# Implementing a data governance program: Four steps to get the most from your organization's data assets

by Darren James

Overview

Introduction

Process

Application

Key learnings

Resources

**What is the issue?**
Many organizations don't know how to manage their data as an asset. It is not like managing tangible physical assets, such as plant and equipment, or intangible assets, such as goodwill and trademarks. Organizations need to appreciate that data represents a unique type of asset. The challenge is the volume of data is quickly getting out of control due to its many different sources, rapid growth, dispersion throughout the organization and non-physical nature.

**Why is the issue important?**
Organizations are increasingly relying on data to implement competitive strategies, manage business risk and comply with applicable regulations. As data becomes ever more important to business strategy and operations, being able to manage this key business asset effectively and efficiently is going to be critical to an organization's survival.

**What can be done?**
Organizations need to put in place governance over their data assets. This guideline discusses some of the key steps and considerations in developing a data governance program, including exploring some of the required roles and responsibilities, policies and procedures, and technology.

**Who is this guideline for and how can it be applied?**
This guide is directed primarily at practising Chartered Professional Accountants (CPAs) in mid- to senior-level roles. It is intended to provide CPAs with a practical perspective on how a data governance program might be created to enable them to play an informed and meaningful role in their organization's data journey.

This guideline can be applied by small- and mid-sized public, private and non-profit organizations. While the underlying concepts will generally be applicable to larger more complex organizations as well, this guideline does not attempt to provide direction on designing or implementing data governance programs for such organizations.

**MANAGEMENT ACCOUNTING GUIDELINE**

# Overview

## What's inside?

This guideline discusses how organizations can employ data governance concepts to help harness the ever-increasing volumes of data they are collecting and generate value from this rapidly growing asset. It provides a four-step process to help CPAs:

1. Determine the need for data governance and desired outcomes
2. Identify stakeholders who will be involved
3. Prepare a roadmap for where they want the program to take their organization
4. Design the program

## How emerging trends impact your business

Almost all organizations regardless of size have started on the digitization journey. From email to electronic bank statements to automation to autonomous vehicles, data is transforming operations across industries and sectors. With this increased reliance on data comes increased focus on how data is acquired, maintained, stored, used, presented and disposed. Organizations must get all these aspects of data management right to stay competitive.

While artificial intelligence and analytics in support of decision-making grab the largest share of data-related headlines (alongside data breaches), organizations implementing much less sophisticated data-driven technologies are also achieving huge impacts. As noted in CPA's guideline *From Data to Decisions: A Five-Step Approach to Data-Driven Decision-Making*:

"Data is not just fueling decision-making, though. It is also transforming the way forward-thinking organizations operate, particularly in the form of automation and improved business processes. In short, data has never been so important."

Improving business processes is a significant opportunity for most organizations as it runs the gamut from changing how we interact with customers and suppliers to changing how we deliver products and services. At the same time, the increased use of data presents many new risks organizations need to navigate to safely deliver the benefits data promises.

Overview

Introduction

Process

Application

Key learnings

Resources

# Introduction to the topic

## What is data governance?

Data has quickly become a key institutional asset to many organizations. As with other assets, organizations need to exercise governance over their data to optimize its value while at the same time ensuring it is protected and used appropriately within ethical and legal bounds.

