DATA GOVERNANCE

Providing assurance regarding data risk management

The Institute of Internal Auditors | Financial Services AUDIT CENTER

Table of Contents

INTRODUCTION

Big data introduces both opportunity and risk

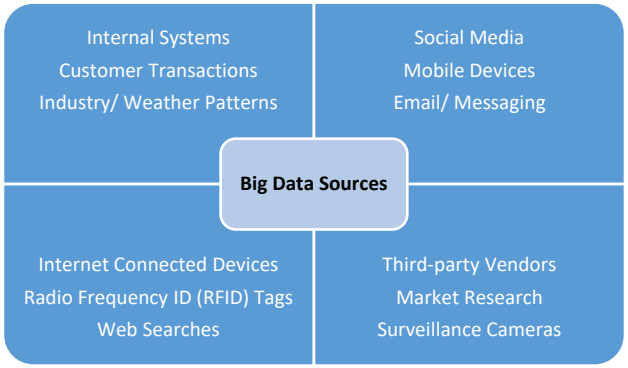
While the concept of risk related to data ethics is relatively new, Chief Audit Executives (CAEs) predict that its relevance will grow rapidly over the next five years. Complexity in the collection, analysis, and use of data is expanding rapidly, complicated by artificial intelligence. Financial services is often on the cutting edge of technological development, so internal auditors in this industry should have some knowledge of these topics:

- Data governance best practices.
- Risks associated with failing to establish proper data governance.
- Potential reputational and financial damages resulting from failed data governance.

What is data?

Big data is a popular term used to describe the exponential growth and availability of data created by people, applications, and smart machines. The term is also used to describe large, complex data sets that are beyond the capabilities of traditional data processing applications. The proliferation of structured and unstructured data, combined with technical advances in storage, processing power, and analytic tools, has enabled big data to become a competitive advantage for leading organizations that use it to gain insights into business opportunities and drive business strategies. However, the challenges and risks associated with big data must also be considered.

Figure 1: Sources of Big Data



Increased demand for data by businesses, immature data risk management frameworks, and emerging risks and opportunities that are not widely understood or systematically managed by organizations have created a need for more direction in this area. Internal auditors, in particular, must develop new skill sets and obtain knowledge of big data principles to effectively provide assurance that risks are addressed and benefits are realized.

STRUCTURED AND UNSTRUCTURED

Why unstructured data is harder to control

Structured data

Historically, the majority of data stored within organizations has been structured and maintained within relational — or even legacy hierarchical or flat-file — databases. Structured data is organized and allows for repeatable queries, as much of the data is maintained in relational tables. It is often easier to control than unstructured data, due to defined ownership and vendor-supported database solutions.

Unstructured data

Unstructured data consists of datasets (typical large collections of files) that aren't stored in a structured database format. Unstructured data has an internal structure, but it's not predefined through data models. It might be human generated, or machine generated in a textual or a non-textual format. This type of data is not confined to traditional data structures or constraints. It is typically more difficult to manage, due to its evolving and unpredictable nature, and it is usually sourced from large, disparate, and often external data sources. Consequently, new solutions have been developed to manage and analyze this type of data. Examples of unstructured data are:

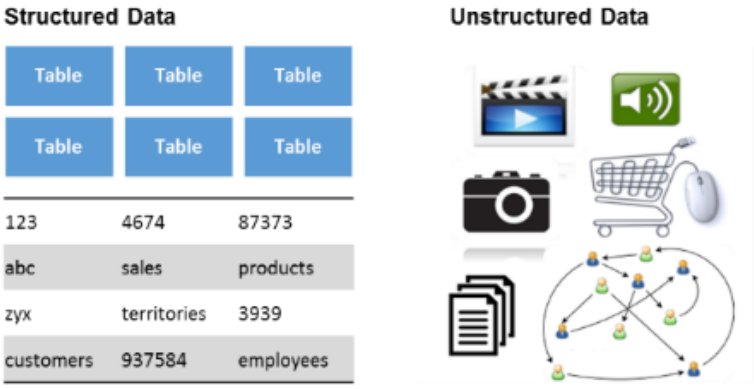
- **Rich media.** Media and entertainment data, surveillance data, geo-spatial data, audio, weather data.
- **Document collections.** Invoices, records, emails, productivity applications.
- **Internet of Things (IoT).** Sensor data, ticker data.
- **Analytics.** Machine learning, artificial intelligence (AI).

