

Preventative and Detective Data Controls

**Why Understanding Information Context
is Critical for Data Governance**



from The Data Governance Institute

The Data Governance Institute
www.DataGovernance.com
2511 E. Colonial Drive #219
Orlando, FL 32803 USA
telephone: 321.438.0774



Preventative and Detective Data Controls

Why Understanding Information Context is Critical for Data Governance

Abstract

Data Governance programs are concerned with “appropriate data usage” – ensuring that the right data is used by a person in the right role only in the right context. It is this last factor – the context in which a person accesses data – that some organizations have trouble articulating. And so, this paper will describe a Data Usage Triangle that all stakeholders – IT, Business, Security, Auditors, and others – can use to describe a specific type of data risk. We’ll explore some scenarios that demonstrate why context is key to understanding risk and interpreting questionable data usage activities. We’ll describe how different types of controls interact with each other, and we’ll discuss some options for preventative and detective data protection controls.

Governance and Data Usage

Two of the main functions of Data Governance are to set policy for the appropriate use of information and to contribute to strategies for controlling and monitoring that usage.

This is especially evident for certain “flavors” of Data Governance. Some programs exist primarily to assist Compliance, Privacy, and Security teams as they work together to manage access to personal, private data. These types of programs focus on ensuring that the organization is in compliance with laws, regulations, and contracts that stipulate access controls for private data.

But all organizations have sensitive data: customer lists, sales data, product information, employee/customer personal and private information. And so, all “flavors” of Data Governance will need to give attention to appropriate data usage.

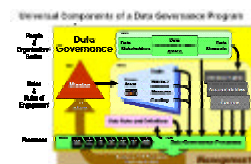
What does this mean? “Appropriate data usage” means that the **right data** is used by a person in the **right role** only in the **right context**.

It is this last factor – the context in which a person accesses data – that some organizations have trouble articulating. And yet, a common understanding is necessary if Data Governance teams are to successfully work with groups from IT, Business, Security, Auditing, etc. to create policy that can be translated into actionable controls and technologies used to manage risk.

And so, this paper will describe a Data Usage Triangle that all stakeholders can use to describe a specific type of data risk. We’ll

Data Governance and Data Protection

Assisting with the protection of sensitive data – and other compliance / security concerns – is only one potential focus area for Data Governance programs. Visit the Data Governance Institute's website at www.DataGovernance.com to learn about other focus areas and to download a copy of the free DGI Data Governance Framework.



explore some scenarios that demonstrate why information context is key to understanding risk and interpreting questionable data usage activities. We'll describe how different types of controls interact with each other, and we'll discuss some options for preventative and detective data protection controls.

Appropriate Data Usage

“Appropriate data usage” means that the **right data** is used by a person in the **right role** only in the **right context**. Even the simplest of data policies generally address two of these three factors: data and roles. Such policies may call out types of data (Financial data, personal/private data, customer records, etc.) and will specify which people filling which corporate roles should have the ability to add, modify, delete, or read these records.

This level of policy is necessary and useful. However, these broad, high-level policies don't address situations where a “trusted user” – someone who has been given credentials to come inside the firewall and see, copy, or change data – is abusing that trust. They don't address the types of risks that keep CEOs awake at night:

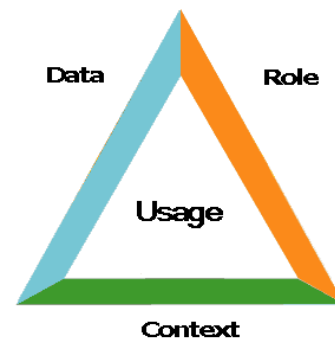
- ◆ Could the salesperson who just gave notice be planning to take copies of our customer data? Could we tell if this were happening?
- ◆ Could one of our contractors be inserting secret code that accesses our financial data? Could we tell if this were happening?
- ◆ Can we 100% trust all the DBAs who have the highest level of database permissions? Do we have audit trails or other records that can't be accessed by technology staff and can be easily interpreted by auditors and business staff?

To answer a CEO's questions – and to manage these types of risks – you need detail-level preventative and detective data controls that work together to accomplish specific data access goals. But first, to create these goals, you need to expand your view of data access risk from a high-level data/role perspective to a perspective that takes into account data, roles, and information context.

Information Context

Let's look at some scenarios in which a trusted user accesses information. Even though the user has been issued appropriate

The Data Usage Triangle



log-in credentials, would any of these scenarios raise concern in your organization? Would you want to know if any of the following events occurred at your company? How soon? Would you want policies or controls to detect them, if you couldn't prevent them?

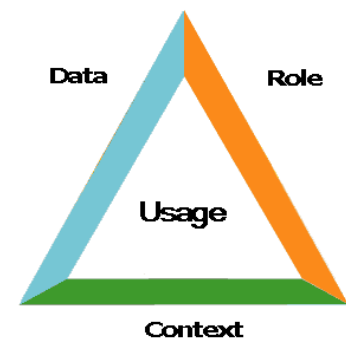
- ◆ An analyst who routinely pulls down records on 20 customers now makes a copy of 20,000 customer records – at home, at 2 A.M.
- ◆ A contractor developing code for supply-chain software runs a “back-door” ad-hoc query to obtain the names and social security numbers of all your employees.
- ◆ A developer bypasses all software application interfaces and instead executes a new SQL routine that downloads a copy of your product specifications.
- ◆ An employee who isn't supposed to know about your top-secret Merger/Acquisition plans begins to monitor sensitive data every day, just before closing time.