Data governance can be defined as a collection of processes, roles, policies, standards and metrics that ensure the effective and efficient use of information in enabling an organization to achieve its goals (CPA Canada, 2020). Put another way, organizations need data governance to enable them to take advantage of data opportunities when they arise while at the same time managing the risks associated with the data value chain (CPA Canada, 2020).
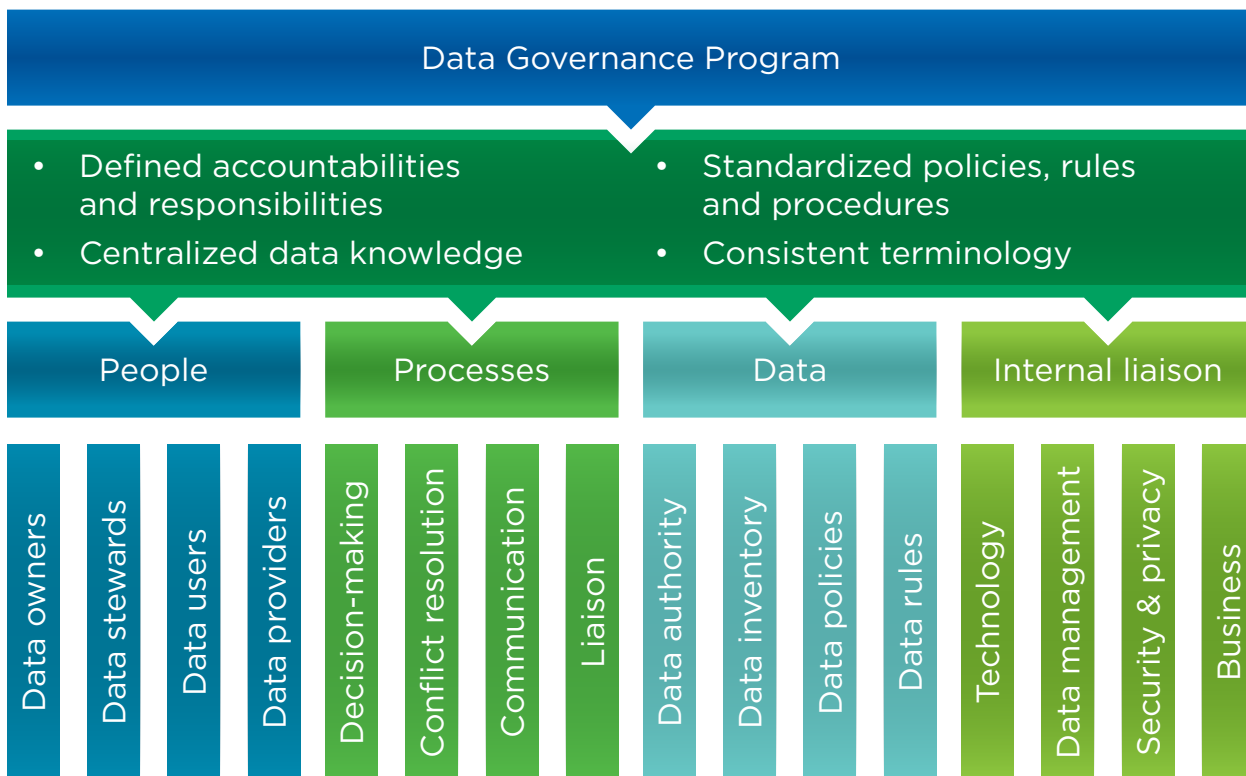


The goals of data governance are to ensure that an organization's data assets achieve the objectives set for them (i.e., contribute towards achieving organizational outcomes, resolve conflicting requirements and manage the risks data may present to the organization).

At a more tactical level some key objectives of data governance are to ensure:

- data across the organization is treated with a consistent approach both in terms of establishing standard terminology and using preferred business practices

- there are clear reporting lines to enable effective and efficient communication on matters and decisions related to data

- data can be effectively integrated across the organization

While some data governance initiatives may be broad and applicable to the whole entity, many initiatives have more narrowly focused purposes or drivers, such as to:

- support specific projects involving data-based decision-making or automation

- respond to regulatory requirements involving data including those associated with specific industries

- support the implementation of new information systems

- support a new business strategy involving data

A well-designed and implemented data governance program will manifest itself in different ways to different people depending on their role and the scope of the program.

- **Users** will generally find it easier to locate, navigate and access the organization's data assets; the data will be of higher quality and have greater consistency across the organization; and they will find guardrails in place regarding how the data can be used.

- **Business leaders** will find it easier to work through situations where their data-related requirements conflict with those of another department; and it will be easier to source new data sets.

- **Systems integrators** will find data has greater integrity and doesn't exist in multiple different versions across the organization; it will be easier to gain an understanding of the data that will interact with and flow through their systems.

## Why is data governance relevant?

To derive value, data assets need to be:

- relevant, high-quality, and readily accessible in a timely manner

- used and interpreted consistently across the organization

- adequately protected from access or alteration by unauthorized individuals

- used for appropriate and ethical purposes

These requirements are continually under threat, which is why data governance is relevant. For the most part, except for data access, these needs are primarily under threat by "business as usual" processes.

For example, data can enter an organization via multiple routes, each of which presents a risk to the quality of data within the organization's systems. Without the necessary mechanisms in place, data coming from third-party business partners, third-party service providers, online customers and even a business's own employees all have the potential to impact data quality.

Overview

Introduction

Process

Application

Key learnings

Resources

One of the roles of data governance is to ensure that appropriate data quality standards have been implemented and the required controls to achieve those standards are in place at all entry points.

