Brought to you by the publishers of COMPLIANCE WEEK

INSIDE THIS PUBLICATION:

The Evolving Era of Big Data

An e-Book publication sponsored by acl transforming audit and risk hp

Inside this e-Book:

COMPLIANCE WEEK

Compliance Week, published by Wilmington Group plc, is an information service on corporate governance, risk, and compliance that features a weekly electronic newsletter, a monthly print magazine, proprietary databases, industry-leading events, and a variety of interactive features and forums.

Founded in 2002, Compliance Week has become the go-to resource for public company risk, compliance, and audit executives; Compliance Week now reaches more than 60,000 financial, legal, audit, risk, and compliance executives.

acl
transforming audit and risk

ACL delivers technology solutions that are transforming audit, compliance, and risk management. Through a combination of software and expert content, ACL enables powerful internal controls that identify and mitigate risk, protect profits, and accelerate performance.

Driven by a desire to expand the horizons of audit and risk management so they can deliver greater strategic business value, we develop and advocate technology that strengthens results, simplifies adoption, and improves usability. ACL's integrated family of products—including our cloud-based governance, risk management, and compliance (GRC) solution and flagship data analytics products—combine all vital components of audit and risk, and are used seamlessly at all levels of the organization, from the C-suite to front line audit and risk professionals and the business managers they interface with. Enhanced reporting and dashboards provide transparency and business context that allows organizations to focus on what matters.

And, thanks to 25 years of experience and our consultative approach, we ensure fast, effective implementation, so customers realize concrete business results fast at low risk. Our actively engaged community of more than 14,000 customers around the globe—including 89% of the Fortune 500—tells our story best. Visit us online at www.acl.com

HP Security Voltage

HP Security Voltage is a world leader in data-centric encryption and tokenization. HP Security Voltage provides trusted data security that scales to deliver cost-effective PCI compliance, scope reduction and collaboration security. HP Security Voltage solutions are used by leading enterprises worldwide, reducing risk and protecting brand while enabling business. For more information see www.voltage.com.

Protecting the World's Most Sensitive Data

» As data has become an increasingly valuable corporate asset, hackers and data thieves continue their relentless drive to thwart protection measures. A serious data breach causes immeasurable damage to corporate reputation. HP Security Voltage helps businesses, such as financial institutions and consumer transacting businesses, comply to specific regulations that mandate data protection.

» Today enterprises must have a complete data protection strategy, using proven encryption, tokenization and data de-identification approaches, that protect the data itself, not just the data containers and network perimeters. HP Security Voltage is the leading expert in data-centric encryption and tokenization.

» HP Security Voltage was created to solve the industry's biggest problem—making encryption of data simple for even the most complex use cases. HP Security Voltage has solved this challenge through cryptographic innovations, specifically HP Identity Based Encryption and HP Format-Preserving Encryption.

» Today HP Security Voltage has over 1,000 enterprise customers and protects sensitive data at the world's largest banks, financial institutions and payment processors.

CCOs Play a Stronger Role in Data Privacy

By Aarti Maharaj

As data privacy laws proliferate around the world, they are creating a web that traps how corporations use personal data in their operations. The challenge for compliance officers: how to play a more strategic role in the organization, ensuring your business doesn't get stuck.

So far that effort hasn't been easy. In the Compliance Trends 2015 report published by Compliance Week and Deloitte, 59 percent of compliance officers are either "somewhat confident" or "not confident at all" that their IT systems can fulfill the data collection and reporting requirements they have. That can cause problems in how your business gathers data, how it uses data, and even how the business recovers from regulatory and reputation risks when it *loses* data, through hackers or otherwise.

"The issue for chief compliance officers is that they are increasingly struggling to connect regulatory requirements to IT issues," says Todd Cipperman, founding principal of Cipperman Compliance Services. "Technology is its own discipline and I don't see compliance officers becoming technologists overnight."

The forward march of technology—specifically, data storage in the cloud—does chief compliance officers few favors. According to research firm Gartner, by 2017 50 percent of an organization's business data will reside outside the physical walls of your corporate data center, up from less than 10 percent today. According to Eurostat, the statistical office of the European Union, about 20 percent of enterprises will rely on cloud computing across the Organization for Economic Co-operation and Development. (Finland is currently leading the race at 50 percent, Poland in the rear at 6 percent.)

The problem is that few compliance officers are involved in high-level discussions around cloud computing and data privacy controls, which can be disastrous for companies as they expand into new locations.

"Employee data is becoming something on the forefront of compliance," says Marie Blake, executive vice president and chief compliance officer at BankUnited N.A. "To avoid an in-house privacy breach, you have to think about what is included in privacy data, like information governance, and you have to move with the direction of the industry."

Currently the industry is moving to play catch up with the risk. One example is the massive breach of financial institutions at JPMorgan and several other large banks last year, where prosecutors and IT security reportedly are still sizing up exactly how the attacks happened and how widespread the damage was (tens of millions of customer records, at least).

IT security tools will help that threat, but often tools address one specific risk. If the process for governing information is weak overall, that leaves the company exposed to any number of other risks your IT security tools don't address. And the moves in Europe and elsewhere around the world to strengthen data privacy laws makes that need for information governance all the more acute.

"From a compliance perspective you want to put policies in place to defend a claim," Cipperman says. "This is sometimes hard for compliance officers to do because it's not easy to understand what are the data security operations in place."

Pressure on CCOs

Janet de Guzman, director of compliance at OpenText, an enterprise information management firm, says, "The CCO mandate is to be aware of and ensure their organization adheres to the laws and regulations relevant to their business. Therefore they should absolutely be at the table for discussions around technology. Data and privacy protection is becoming a critical part of the compliance function because it's not only their own data at stake but the data of their customers and other stakeholders."

