically unsafe, but also possibly non-compliant with a variety of data security and privacy regulations.

The good news is that conventional security best practices can be applied to big data, while serving data scientists equally well. The first step is to assess your current security state.

Dusting Off Tried-And-True Security Best Practices

The good news is that conventional security best practices can be applied to big data, while serving data scientists equally well. The first step is to assess your current security state. This is best accomplished by using an external organization to perform a compliance audit or, if not regulated, a risk assessment based on an authoritative standard. This process can highlight situations that expose your organization to data breach risks, and reveal necessary steps for reducing that risk to acceptable levels.

Here are some recommended actions you can take right now to reduce risk:

Hardening operating systems. Modern operating systems provide end users with a wide variety of configurable options, many with security implications. Security professionals can recommend and implement configuration options to optimize the security and utility of the operating systems underlying the big data environment.

Tightening up user authentication and authorization. Authentication, typically a user ID and password, confirms or authenticates the user. Authorization is what the authenticated user is allowed to do. These standards should be applied to big data environments across the organization, including discovery projects. Applying rigorous access controls helps safeguard sensitive data resulting from the aggregation of seemingly disparate and non-sensitive data.

Implementing the concept of least privilege. Least privilege means users are only granted access to systems they need to accomplish their job-related tasks, and provided the minimum level of read, write and execute permissions on those systems. Pooling large volumes of data without controlled access is an accident waiting to happen.

Implementing conventional security practices will not eliminate all risk of data breach or loss, but it will put you in the driver's seat for minimizing risk exposure within your organization.

Big data, along with properly implemented security practices, can yield exciting new insights for better customer experiences, enhanced loyalty and retention. It can also help reduce losses due to fraud and financial crimes, and lead to powerful discoveries that support your organizational objectives in new and innovative ways.

Teradata Cyber Defense Solutions

Sam Harris

Director, Cyber Security

+1 (919) 341-2463

sam.harris@Teradata.com

www.linkedin.com/in/samharris

@samuellharris

10000 Innovation Drive, Dayton, OH 45342 **Teradata.com**

5.15 EB8895

Forbes TeradataVoice