Example: Marketing pulls a new customer report, and another business unit pulls a new customer report, and the numbers are different. This can affect how decisions are made. Big data is not always clearly defined and it's possible to get multiple outcomes from the same scenario.

Businesses collect and use big data for a variety of reasons including:

- Competitive advantage.
- Increased revenue.
- Innovation and faster product development.
- Market demands predictions.
- Informed business decisions.
- Operational efficiency.

Figure 2: Examples of Structured and Unstructured Data



THE REGULATORY ENVIRONMENT

Governments have stepped in to protect data

Regulations deal with use, transparency, and ownership of data

As big data collection and use becomes more prevalent, countries, regions, and states have passed legislation governing how consumer data can and cannot be used. Some examples include:

- The European Union's General Data Protection Regulation (GDPR).
- The U.S. Health Insurance Portability and Accountability Act (HIPAA).
- Children's Online Privacy Protection Act (COPPA).

In addition, individual states, regions, and provinces are developing more expansive regulations around data ethics, such as the California Consumer Privacy Act (CCPA). Data breach notification requirements can vary from state to state in the United States. For example, if 500 people from California are affected in a breach, the organization is required to notify the California State Attorney General's office. Questions to ask are:

- Which states are our customers in?
- What are the corresponding notification requirements?
- Have any of the states in which we operate enacted their own data privacy regulations?

The rise of unstructured data and the need for a framework governing the ethical handling of that data has driven this new level of regulation. In general, the regulations try to communicate the following principles:

- **Ownership**. Individuals own their own data. (Laws such as the General Data Protection Regulation, indicate that individuals own their own personal data.)
- **Transaction transparency**. If an individual's personal data is used, they should have transparent access to the algorithm design used to generate aggregate data sets.
- **Consent**. If an individual or legal entity would like to use personal data, one needs informed and explicitly expressed consent of what personal data moves to whom, when, and for what purpose from the owner of the data.
- **Privacy**. If data transactions occur, all reasonable effort needs to be made to preserve privacy.
- **Currency**. Individuals should be aware of financial transactions resulting from the use of their personal data and the scale of these transactions.
- **Openness**. Aggregate data sets should be freely available.

DATA GOVERNANCE BEST PRACTICES

Avoiding customer notification pitfalls

The role of codes of conduct, privacy policies, and CAN-SPAM

One way an organization can ensure their employees are informed of their data governance expectations is through their code of conduct. In data intensive businesses such as financial services, codes of conduct should be updated with data governance-related requirements. These requirements could include:

- List or inventory of data covered by the organization's data governance policies.
- Employee security requirements to ensure employees are aware of best practices when securing company equipment and mobile devices.
- Restricted access tied to job roles.
- List of actions that could negatively affect data governance with penalties.

Obviously, code of conduct elements will differ based on industry, geography, etc. However, employees should understand that they may be held personally accountable for misusing data.

Another important touch point affecting an organization's data governance practices is its privacy policies. Codes of conduct communicate data governance expectations and requirements to employees, but privacy policies communicate those elements of data governance to customers. Segmented privacy policies help customers understand what data collection, usage, and storage practices apply to them. Customer notification practices ensure your customers know what your privacy policies mean.

Figure 3: Privacy Policies and Customer Notification Best Practices

Best Practices for Privacy Policies	Best Practices for Customer Notification
• **Consider segmenting policies (e.g., GDPR policy, CCPA policy, in addition to general privacy policy).** • **Read the policies to ensure language is clear.** • **Opt-in and opt-out provisions should be clear.**	• **Disclosing data collection at specific points in the process.** • **Link the privacy policy at the point of collection.** • **Ensure emails comply with the CAN-SPAM Act.** • **Unsubscribe and opt-out options should be clear at point of collection and in marketing materials.**

While most of the best practices for privacy policies are common sense. The opt-in/opt-out provisions can become confusing. From a legal perspective, much of what financial services organizations do to market their products is not considered vital to delivery of services. So, when data is collected for marketing purposes or through marketing programs, opt-in and opt-out clarity can become crucial. Some questions to ask may include:

- What do you do with data for people who opt out?
- Do you delete data stored in backup servers?
- Should you disclose how you might delete backup data?