Blake says that not many CCOs are involved in their companies' data privacy committees, and she expects that to change over time as companies realize that CCOs bring critical knowledge about regulatory requirements to their cyber-security discussions.

"The inclusion of the CCO function in defining controls related to things like cloud computing has yet to hit maturity," Blake said. She compared it to vendor management,

Continued on Page 15

EU ON DATA PROTECTION REFORM

The following is an excerpt from the European Council on data protection reform.

In the last few decades, the European Union has adopted several pieces of legislation to protect personal data, the main one being the 1995 data protection directive. However, since the Lisbon Treaty, protection of personal data has been a fundamental right under EU law, recognized by the Treaty on the Functioning of the European Union and the EU Charter of Fundamental Rights. This means the Union now has a specific legal basis to adopt legislation to protect this fundamental right.

Rapid technological developments in the last 2 decades have brought new challenges for the protection of personal data. The scale of data sharing and collecting has grown exponentially, sometimes taking place on a global level, and individuals are increasingly making personal information publicly available. The economic and social integration resulting from the functioning of the internal market has also led to a substantial increase in cross-border flows of data. To take all these developments into account and promote the digital economy, there is a need to ensure a high level of protection of personal data, while at the same time allowing for the free movement of such data within the European Union.

In the case of personal data used for law enforcement purposes, there is a growing need for authorities in the member states to process and exchange data as part of the fight against transnational crime and terrorism. In this context, clear and consistent rules on data protection at EU level are fundamental to improving cooperation between those authorities.

Source: European Council.

Facebook's Big Data Fail Calls for More Ethics

By Joe Mont

A flap over a controversial use of Big Data analysis techniques by Facebook has once again spurred calls for the more ethical use of data-gathering tools by companies.

In July 2014, Facebook drew fire for an "experiment" that studied how users' news feeds could be manipulated to affect their moods. The so-called "emotional contagion experiment" involved nearly 700,00 users, all of whom, according to Facebook, agreed to the privacy parameters set forth in its terms-of-use agreement that allow such tinkering.

With an apology, Facebook insisted it had considered the ethics of conducting the project. "The reason we did this research is because we care about the emotional impact of Facebook and the people that use our product," Adam Kramer, one of the Facebook researchers involved in the study, said in a statement.

The research and the reaction to it, however, illustrate a dilemma companies have faced since the early years of the Internet: How do we use the powerful data-gathering abilities that the online environment affords without trampling on the privacy of customers and others. "Just because we can do something, doesn't mean we should do it," warns Deborah Johnson, a professor of applied ethics in the School of Engineering and Applied Sciences at the University of Virginia.

Unfortunately, many companies on the cutting edge of social media and Big Data give into the rationale "If the law doesn't tell me I can't, why shouldn't I?" says Johnson. "The way the information world has developed has been a free for all."

The episode may prove to be a tipping point for the consideration of ethics in Big Data. Moving forward, companies will need to give greater consideration to ethics when finding new and creative ways to collect and parse data. Facebook itself, through Kramer's apology, acknowledged as much. "While we've always considered what research we do carefully, we have been working on improving our internal review practices," he said. "The experiment in question was run in early 2012, and we have come a long way since then. Those review practices will also incorporate what we've learned from the reaction to this paper."

The power of Big Data—piecing together clues that reveal more about the personal lives of customers and potential customers—should not be treated cavalierly. Many companies have created privacy committees and installed privacy officers. This is important to insulate the brand from negative repercussions and associations.

Developing a Process

Neil Richards, professor of law at Washington University School of Law and co-author of a research paper, "Big Data Ethics," says: "Companies have to, at least, take the ethical perceptions of what they are doing into account in the short term. If they do something perceived as outrageous, they are going to suffer a short-term business hit. In the long term, developing a process for ethical data usage is essential to ensure productive and profitable relationships with customers, users, and business partners."

Companies will need to give greater consideration to what they plan to do with the data, even before they begin collecting it. "We are starting to realize that, when it comes to data, the era of digital strip mining is over," Richards adds. "We can't just, as companies, exploit for the immediate short-term gain. Things we do in the short term have long-term consequences."

Continued on Page 14

A SENATOR'S QUESTION TO THE FTC

The following is from a letter written by U.S. Sen. Mark Warner (D-Va.) requesting that the Federal Trade Commission provide more information on recent reports that the social network Facebook conducted an experiment involving nearly 700,000 users to study the emotional effect of manipulating information on their news feeds.

While Facebook may not have been legally required to conduct an independent ethical review of this behavioral research, the experiment invites questions about whether procedures should be in place to govern this type of research.

I am not convinced that additional federal regulation is the answer. Public concerns may be more appropriately addressed through industry self-regulation. As the federal regulator with oversight of privacy and consumer protection policies, I would be interested in your responses to the following questions:

Does the FTC have a role to play in improving transparency, accountability, and consumer trust in industry's use of Big Data?

Are there better ways to educate consumers or otherwise improve transparency, about the practices consumers agree to through their use of social media platforms? Are there incentives in place-for companies to voluntarily create, or to consult with independent review boards, or to utilize other means of self-regulation before conducting studies such as this? Additionally, are there incentives that could encourage the hiring or designation of chief privacy officers at social media companies, or to establish other credible review programs?

Does the FTC make any distinction between passively observing user data versus actively manipulating it? Should consumers be provided more of an explicit opt-in or opt-out of such studies? Additionally, is it appropriate for any research findings to be shared with participants prior to public dissemination?

Does the FTC or another federal entity require any additional regulatory authority or technology in order to monitor this type of data-mining?

Source: Sen. Mark Warner.