In all of these scenarios, there's the possibility that no wrongdoing is taking place. But there's also the possibility of “Rogue Data Usage” – improper information access by an individual. What makes the activity improper is the context in which the user accessed the data.

Information Context Factors

1. **Amount of data**
(Were an unusual number of records accessed?)
2. **Type of data**
(Was the information accessed a type of data, such as financial, that the user doesn't generally need to access?)
3. **Business transaction**
(Was the access part of a database call that is contained within an approved business process?)
4. **Assignment of person/role**
(For unusual data access, is the access part of an approved project or special effort?)
5. **Type of acquisition method**
(Was access via an application interface or an atypical method such as an ad-hoc query?)
6. **Location of user**
(Was the user accessing the data inside or outside of the firewall, at the office or at home?)
7. **Time of day**
(Did the access take place during normal operating hours or during a period where observers might not be present?)

The Data Usage Triangle



1. Amount of data
2. Type of data
3. Business transaction
4. Assignment of person/role
5. Type of acquisition method
6. Location of user
7. Time of day

Your Data Governance team may be asked to work with your organization’s Internal Auditors, Compliance, Privacy, and Security teams to establish guidelines for what constitutes “rogue” data usage – or at least activity that warrants investigation. You may be called upon to create more detailed data policy, so that such usage is actionable by your company. You will probably be asked to help recommend controls and technologies to prevent rogue data usage when possible, to detect it when prevention can’t be assured, and to speed up auditing processes.

A common understanding of concepts

To make good decisions, your cross-functional team will need a common understanding of terms, concepts, and approaches. For example, they need to understand the connection between risk and controls.

- ◆ How do we address risk? We choose between four options.
 1. We can **transfer the risk**. This is what we do when we purchase insurance.
 2. We can **accept the risk**. We decide to do nothing, hoping the risk event will not occur. If it does, our plan is to deal with the consequences then.
 3. We can **prevent risk**. We take steps to lessen the chances of a risk event from occurring.
 4. We can **detect/correct risk**. We take steps to determine whether a risk event has occurred, and we put measures in place to reverse the outcome of that event or to lessen its impact.
- ◆ Controls are how we operationalize our risk management strategies. Controls can be preventative, detective, or corrective. They can be human processes, embedded in code or automated processes, or technology-aided human-initiated workflows.

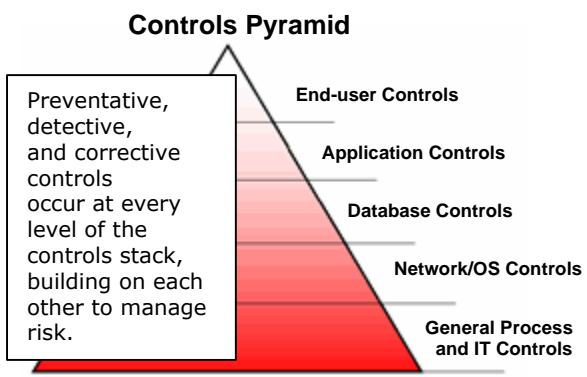
Controls build upon each other. Most risks require a “stack” of preventative, detective, and corrective controls, with those at the top of the stack depending upon foundational controls below them.

For an example of how controls build on each other, consider a case where we are

Involving Stakeholders

“Protecting sensitive data is a business problem, not an ‘IT issue.’ It’s a balancing act that involves input from many stakeholders. The result is a comprehensive stack of interdependent controls that help you prevent problems and detect data issues so you can address them before they become problems.”

- Gwen Thomas
President,
The Data Governance Institute



Source: The Data Governance Institute

managing a particular risk event: that an unauthorized user could download a list of customers along with their social security numbers or other sensitive data.

A stack of five levels of controls helps manage this risk. The top three levels focus on “trusted users” who have been given some level of credentials for working within the organization’s firewall. The fourth level down keeps out unauthorized users. And the bottom level provides a control environment.

How do they work together? At the top level of the controls pyramid – or controls stack – are user-based controls, ones that we expect a person to carry out. An example is keeping a password confidential. This protects against one scenario where the risk event could occur, but by itself it doesn’t manage the entire risk.

For example, what if it were possible to access the application anyway? Application-level controls – such as log-ins and role-based views of information – ensure that only credentialed users can open the application and that they can create/read/update/delete data only as appropriate.

But what if it were possible to bypass the application, and to simply open the underlying database? Application-level controls are always supported by database controls. These may include preventative controls, such as mechanisms for locking out classes of users except those who access the database through an application or valid business process. They may include data architectures that separate sensitive data from less-sensitive data. They may even include technologies that can detect in real-time which users are accessing data – and through what application or process – and can send alerts for unusual activity.