Customer Notification Pitfalls

- **Making the privacy policy difficult to find, read, and retain.**
- **Confusing opt-out process.**
- **Opt-out processes stated but not implemented.**
- **No visible "unsubscribe" option in your emails, especially marketing specific emails.**
- **Companies may not have the ability to match a cookie with a person.**
- **Selling cookies.**

In the United States the CAN-SPAM Act is a key piece of legislation that internal auditors should understand. Each separate email in violation of the CAN-SPAM Act is subject to penalties of up to $42,530, so non-compliance can be costly. But following the law isn't complicated. Here's a listing of CAN-SPAM's main requirements:

- **Don't use false or misleading header information.** Your "From," "To," "Reply-To," and routing information — including the originating domain name and email address — must be accurate and identify the person or business who initiated the message.
- **Don't use deceptive subject lines.** The subject line must accurately reflect the content of the message.
- **Identify the message as an ad.** The law gives you a lot of leeway in how to do this, but you must disclose clearly and conspicuously that your message is an advertisement.
- **Tell recipients where you're located.** Your message must include your valid physical postal address. This can be your current street address, a post office box you've registered with the U.S. Postal Service, or a private mailbox you've registered with a commercial mail receiving agency established under Postal Service regulations.
- **Tell recipients how to opt out of receiving future email from you.** Your message must include a clear and conspicuous explanation of how the recipient can opt out of getting email from you in the future. Craft the notice in a way that's easy for an ordinary person to recognize, read, and understand. Creative use of type size, color, and location can improve clarity. Give a return email address or another easy internet-based way to allow people to communicate their choice to you. You may create a menu to allow a recipient to opt out of certain types of messages, but you must include the option to stop all commercial messages from you. Make sure your spam filter doesn't block these opt-out requests.
- **Honor opt-out requests promptly.** Any opt-out mechanism you offer must be able to process opt-out requests for at least 30 days after you send your message. You must honor a recipient's opt-out request within 10 business days. You can't charge a fee, require the recipient to give you any personally identifying information beyond an email address, or make the recipient take any step other than sending a reply email or visiting a single page on an Internet website as a condition for honoring an opt-out request. Once people have told you they don't want to receive more messages from you, you can't sell or transfer their email addresses, even in the form of a mailing list. The only exception

is that you may transfer the addresses to a company you've hired to help you comply with the CAN-SPAM Act.

- **Monitor what others are doing on your behalf.** The law makes clear that even if you hire another company to handle your email marketing, you can't contract away your legal responsibility to comply with the law. Both the company whose product is promoted in the message and the company that actually sends the message may be held legally responsible.

So, if your organization is following the legal rules and best practices, the next key consideration is record retention. Record retention laws vary from state to state and can even be governed by local laws and regulations. For financial services organizations operating in multiple jurisdictions, record retention can be a very difficult policy to develop and enforce.

There are some general rules of good practice that organizations can follow for record retention that will cover most situations.

- Organizations have the right to retain records in accordance with the law.
- If you are unsure whether or not you can retain data, utilize outside counsel and get a confirmation from them in writing of their advice.
- Always remember personal liability is a possibility.
- Organizations can store data for analytics in the future, but you should have processes in place to de-personalize the data, so it is not linked to a person.

Record Retention Confusion

Q: If there are differences in requirements between lawmaking bodies — state law versus federal law, local law versus state law, which law takes precedence?

A: Unfortunately, there is no easy answer to that question. Larger organizations may need to do a case-by-case analysis by data type, collection point or other criteria.

Overall, there are two key ideas: 1) Do not collect data that you do not need and 2) Do not store data longer than required.