The Big Data Opportunity for Audit, Risk Management, and Compliance

By John Verver, CPA CA, CISA, CMC, Strategic Advisor to ACL

Much has been written and discussed about big data in recent years, often focusing on the techniques and concepts that can help businesses better understand their markets and drive new revenues and more profitable opportunities. Big data includes traditional business sources of structured data, such as the millions of daily transactional records from retail, financial, manufacturing, or transportation/logistics industries. And now there's also non-traditional or unstructured data that might come from social media sources, human responses, or physical sensory recordings, for instance.

Of course, along with the benefits of improved market understanding and decision making arising from new data insights, increasing attention is also being paid to the risks and complexities of managing big data. These include maintaining security, ensuring compliance with privacy and other regulatory requirements, as well as dealing with the multiple challenges of maintaining and working with massive volumes of data. While auditors and those involved in risk management and compliance have their respective roles to play in addressing the risks of big data, many of the concepts of big data also provide great opportunities for transforming the way that audit, risk, and compliance professionals perform their work.

Not so new for audit and compliance professionals

In fact, many of the fundamental techniques and concepts of big data have much in common with those that leading audit and compliance teams have been applying for decades. The use of data analysis to support the audit process typically involves obtaining entire populations of data, often from a variety of sources and databases. All that data is analyzed in order to gain new insights into risks and to identify fraud, error, abuse, and many other forms of internal control and compliance failures.

Increasingly, auditors are now looking for new and innovative sources of data—both internally and externally generated—in order to broaden and deepen the search for risk and compliance issues. These sources can include, for example, e-mails and the content of social media and other sources of relatively unstructured data. There is considerable overlap between these approaches and those that underlie big data processes.

> *Increasingly, auditors are now looking for new and innovative sources of data—both internally and externally generated—in order to broaden and deepen the search for risk and compliance issues.*

Big data techniques across the spectrum of audit, compliance, and controls

At present, much of the use of data analysis in audit is relatively basic. Approximately a third of audit departments currently use data analysis extensively within audit procedures. However, within the profession, there is broad agreement that data analysis should be used far more widely and should form the basis for transforming audit procedures.

Increasingly, there is also agreement that the use of data analysis by auditors should be performed in conjunction with usage by those involved in risk management and compliance, as well as those who have direct daily responsibility for maintaining controls and compliance in operational and financial systems. The Institute of Internal Auditors refers to the *Three Lines of Defense*[1] model as a way of recognizing the respective responsibilities of auditors, controls and compliance professionals, and operational management. Data analysis has a role to play in each one of these areas as a means of helping to ensure and improve the effectiveness of compliance procedures—both in terms of regulatory compliance as well as compliance with internal controls and procedures. Data analysis not only enhances the audit of what has already occurred, but also enables improved risk management by providing insights into trends and what is likely to occur in the future.

Driving continuous operational risk monitoring and assessment

Much of the use of data analysis for audit, risk, and compliance

[1] Institute of Internal Audits, "Three Lines of Defense in Effective Risk Management and Control"

commences with procedures that are relatively ad hoc: often one-off explorations or profiling of data in order to determine risk exposure and identify compliance problems around a specific business process area. In most cases, the end goal is to perform similar procedures (once proven to be of value) on an ongoing sustained basis, using automated techniques. In this way, data analysis enables continuous risk assessment and automated operational risk management that can be used by various stakeholders including senior management, audit and risk committees, and even external regulatory and audit organizations.

Practical implementation

There are common practical stages in applying big data concepts to audit, risk, and compliance:

1. Data acquisition

The identification of the most appropriate data to support a specific risk and compliance analysis procedure is one of the most critical stages—followed by the gathering of that data from either structured or unstructured sources. Multiple data sources should be considered, often because the most effective analytic procedures involve comparing data from a variety of systems in ways that do not normally occur. Appropriate, platform-independent technology is essential for providing rapid connectivity to a wide range of data types and structures.

> *There are many issues to consider in ensuring big data does not become a risk that damages an organization. However, big data techniques themselves also represent a great opportunity to enhance and transform the core processes of audit, risk management, and compliance.*

2. Analysis procedures

Typically, a variety of analysis types are performed in order to achieve two objectives. The first objective involves examining entire populations of transaction data for a business process in order to determine if each transaction complies with a specific internal control or regulatory requirement. The second involves examining the same data in order to determine if there are indications of risks or compliance failures for which no control has been established. The end result is often to implement a suite of tests, ranging from tens to hundreds of specific analytics, which can be applied as needed.

An important characteristic of technology used for these forms of data analytics is the ability to automatically create secure logs of all procedures performed. Such logs are a form of documentation that's often essential for audit and regulatory compliance procedures.

3. Dealing with massive data volumes

Managing the data volumes and processing requirements for audit, risk, and compliance analytics sees similar challenges as for any big data application. However, in practice, the data sets involved, though large in terms of record volumes, typically only include a relatively small number of the total data elements that exist within corporate and other databases. Maintaining security and control over data used for audit and risk purposes is clearly as important as for any dataset including critical, sensitive, and valuable information.

4. Automation and continuous risk monitoring

Once a library of data analytics and compliance tests have been established, a decision needs to be made as to how often they should be run. In some cases the best alternative may be to run tests on a periodic basis of months or quarters in order to identify anomalies. However, it often makes sense to run tests for critical exposures on a more frequent regular basis, either daily or even close to real-time. The most significant issue for performing regularly scheduled continuous monitoring is determining responsibilities for responding to exceptions and addressing control and compliance risks.

Transforming not just the organization, but also risk management and compliance processes

Big data is a topic that is likely to only get bigger as businesses and other organizations increase their ability to apply analysis techniques to obtain many new forms of insights and understanding. From a perspective of audit, risk management, and compliance, there are many issues to consider in ensuring big data does not become a risk that damages an organization. However, big data techniques themselves also represent a great opportunity to enhance and transform the core processes of audit, risk management, and compliance, in the same way that big data can benefit organizations overall and help them to better achieve their strategic objectives.

