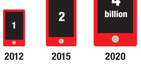
Securing the Big Data Life Cycle

There's a Lot of Data Out There... And It's Growing Exponentially

The big data phenomenon is a direct consequence of the digitization and "datafication" of nearly every activity in personal, public, and commercial life.

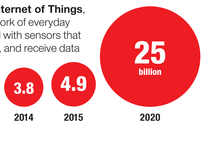
Almost a quarter of the world's population now uses **smartphones**. Data from all those calls and apps are being warehoused and mined.

Global Smartphone Users in billions



Add to that the **Internet of Things**, the growing network of everyday objects equipped with sensors that can record, send, and receive data over the Internet.

Growth in Internet-Enabled Things in billions



All That Data Is at Risk

Security breaches affect organizations of all sizes across all industries. Just a few years ago, an individual breach affected 1 million to 10 million records. Today, in the age of mega-breaches, a single incident can involve 200 million records—or more.

Survey of Security Practitioners* at midsize-to-large companies, 2014

The Verizon 2015 Data Breach Investigations Report (DBIR)

Big Data Comes With Big Responsibility for Consumer Privacy

Don't compromise customers' privacy when collecting big data. Avoid these pitfalls (and the resulting regulatory consequences):

Ubiquitous and indiscriminate data collection from a wide range of devices

Unexpected uses of collected data, especially without explicit customer consent

Unintended data breach risks with larger consequences

Security Concerns

Common Mistakes

Companies focus too much on reaping the potential value of big data and too little on sufficiently securing that information. They suffer from:

OUTDATED APPROACHES

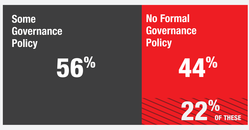
Securing the network perimeter is not sufficient. Fewer than 1 percent of breaches were detected using perimeter security controls such as switches, firewalls, and routers. Yet two-thirds of security budgets are used to protect the network, leaving little for data and intellectual property protection.

SECURITY BUDGET ALLOCATIONS

INSUFFICIENT GOVERNANCE

Unable to manage the availability, integrity, and security of enterprise data, organizations struggle to meet privacy and regulatory mandates. Yet 44 percent of organizations have no formal data governance policy, and 22 percent of those firms have no plans to implement one.

DATA GOVERNANCE POLICIES



Concerns About Hadoop

Hadoop wasn't built with security in mind. However, it is widely used and faces the following threats:

UNAUTHORIZED ACCESS

Built to make all data accessible by all users, Hadoop lacks permissioning, password controls, or auditing, and cannot comply with rigorous compliance standards like HIPAA and PCI DSS.

DATA PROVENANCE

In traditional Hadoop, determining where a particular dataset originated has been difficult, creating potential "garbage-in, garbage-out" problems. Business decisions may be based on suspect or compromised data.

DIY HADOOP

Do-it-yourself Hadoop clusters introduce risk as it scales from small to enterprise-wide projects. At every stage of growth, data verification and user management become more difficult, leaving fewer resources available to address security and stability.

Big Data with Security in Mind

Defense-In-Depth Security

Defense-in-depth security safeguards against malicious attacks and protects information assets by ensuring only authorized access. In the age of big data, organizations need to employ three layers of security controls:

PREVENTIVE

Secure the data itself with controls such as encryption of data at rest and in motion, redaction of data in applications, and use of identity and access management.

DETECTIVE

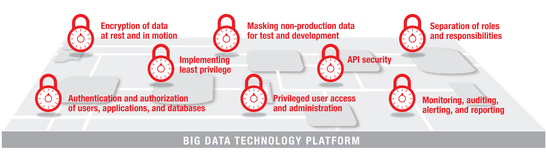
Detect and alert on suspicious behavior after and during an incident by auditing operating systems, Hadoop services, and network activity. Follow up with appropriate reports for auditing and regulatory compliance.

ADMINISTRATIVE

Implement tools that enable the processes and procedures for security, such as sensitive data discovery, privileged user analysis, configuration management, and encryption key management.

The Case for a Big Data Technology Platform

Big data works best in an environment that integrates Hadoop, NoSQL, and relational databases onto a single, secure platform that makes data and all critical business systems securely available across the organization, with proper governance in place.



Oracle for Enterprise Big Data

ORACLE

With Oracle, organizations can take advantage of the future of big data while preserving the value of existing investments in technology and skills.

Remove the barriers between Hadoop, NoSQL, and relational databases — in the cloud and on-premises.

Quickly discover and predict real-world patterns in data — and apply those insights immediately.

Simplify data access across the big data environment — including third-party data from customers and partners, mobile apps, and connected devices.

Achieve all the potential of big data — with a comprehensive security approach that gives the right people the right access at the right time.

Bringing It All Together

There's no question: companies that use big data effectively can gain significant competitive advantage. But despite big data's many benefits, organizations are exposing their sensitive information to increased risk as they integrate open-source Hadoop into their IT environments.

That's why companies serious about using big data effectively need to make sure they're doing so securely, protecting their valuable information and securing private data so that it stays private.

For more information about about securing big data, visit www.oracle.com/bigdata.

* For source information, see the MIT Technology Review Custom and Oracle white paper **"Securing the Big Data Life Cycle"**: http://www.oracle.com/us/technologies/big-data/securing-big-data-life-cycle-2543085.pdf