These top three levels of controls manage risks that can occur once a user is inside a firewall. The next level in the controls stack, or controls pyramid, is the network/operating system level. Here is where technologies are employed to keep out unauthorized parties or to detect intrusions.

The base level of the controls pyramid – the foundational level – includes general process and IT controls. Auditors looking for effective controls environments look at these types of controls. They ask whether a Security Program and Access Management processes are in place. They look at the effectiveness of risk assessments, and whether appropriate stakeholders were included in the design of controls.

Data and Risk

“Not only is data AT risk; data REPRESENTS risk. The personal, private data that you collect and store can become a great liability if it is improperly disclosed.

You can't ignore this risk; you have to manage it, and this means accepting that you can't always prevent problems.

So the question becomes one of detection. Suppose your private data were accessed improperly from someone within your firewall. When would you want to find out – when the data is used to perpetrate identity theft? Or immediately, before the perpetrator leaves the building?

- Gwen Thomas
President,
The Data Governance
Institute

Defining and locating sensitive data

Once your teams addressing policy for data usage in context have a common understanding of how controls build on each other, they need to establish the scope

of their efforts. What does “sensitive data” mean in your organization? Teams often take a scenario-based approach to define sensitivity in context.

Note: Of course, you can’t protect what you can’t find. Part of your organization’s efforts will include locating your sensitive data. Once you have defined sensitivity, located sensitive data, and decided on an approach to managing risk, then your team may be ready to let smaller groups work with actual controls.

Accountabilities and controls

It’s important to remember that controls to protect sensitive data – like any other type of controls – have lifecycles. They are designed, implemented and tested, deployed, monitored, and eventually retired. Accountability must be assigned for each phase of the lifecycle for each control.

Accountability should be role-based. Individual people, then can be assigned to roles with responsibilities for preventative, detective, and corrective controls.

Sound like a lot of work? It’s not as bad as it sounds.

- ◆ First, you’ll want to determine which are the key controls for the scenarios you’re addressing. You’ll find that certain controls – often your “first defense” or your “last defense” – are the ones that require closest attention. The best ones manage many risks at once, at the lowest possible level in the controls pyramid.
- ◆ Controls tend to occur in sets. Once you’ve established a pattern you’ll be able to assign accountability for manageable sets of controls.
- ◆ The same set of controls can generally be used to manage many instances of data. You’ll be applying them to different databases or applications, but the work should be easy to assign and monitor. Once again, controls nearest the bottom of the stack are often the most economical in the long run.

“Trusted Users”

“Often, risk comes in the form of ‘trusted users.’ Employees, contractors, and outsourcers have all been known to misuse data. If you can’t prevent access to your most sensitive data, it’s wise to have the capability to monitor data usage so you can detect problems and address them before you have a reportable security breach.”

- Gwen Thomas
President,
The Data Governance
Institute

Communications and Disclosures

Someone will have to keep track of this, of course. And someone will have to communicate control-related decisions, status, and other information to all stakeholders. This activity is critical. Over time, it’s natural for gaps and overlaps in your controls framework to occur. Effective, comprehensive communication – apart from serving its primary mission of sharing information – serves as a control itself, since it provides a mechanism for unearthing those gaps and overlaps.

Such coordination and communication work is a natural fit for most Data Governance Offices. It’s not as important who does this work, however, as it is to make sure that it gets done.

Just as important is a plan for identifying trigger events and for disclosing them in a timely manner. Consider, for example, one of our “rogue data access” scenarios – an employee who has given notice and then starts accessing data for which he has no business need. When will this be detected? How will it be communicated to someone within the organization? Can such disclosure occur soon enough to avert a wholesale problem?

Taking advantage of new technologies

Yes, protecting sensitive data is a business problem, not an “IT issue.” Yes, it requires a business approach – setting decision rights through governance, setting scope and approach, deciding on strategies, then building a comprehensive stack of interdependent controls.

But it also requires technology. And that’s always changing. New options for preventative, detective, and corrective controls are available that a few years ago were just dreams.

If you haven’t examined your technology options in the last year, it’s probably time to look again. Technologies that can be applied near the bottom of the pyramid – at the network and database level – may offer more cost-effective and comprehensive approaches to managing risk than your legacy options. And, they may offer new options for detecting problems, so you can correct issues before they turn into data disasters.



The Data Governance Institute
www.DataGovernance.com
2511 E. Colonial Drive #219
Orlando, FL 32803 USA
telephone: 321.438.0774



About the Data Governance Institute

The Data Governance Institute is the premier provider of in-depth, vendor-neutral information about – and assistance with – tools, techniques, models, and best practices for the governance of data and information.

The Institute provides a wealth of resources: the free DGI Data Governance Framework, information on data laws, regulations, and standards, whitepapers, case studies, best practices, data humor, and non-technical briefings on data-related issues and disciplines.

The Institute also provides community for governance practitioners, and it directly assists clients through training, issue analysis, and program assistance as they roadmap, design, and implement governance programs.

The Data Governance Institute also publishes www.SOX-online.com, the web’s largest source of vendor-neutral Sarbanes-Oxley information.