About the author

John Verver, CPA, CISA, CMC, Strategic Advisor to ACL

John Verver is an acknowledged thought leader, writer and speaker on the application of technology for audit, fraud detection, risk management and compliance. He is recognized internationally as a leading innovator in continuous controls monitoring and continuous auditing and as a contributor to professional publications. He is currently a strategic advisor to ACL, where he has also held vice president responsibilities for product strategy, as well as ACL's professional services organization. Previously, John was a principal with Deloitte in Canada.

www.acl.com

Auditing in the Era of Big Data

By Tammy Whitehouse

Get ready for the era of Big Audit.

The auditing profession is starting to look at how to leverage Big Data in audits with some big investments into cutting-edge data analytics that could dramatically deepen the reach of external auditors into corporate books and records.

The push into the next generation of auditing, however, is mired in regulatory and legal complexity that promises to bog down the transformation.

Under the future vision of auditing, public companies would give auditors access not just to a sample of their transactions, but to their entire general ledger and their databases. "With these tools, auditors will have the capability to look at the underlying data, not just the summary data," says Brian Fox, president of audit services firm Confirmation.com. "They will want all the transactional data, even if it's millions or tens of millions of records. It will be a different conversation."

But don't expect the transition to happen overnight. External auditors have been slower than others to jump on the Big Data bandwagon, says Kelly Todd, shareholder with audit firm Forensic Strategic Solutions, which uses data analytics to conduct investigations. It's a massive leap to go from traditional audit approaches, which are based on sampling transactions, to an audit that would look at literally everything.

Todd

"The reality is with data analytics, you have the ability to look at 100 percent of the transactions," she says. "You can see the footprint of the beast, the unusual patterns, and the things that don't make sense."

As analytic technology becomes more readily available and as audit firms take a beating from regulators and the capital markets over audit quality issues, external auditors are turning to Big Data for answers.

"Virtually all of the Big 4 and others have very sizable projects around transforming the effectiveness of external audit procedures through the use of technology," says John Verver, vice president of audit technology firm ACL. "They're focused on the quality of the audit, reducing the risk, and improving the efficiency and effectiveness."

Auditors have long used analytical procedures in their audit work, says Dorsey Baskin, managing partner at Grant Thornton and a member of the assurance services executive committee of the American Institute of Certified Public Accountants. Existing auditing standards require the use of analytical procedures to plan the audit and to wind up an engagement, or to perform "the smell test" at the end of the audit, he says. The kind of data analytics firms are now investigating are far more complex, he says. "The audit tool kit looks the same as it did 50 or 60 years ago," he says. "If we were doctors, that would be pretty frightening. This has tremendous potential, but it's still early. We're still experimenting."

Deloitte & Touche, for example, says it is looking at the potential to leverage tools in three different ways. The first, says Joseph Ucuzoglu, national managing partner, is to audit large or complete sets of data, rather than just samples of data. The second is to leverage artificial intelligence to search not just data but also text, looking for red flags and tell-tale terms. The third area is to look beyond the data a company produces to examine data available elsewhere.

The firm is developing a series of trials and testing them on a small scale to assure the techniques work, Ucuzoglu says. "We are still doing traditional audit techniques, but once we prove the concept, we can take it to a larger stage," he says. The technology won't replace human auditors, but will remove the rote tasks, examine more data, and give auditors better information to consider, he says. "It will free up professionals to spend their time on the highest value areas," Ucuzoglu says. That aspect is actually exciting to auditors, he says, who are looking for ways to engage and retain more young talent in the profession.

Audit firms are tooling up for the transformation in a variety of ways, not the least of which is buying consulting businesses where the technology and the analytical skills reside. The Public Company Accounting Oversight Board, however, has expressed some concern over the

Continued on Page 15

AUDIT DATA STANDARD WORKING GROUP

Below the AICPA's Audit Data Standard Working Group provides information on its voluntary, uniform audit data standards.

ASEC's Emerging Assurance Technologies Task Force established the Audit Data Standard working group to help develop new technologies that will contribute to the effectiveness, timeliness, and efficiency of the audit process. One of its main projects is developing a standardized data model that management, internal auditors, and external auditors could utilize for enhanced analytics that would further improve the timeliness and effectiveness of the audit process.

One of the challenges that management and auditors face is obtaining accurate data in a usable format following a repeatable process. As a result, the working group has developed a voluntary, uniform audit data standards that identifies the key information needed for audits and provides a common framework covering: (1) data file and field definitions and technical specifications, and (2) supplemental questions and data validation routines to help auditors better understand the data and assess its completeness and integrity The standards are offered in two file formats: (1) flat file format (pipe-delimited UTF-8 text file format) and (2) eXtensible Business Reporting Language Global Ledger Taxonomy Framework (XBRL GL).

The first issuance of the Audit Data Standards includes: Base Standard, General Ledger Standard, and Accounts Receivable Subledger Standard.

Source: AICPA.

The bottom line

What: Hadoop's big benefit also provides opportunities for attacks.

Why: Centralized data can give hackers easier access.

How: Most importantly, protect data before it's ingested.

Hadoop's Potential Also Comes With Big Security Questions

New Challenges

Today, companies have implemented every type of deterrent, policy, training, intrusion prevention, and firewall, but it is not enough. New challenges exist because businesses are increasingly driving initiatives that push sensitive data into more business areas.

Hadoop is ground zero for the battle between the business and security. The business needs the scalable, low-cost Hadoop infrastructure so it can take analytics to the next level—a prospect with myriad efficiency and revenue implications. Yet Hadoop includes few safeguards, leaving it to organizations to add an enterprise security layer.