DATA GOVERNANCE

Building a robust governance program

Roles and responsibilities in data governance

The adoption of big data in an organization requires strengthening data governance to ensure that information remains accurate, consistent, and accessible. Key principles of data governance include:

- Proper data inventory and mapping.
- Identifying roles and responsibilities.
    - Issue management.
    - Escalation protocols.
    - Consistent reporting.

Data Governance Definition

The exercise of authority, control, and shared decision-making (planning, monitoring, and enforcement) over the management of data assets.

Organizations must implement appropriate controls to ensure that all necessary data quality dimensions (e.g., completeness, validity, uniqueness) are properly maintained.

Data inventory

The first step toward a robust data governance program is creating an inventory that is simply a log of the most important data assets. Internal audit needs to support its functional partners in evaluating the process for creating and maintaining the inventory. Some of the activities involved in creating a data inventory are:

- **Data mapping.** What data are you collecting at specific points, storage, and collection?
- **Categorizing data.** There is guidance on how to categorize data through rating scales showing criticality, value, and other factors.

Data office

Overall, the organization should establish a data office with a chief data officer that sets policy and standards with requirements for the business to implement data governance protocols. The office should lead a data governance committee where key decisions, escalations, and issues related to data governance are reviewed and approved by the committee. Some considerations:

- An executive sponsor could be employed to beef up the credibility and authority of the program.
- The sponsor could also make sure people and resources are allocated to get the basics down before data use ramps up.

There should be data owners within the businesses and functions who understand these elements and participate in working groups and councils to escalate issues and concerns. The owners could also work with other data owners on projects to implement updated data standards. These groups should report up through the data governance committee who is providing oversight.

Smaller organizations should know that a separate data office with its own personnel is not necessarily required. These responsibilities can be accomplished a variety of ways. The key is to make sure there is someone in the company who understands these issues and can escalate them to senior management without any conflict.

Figure 4. Data Governance Roles and Responsibilities

Data Officers	Data Office	Data Governance Committee
• **Individuals assigned to manage the data governance in a specific area.** • **Experience dealing with data including "big data".** • **Knowledge of the business and how it uses data.** • **Knowledge of data risks and controls.** • **Implement policies and procedures for data governance including taking action when data quality or data use issues arise.** • **Train employees on the organization's Data Governance Framework.** • **Escalate data issues.**	• **Chief Data Officer** • **Establishes policies and standards for data governance.** • **Provides oversight of business implementation of data governance protocols.** • **Serves as one level of escalation.** • **Leads the Data Governance Committee.**	• **Cross-functional committee.** • **Owns the organizationwide Data Governance Framework.** • **Serves as second level of escalation and reviews exceptions/ issues.** • **Reports on data governance to senior management and the board.**

AUDITING DATA GOVERNANCE

Uncovering data privacy and ethics controls

Key questions to ask during a data governance engagement

The role of internal audit is to identify data ethics-related governance bodies, policies, procedures, etc. within the organization and assess their effectiveness. Internal auditors should be aware that data ethics may be a part of a larger audit engagement. Data governance may be integrated into a variety of other processes, policies and procedures, so internal auditors may have to perform an identification exercise to locate specific data governance risks and controls and then dig a level deeper to identify the data ethics-related subsets of the broader procedures.

Good data governance audit programs should not only cover the organization and effectiveness of the operation of the data governance bodies as discussed above, but should also evaluate the minimum control expectations for data factors such as:

- Data storage.
- Data quality.
- Data transfer.
- Data collection and use reporting.

This goes back to mapping and inventorying data. Internal auditors will need to know where the data is going and what rules and regulations are required for collection, use, and storage. Internal auditors should do ongoing monitoring to make sure those processes are operating as desired. This is not something internal audit wants to look at every two or three years. Data governance is something the organization will be doing every day, so it is critical to continuously keep an eye on it.