Security cannot afford to lose this fight: Implementing Hadoop without robust security in place takes risk to a whole new level. But armed with good information and a few best practices, enterprise security leaders can put an end to the standoff and ensure attackers will glean nothing from their attempts to breach Hadoop.

Risk, reinvented

"Hadoop is the biggest cyber-crime bait ever created," says Reiner Kappenberger of global product management for HP Security Voltage. "In the past, attackers had to map the network and go to a lot of work and expense to find the data they wanted to retrieve. With Hadoop, organizations consolidate all their information into a single destination, making it very easy to find all the information criminals want—and more."

It isn't just the size of the bait that makes Hadoop breaches so treacherous. Hadoop environments are inexpensive to replicate and require no prior knowledge of the data schema used. In just a few days, terabytes of data can be siphoned up and replicated elsewhere.

An expert offers 5 steps to take back control

Market solutions for Hadoop security are beginning to emerge, delivering data masking features that make it possible to obscure sensitive data. But whether you leverage a commercial solution or create a homegrown approach, Kappenberger suggests the following five steps to identify what needs protecting and apply the right techniques to protect it—before you put Hadoop into production.

1. Audit and understand your Hadoop data.

To get started, take an inventory of all the data you intend to store in your Hadoop environment. You'll need to know what's going in so you can identify and rank the sensitivity of that data. It may seem like a daunting task, but attackers can take your data quickly and sort it at their leisure. If they are willing to put in the time to find what you have, you should be too.

2. Perform threat modeling on sensitive data.

The goal of threat modeling is to identify the potential vulnerabilities of at-risk data and to know how the data could be used against you if stolen. This step can be simple: For example, we know that personally identifiable information always has a high black market value. But assessing data vulnerability isn't always so straightforward. Date of birth may not seem like a sensitive value alone, but when combined with a zip code, a date of birth gives criminals a lot more to go on. Be aware of how various data can be combined for corrupt purposes.

3. Identify the business-critical values within sensitive data.

It's no good to make the data secure if the security tactic also neutralizes its business value. You'll need to know if data has a characteristic that is critical for downstream business processes. For example, certain digits in a credit card number are critical to identifying the issuing bank, while other digits have no value beyond the transaction. By identifying the digits you need to retain, you can be sure to use data masking and data encryption techniques that make re-identification possible.

4. Apply tokenization and format-preserving encryption on data as it is ingested.

You'll need to use one of these techniques to protect any data that requires re-identification. While there are other techniques for obscuring data, these are particularly suited for Hadoop because they do not result in collisions that prevent you from analyzing data. Each technique has different use cases; expect to use both, depending on the characteristics of the data being masked. Format-preserving technologies enable the majority of your analytics to be performed directly on the de-identified data, securing data-in-motion and data-in-use.

With Hadoop, you must protect sensitive data before it is ingested. Once data enters Hadoop it is immediately replicated inside your cluster, making it impossible to protect after the fact. By applying your tokenization and format-preserving data encryption during the ingestion process, you'll ensure no traces of vulnerable data are floating around your environment.

5. Provide data-at-rest encryption throughout the Hadoop cluster.

As just mentioned, Hadoop data is immediately replicated on entering the environment, which means you'll be unable to trace where it's gone. When hard drives age out of the system and need replacing, encryption of data-at-rest means you won't have to worry about what could be found on a discarded drive once it has left your control. This step is often overlooked because it's not a standard feature offered by Hadoop vendors.

Timing is everything

The perfect time to undertake this process, says Kappenberger, is after you've done a pilot and before you've put anything into production.

"Because they've done the pre-work, companies at this stage know what they want. They understand their queries, and adding the format-preserving encryption and tokenization to the relevant fields can be done very easily," he says. In fact, it can take just a few days to create a proof of concept.

Data-centric Security

The obvious answer for true Hadoop security is to augment infrastructure controls with protecting the data itself. This data-centric security approach calls for de-identifying the data as close to its source as possible, transforming the sensitive data elements with usable, yet de-identified, equivalents that retain their format, behavior, and meaning. This protected form of the data can then be used in subsequent applications, analytic engines, data transfers, and data stores, while being readily and securely re-identified for those specific applications and users that require it.

For Hadoop, the best practice is to never allow sensitive information to reach the HDFS in its live and vulnerable form. De-identified data in Hadoop is protected data, and even in the event of a data breach, yields nothing of value, avoiding the penalties and costs such an event would otherwise have triggered.

The Solution–HP SecureData for Hadoop

HP Security Voltage provides maximum data protection with the HP SecureData for Hadoop, with industry-standard, next-generation HP Format-Preserving Encryption (FPE), (see NIST SP-800-38G) and HP Secure Stateless Tokenization (SST) technologies.

With HP SecureData FPE and SST, protection is applied at the data field and sub-field level, preserves characteristics of the original data, including numbers, symbols, letters, and numeric relationships such as date and salary ranges, and maintains referential integrity across distributed data sets so joined data tables continue to operate properly. HP FPE and SST provide high-strength encryption and tokenization of data without altering the original data format.

De-identified data in Hadoop is protected data, and even in the event of a data breach, yields nothing of value, avoiding the penalties and costs such an event would otherwise have triggered.

HP SecureData encryption/tokenization protection can be applied at the source before it gets into Hadoop, or can be evoked during an ETL transfer to a landing zone, or from the Hadoop process transferring the data into HDFS. Once the secure data is in Hadoop, it can be used in its de-identified state for additional processing and analysis without further interaction with the HP Security Voltage system. Or the analytic programs running in Hadoop can access the clear text by utilizing the HP Security Voltage high-speed decryption/de-tokenization interfaces with the appropriate level of authentication and authorization.

If processed data needs to be exported to downstream analytics in the clear—such as into a data warehouse for traditional BI analysis—there are multiple options for re-identifying the data, either as it exits Hadoop using Hadoop tools or as it enters the downstream systems on those platforms.

To implement HP Security Voltage data-centric security requires installing the HP SecureData infrastructure components and then interfacing with the appropriate applications and data flows. SDKs, APIs, and command line tools enable encryption and tokenization to occur natively on the widest variety of platforms, including Linux, mainframe, and mid-range, and supports integration with a broad range of infrastructure components, including ETL, databases, and programs running in the Hadoop environment. HP Security Voltage has technology partnerships with Hortonworks, MapR, Cloudera and IBM, and HP SecureData for Hadoop is certified to run on each of these. In addition, HP SecureData protects sensitive data on HP Vertica, Teradata, and other Big Data platforms.

For more information please visit www.voltage.com/hadoop

Next Up: Regulating the ‘Internet of Things’

Regulators struggle with controlling what data is being collected, how it is aggregated, and how it’s being used

By Joe Mont

Despite its rather inelegant name, the “Internet of Things” is revolutionizing the business world and presenting regulators with some weighty challenges.

The “things” in question are consumer products that can share data over the Internet– from automobiles to thermostats, dishwashers to slow cookers, pacemakers to insulin pumps. The benefits to consumers include convenience and efficiency; the perk for companies is a treasure trove of data on product usage and the habits of consumers themselves.

The threat: Data has broken free from the confines of computers and mobile devices, making it hard for pretty much anyone to control what is collected, how it is aggregated, and how it can be used. With regulators already struggling to keep order on the gathering of online consumer data, the Internet of Things threatens to make the Wild West of Internet data gathering that much wilder.

While regulators in the United States and abroad already have their hands full policing privacy and security issues inherent to Websites, mobile apps, and retail “Big Data” collections, this new breed of connected devices is a far more difficult area to police. For example, the Federal Trade Commission has structured many of its online efforts around mandatory safeguards for “personally identifiable information” data points such as Social Security numbers that can directly single out an individual. The Internet of Things redefines the very concept of what is personally identifiable.

“The truth is that personally identifiable information is a mathematical construct and no longer a list of specific items,” says Theodore Claypoole, a privacy expert and partner with the law firm Womble Carlyle Sandridge & Rice. Identities can be gleaned purely from location, piecing together work and home. Studies have also shown that just three pieces of data—birth date, zip code, and gender—can be enough to zero in on nearly any individual, he says.

“We built our regulatory scheme around pieces of personally identifiable information, when the truth is it can mean nearly anything we want it to depending on what kind of information I am collecting,” Claypoole says. “When we are talking about the Internet of Things, more things will know where you are and if we know where you are, we know who you are.”

Another obstacle for regulators is that high-tech security “doesn’t lend itself to a list of rules where, if you do a to z then you have good procedures and you are going to be safe,” says Christopher Clearfield, a principal at System Logic, a risk consultant. “Cyber-security doesn’t really lend itself to a rule-based approach and it will be really hard for agencies to actually regulate this.”

Regulatory Fits and Starts

It may not be easy, but regulators—specifically the FTC—are nevertheless trying. In September 2013, the agency took its first steps toward cracking down on the Internet of Things when it reached a settlement with TRENDnet, a California-based company that markets security cameras that can be monitored remotely by users over the Internet. The FTC’s complaint said that TRENDnet failed to implement reasonable security measures, and as a result the live feeds for nearly 700 cameras were publicly accessible online, with illicit viewers watching and recording unaware families from inside their homes.

> “What companies should think about as a starting point, and it is not an end point, is looking at the core FTC principles and privacy-by-design notions as a jumping off point and asking if they apply and make sense, in some form, in this new space.”
>
> Gerard Waldron, Partner, Covington & Burling

“The exposure of sensitive information through respondents’ IP cameras increases the likelihood that consumers or their property will be targeted for theft or other criminal activity, increases the likelihood that consumers’ personal activities and conversations or those of their family members, including young children, will be observed and recorded by strangers over the Internet,” the complaint stated.

The center of the FTC complaint is that TRENDnet failed to use reasonable security measures, despite implying to customers that it was doing so. It also failed “to employ reasonable security in the design and testing of its software.”

A lesson for companies: If you have a stated privacy policy, you need to abide by it fully. “Over time, the more responsible players in the industry have developed best practices and the FTC has said that if you have a policy they are going to hold you responsible for it,” says Gerard Waldron, a partner with the law firm Covington & Burling.

One obstacle, however, is exactly how a company can convey a privacy policy and ongoing updates given the wide variety of appliances. For instance, do you have to tell a driver that data is collected every time he starts his car? Do privacy expectations differ based on the product being used?

“What companies should think about as a starting point, and it is not an end point, is looking at the core FTC principles and privacy by design notions as a jumping off point and asking if they apply and make sense, in some form, in this new space,” Waldron says. “They may not make sense in the same way they do in the online world, but I don't think that means you rip them up, throw them away, and do whatever you want. It means you need to be smart and think about how the principles and general policy goals make sense for your particular product or service. It may be that they are not all adaptable, but some are."

Do What You Say You Do

Companies that adopt boilerplate online security language but don't ensure that the proper data security and privacy safeguards are in place for Internet appliances could be exposed. "The FTC has made it [its] mission to go after people who say they are secure, when they really aren't making much effort to be secure," Claypoole warns . "The problem is that the government cannot set standards for data security because the technology changes all the time. Regulators can only hold you to your word that you are going to live up to what you are promising you are doing."

Claypoole suggests that companies pay closer attention to the terms-of-use agreements they present to consumers and avoid the temptation to turn them into a marketing document full of vague, feel-good promises.