Some key questions internal auditors should ask stakeholders during the **Gather Information** stage of the engagement could include:

- Does our customer know we are collecting their data? Do we honor, complete, and document opt-outs?
- How do we make our customers aware of the data we are collecting on them and how we use it?
- How have we constructed governance around data collection and management?
- Do we have an inventory of the data we capture?
- Is our code of conduct reviewed and acknowledged in writing by employees on a regular timeframe? Does it include data-related elements?

- Do we require employees to complete ethics training?
- When we seek to introduce a new product or service, is a risk assessment done regarding the data that could be collected, and what we would use it for once the product is in place? Is that information included in risk reporting to executive management and the board? Do you have processes and procedures to update your data mapping?

In the **Risk Assessment** stage of the engagement, internal auditors working with data should engage with the organization's chief information officer (CIO) and other key leaders to better understand the risks in terms of data collection, storage, analysis, security, and privacy. As of right now, many organizations do not have a clear picture of where their data lives. They have unidentified, inconsistent and disparate data spread across various functions, units, and devices. This presents a pervasive security liability, leaving companies vulnerable to potentially costly security breaches.

Some key general data risks to consider are:

- Inappropriate data collection protocols.
- Poor data quality.
- Inadequate technology solutions and/or configurations.
- Inadequate storage.
- Insufficient security.
- Immature data governance practices.

Some data use key performance indicators (KPIs) and key risk indicators (KRIs) include a variety of data governance issues. Internal auditors should be aware of and include in their data governance-related work programs KPIs or KRIs such as:

- Board has established guideposts on how the organization can ethically use data.
- Management has established and/or communicated policies on how data can be ethically used.
- Unfavorable social or traditional media news coverage of how the organization or other similar organizations use data.
- New or proposed regulations or legislation related to data collected and used by the organization.
- Fines and penalties incurred for data ethics-related violations.
- Number of employees completing ethics training on time and with satisfactory performance on the associated exam.
- Number of employees with out-of-date or incomplete conflict-of-interest or code-of-conduct acknowledgements.

As mentioned, the scope and objectives of an audit to address data governance may include elements embedded into larger governance structures and broader data management practices. In addition to the code-of-conduct considerations, privacy policies, customer-notification processes, and rules and regulatory requirements considered above, internal auditors may want to include reviews of the following in their work programs (or review the results of audit engagements that have already covered these areas):

- Documentation demonstrating ethics complaints, whistleblowing situations, or other incidents involving members of management — individuals or as a group — are investigated and addressed promptly and in a manner consistent with the organization's ethics policies, escalation protocols, code of conduct, etc.
- Perception or evidence of management or other employees retaliating against those who report issues.
- Statistical trending of complaints, whistleblowing situations, or other incidents to determine the effectiveness of controls in place.
- Rates of completion and pass rates for electronic training programs including ethics, code of conduct, core values, etc.

Conclusion

Data governance can be a big risk area for organizations and it may be largely unexplored by internal auditors. However, the current increase in regulations, fines, penalties, and reputational risk exposure make it a critical area for internal auditors to cover as quickly as possible.

About The Financial Services Audit Center

The Financial Services Audit Center (FSAC) is a specialty offering of The IIA for financial services auditors. FSAC was established to provide financial services auditors with targeted high-quality professional development; networking opportunities for knowledge sharing among financial services stakeholders; and ongoing, timely, and relevant reporting on trends, benchmarking, and thought leadership in the internal audit profession. This report is reserved for your exclusive use as a member of the Financial Services Audit Center. For more information, visit www.theiia.org/FSAC.

About The IIA

The Institute of Internal Auditors (IIA) is the internal audit profession's most widely recognized advocate, educator, and provider of standards, guidance, and certifications. Established in 1941, The IIA today serves more than 200,000 members from more than 170 countries and territories. The association's global headquarters is in Lake Mary, Fla. For more information, visit www.theiia.org.

Disclaimer

Copyright

August 2020

Financial Services
AUDIT CENTER

Global Headquarters
The Institute of Internal Auditors
1035 Greenwood Blvd., Suite 149
Lake Mary, FL 32746, USA
Phone: +1-407-937-1111
Fax: +1-407-937-1101
www.theiia.org/FSAC