"It's a matter of keeping your promises," he says. "One of the most important things when you are dealing with privacy and security and writing something for the general public is to be accurate. Describe what you are doing precisely and don't overstate anything. This isn't a sales or marketing document. You need to tell people exactly what you are doing. It is bad for businesses to have broad statements like, 'We care deeply about your privacy and do everything possible to protect that information.' No you don't, because you can't afford to. Nobody does everything possible."

Empty promises, despite their public relations value, can open a door for regulators to take action and for judges to side with aggrieved plaintiffs, says Claypoole.

According to John Hutchins, a partner with the law firm Troutman Sanders, companies must know the data they are collecting from online appliances and how it is used. "The first question you have to ask is what information you are collecting, then ask how you are using it. That includes how it is stored and what are the security protocols you have in place. And those same questions apply to information that is going to be collected in these less traditional ways."

An ideal for regulators, one that is also the basis of privacy standards under consideration in the European Union, is that consumers should have meaningful opportunities to review and accept a privacy policy and "own" their data. "That is a laudable goal, but it is not realistic," Hutchins says. "Two, three, or five years from now there are going to be so many devices connected to the Internet that are not the kind of devices we are used to seeing connected." ■

WHAT TRIGGERS AN AC INVESTIGATION?

As it looks to regulate the "Internet of Things" the Federal Trade Commission has been considering a variety of questions, many of which were posed to businesses as part of a public comment process earlier this year. Among those questions:

1. How can consumers benefit from the Internet of Things?
2. What are the unique privacy and security concerns and solutions associated with the Internet of Things?
3. What existing security technologies and practices could businesses and consumers use to enhance privacy and security in the Internet of Things?
4. What is the role of the Fair Information Practice Principles in the Internet of Things?
5. What steps can companies take (before putting a product or service on the market) to prevent connected devices from becoming targets of, or vectors for, malware or adware?
6. How can companies provide effective notice and choice? If there are circumstances where effective notice and choice aren't possible, what solutions are available to protect consumers?
7. What new challenges does constant, passive data collection pose?
8. What effect does the Internet of Things have on data de-identification or anonymization?
9. How can privacy and security risks be weighed against potential societal benefits (such as improved healthcare decision making or energy efficiency) for consumers and businesses?
10. How can companies update device software for security purposes or patch security vulnerabilities in connected devices, particularly if they do not have an ongoing relationship with the consumer? Do companies have adequate incentives to provide updates or patches over products' lifecycles?
11. How should the FTC encourage innovation in this area while protecting consumers' privacy and the security of their data?
12. Are new use-restrictions necessary to protect consumers' privacy?
13. How could shifting social norms be taken into account?
14. How can consumers learn more about the security and privacy of specific products or services?
15. How can consumers or researchers with insight into vulnerabilities best reach companies?

Source: Federal Trade Commission.

Facebook's Big Data Fail Calls for More Ethics

Continued from Page 5

Increasingly, companies will likely find that they need an internal arbiter of what is not just legal when it comes to data, but what is ethical. "Organizations are realizing data ethics are not going away, says Kord Davis, a digital strategist, business consultant, and author of the book *Ethics of Big Data*.

"The question, however, is who should be in charge of parsing those ethical quandaries? One of the first places companies may turn to is the compliance function," he adds. "Organizations already have compliance capabilities, legal capabilities, and program managers capable of taking new enterprise initiatives and developing programs around them."

While compliance may be "a fine place to start" initially, Davis says companies need to dig deeper. "We are on the cusp of organizations realizing what skill sets and business processes they need to develop," he says. "This can be formed by compliance, but ethical data handling is not just a compliance issue and organizations are starting to realize that."

Leading the Ethical Discussion

Companies will also need to embrace values-based management. "There should always be somebody outside the system who is observing, validating, and analyzing how the system is working and whether it was doing what it was intended to do," Davis suggests. Organizations are going to realize this idea of having a '10th man' that puts them in a position to do that internal review, analysis, and reporting."

It may be easier said than done, however. "One of the big challenges I've seen, is that organizations just don't know how to have ethical discussions in the context of business," he says. "Why is at ethics so hard? It is a loaded word. It makes people uncomfortable and it implies that you and your values are gong to be judged." Those fears, however, can quickly dissipate when a company commits to having ethical discussions. "At a minimum, if you just create a space for the explicit conversations, you are gong to be in a better position," he says.

"All the headlines out there talk about how data is the new currency," says Dave Deasy, a vice president at TRUSTe, a data privacy management company. "Yes, but the old currency is trust. Companies built their brand on this idea of building a trusted relationship with their customers. Yet, you make a couple of missteps with regard to how you are collecting data, and you wipe out years of brand trust you built up over time."

While companies focus on data from a legal compliance perspective, at least initially, the "ultimate driver" needs to be making sure they can continue to have a trusted relationship with their customers. "Companies are in an unprecedented place in terms of their ability to do creative things from a marketing perspective, but at the same time that creates lots of challenges, and it is only going to get harder," Deasy says.

Deasy's advice for companies is to start with privacy and transparency as a cornerstone for all business decisions. "It is all about letting people know what data you are collecting, what you are doing with that data, and giving them the ability to control it," he says.

"Step one is making sure the company understands where all the data is being collected and conducting a data audit," he suggests. The second step is putting internal procedures and guidelines in place around who gets access to data and what they can do with it. Next, there must be a "big focus on training" so those procedures and policies are communicated throughout the company.

Companies also need to carefully vet and audit any third parties that gain access to customer data. "Sometimes they may know exactly who those third parties are and how they got there," Deasy says. "It can be a complex thing for a company to understand all that tracking activity and be able to manage it."

"One of the fundamental focus areas of compliance is having proper vendor management procedures in place," he adds. "These third parties are not as easy to figure out and, in a lot of cases, there are fourth parties that can have access to your Website through other third parties. If you don't have the right tools to see that, and manage it, there can be unintended consequences." ■

FACEBOOK CODE OF CONDUCT

Below is an excerpt from Facebook's Code of Conduct, as amended on June 15, 2015.

If you learn about or suspect a violation of this code, another Facebook policy, or any law, you shall promptly report it to your manager, another manager, Human Resources, Internal Audit, or the Legal Department. If you are uncomfortable making such a report, you may do so anonymously. For more information on such anonymous submissions, please see Facebook's Whistleblower and Complaint Policy on the wiki here.

In cases in which an individual reports a suspected violation of policy or law in good faith and is not engaged in the questionable conduct, Facebook will attempt to keep its discussions and actions confidential to the greatest extent possible and in compliance with applicable laws and regulations governing privacy. Facebook will not retaliate against anyone making a good-faith report of a potential violation. Facebook will investigate any report of a violation. You must cooperate fully with any investigation, but should not investigate independently, as alleged violations may involve complex legal issues, and you may risk compromising the integrity of a formal investigation.

Conduct that violates the law or company policies is grounds for prompt disciplinary or remedial action. In addition, your failure to report a known violation of law or company policy by someone else may result in disciplinary action for employees and/or termination of employment/your relationship with Facebook. Discipline for a violation of Facebook policies or applicable law may range from a warning up to and including summary termination of employment/your relationship with Facebook (in accordance with applicable law). Where laws have been violated, we will cooperate fully with the appropriate authorities.

Source: Facebook.

CCOs Play a Stronger Role in Data Privacy

Continued from Page 4

where initially compliance officers were not involved but are now vital voices at the table. (Think of all the trouble third parties can bring to your business.)

"I see that evolution in the information security and data protection space as well," Blake says. "It's simply a matter of time for banks to further include the CCO into that realm of information governance."

Ground Zero for privacy regulations complicating business operations is, of course, France. French data protection laws date back to the 1970s, and the tough stance of the Commission Nationale de l'Informatique et des Libertés, its data protection authority, has flummoxed many U.S. businesses. Last year CNIL fined Google €150,000 ($164,000) for changes the company made to its privacy policies.

The enforcement was triggered by an announcement that Google planned to replace product-specific privacy policies with single, overarching terms without notifying users ahead of time. An investigation by the Article 29 Working Party, an advisory body comprised of DPAs from 28 European member states, ruled that Google's privacy policy violated the European Data Privacy Directive because users were not informed of what data would be collected, or why, and data retention timelines were not public.

Regulatory skirmishes like that will force companies to consider data privacy compliance more seriously as they plot business moves. Europe is simply the biggest example, not the only one.

"The data security laws in the EU are complex, with non-EU countries beginning to follow suit," says Meena Elliot, chief legal officer at Aviat Networks, a $350 million maker of wireless transmission systems. "Google is facing challenges concerning the EU's views on the right to be forgotten from the Web. At the moment, there is no such requirement in the United States."

But Google has been facing intense heat, especially from France. Recently the company received a formal notice from CNIL calling for Google to delist links from all European versions of Google Search and all global versions as well.

In response, Google argues that while European law enforces the right to be forgotten, its scope is limited and can't be applied globally. In fact, content (read: data) that is illegal in one country may be legal in another—one more challenge that companies face as they grapple with data privacy compliance.

"Google's stance in this case means a lot for compliance officers, and it serves as a warning for companies as they expand into new regions," de Guzman says. "It shows that the chief compliance officer constantly needs to be aware of new legal developments and have strong policies in place as governments around the world roll out new or more stringent data privacy laws."

"The compliance function has dramatically evolved over the years," Blake says. "Now we are engaging more in IT solutions that help to protect customer information. Although the functions between compliance, IT, and information security are still somewhat separated, we work very closely with these areas to have a sense of the overall controls in place to protect consumer and employee data." ■

Auditing in the Era of Big Data

Continued from Page 8

firms' return to the consulting business with an eye on whether it compromises the auditor's ability to perform an independent audit.

Regulatory Skepticism

PCAOB member Lewis Ferguson recently said regulators are concerned about the economic model for audit firms—fees for audit are flat while the real growth lies in consulting services—and whether that could jeopardize audit quality. "Part of what's driving the acquisition binge is to acquire the businesses that have those analytical skills," he says. "To that extent, I understand why the firms are driven to make these acquisitions, but that's not the only kind of acquisitions they're making." The firms make a valid argument that they need to invest in technology and analytics for the sake of the audit, he concedes. "This could fundamentally change the way we do audits," he says. "If anything it is likely to make the audit better."

Another concern, says Ferguson, is whether auditing standards need some revision to facilitate the use of more advanced technology that would make traditional sampling techniques unnecessary or even obsolete. "We have to assure our standards are not forcing auditors to do things that are simply no longer relevant," he says.

That's a concern for auditors as well, says Baskin. "The standards are based on what we could do 50 years ago," he says, which is sampling, not examining all transactions. So even if auditors could look at all transactions, the standards would still require sampling, he says, leading to inefficiency. The PCAOB has also insisted through its inspection process that auditors test the completeness and accuracy of any database auditors rely on for audit evidence, an impossibility if auditors were to use externally available market data in their analysis, he says.

Some auditors may also resist over fear of a new litigation risk, says Peter Bible, a partner with audit firm EisnerAmper. "In hindsight, if something goes wrong for whatever reason, you can always be challenged or criticized or found at fault for not doing something," he says. Auditors will have to wrestle, for example, with what to do about small or immaterial mistakes that are bound to turn up with more detailed analytics, says Baskin. "The software might overwhelm you with anomalies that have to be investigated," he says.

Still, folks like Ucuzoglu are excited about the potential. "Audit firms haven't substantively changed the way the profession goes about doing an audit in a long time," he says. "This is frankly overdue." ■