With the explosion of data being collected and the growth in computing power, users of data can analyze data in ways they had never dreamed of before. These innovations also increase the risk of unethical uses of data.

As the volume of data that businesses collect continues to grow and organizations find ever more innovative ways of deriving value from it, the relevance of data governance is only going to increase.

## What is a data governance program?

A data governance program comprises the people, processes, technologies and policies required to exercise governance over an entity's data assets.

Its mandate includes:

- implementing a decision-making process as it relates to data, including formalizing the roles of data decision-makers (i.e., data owners)
- establishing data-related rules and policies
- facilitating the effective and efficient use of data across the organization and with business partners

Data requirements are driven by business strategies. As new strategies are developed and implemented, data requirements will change and the data assets of an organization will need to evolve over time. These changes need to be agreed upon within the organization.

Data owners and to some extent other stakeholders must approve:

- the data requirements (the "what")
- the source of the data (the "where")
- the processes for looking after the data (the "how")
- the users who can access the data (the "who")
- the data's uses (the "why")

In some cases, the scope of data requiring approval may be very specific to a particular business process. However, in many cases the scope will be broader and governed by enterprise and departmental policies. Unless there are existing data policies in place, new policies may need to be created and maintained to address areas such as access, usage and quality.

A data governance program oversees the creation of these policies and rules, ensuring that they are maintained over time and that relevant stakeholders have access to them when needed.

Overview

Introduction

Process

Application

Key learnings

Resources

While certain policies and rules may be specific to a given data domain, there will be many cases where the policies and rules will span across more than one part of the organization. This may mean the involvement of more than one data owner and potentially other stakeholders from across the organization, increasing the potential for conflict.

A data governance program designs the decision-making process and establishes the roles and responsibilities associated with that process as well as other processes relevant to data.

More and more data is being shared within organizations: from enterprise data (e.g., customer information) to departmental data (e.g., sales returns data) to third-party data (e.g., weather information one area of the organization has purchased and is willing to share with others). There are significant benefits to sharing data efficiently across the organization but there are also significant challenges. How do users find out about the data's existence? How do users know whether the data is reliable and can be trusted? How do users know whether the data is the definitive source or whether there are multiple different versions of the same data? How do users know how to interpret the data correctly? Answering any of these questions can be extremely difficult without a central authority on the organization's data assets.

A data governance program establishes that central authority and source of knowledge about the organization's data assets. Setting up good communications channels and knowledge sharing about the data assets is another key element of a data governance program. Where multiple versions of the same data exist, the processes put in place by the data governance program will help establish one version of the truth and handle any related conflicts within the organization.

## Data management and other supporting functions

The data governance program operates in conjunction with a data management function, which is required to implement the policies, rules, controls and processes to maintain data assets. The data management function will typically manage the data quality of the data assets and play a role in the systems and technology users employ to access the data, including the necessary data stores. Whereas the data governance program tends to be more a business function, the data management function is more an information technology (IT) function.

Similar to how the data management function looks after data quality and accessibility, other areas of the organization will typically play supporting roles in relation to data security and privacy. The data governance program is responsible for defining who should be able to access the data, but it is these other functions that implement the controls and processes to ensure the organization complies with the respective policies and rules.

## What is the risk of not having a data governance program?

Below are some illustrations of the impact that not having a data governance program could have on an organization's operations.

Overview

Introduction

Process

Application

Key learnings

Resources

### Speed of executing data initiatives

Imagine an appliance manufacturer is developing an exciting new appliance that combines a refrigerator and freezer into one appliance promising to save customers space and cut energy costs. The head of product development contacts the company's data analysis expert to understand the size of the potential market within the existing customer base.

Given the *simple* request the head of product development hopes to have the results by the end of the week.

The data analyst contacts the IT department to get access to the past five years of sales data. The IT manager asks whether the analyst wants commercial, big box or independent retail sales data as they are stored in different systems. He also lets the analyst know they don't have the authority to provide the data without first getting approval from each of the respective sales departments. The analyst asks what the approval process is but the IT manager doesn't know if there is an established process.

The analyst contacts the sales director for each department. Upon learning of the requirement, the commercial sales director immediately provides an email authorizing access to their sales data. The big box sales director requests a full business case to be completed as she'll need to get further approval from their big box customers. The independent retail sales director responds that due to the personal nature of the data the company's privacy officer will need to be consulted.

At the end of the week, instead of providing the results to the head of product development, the analyst reports back that they are in the process of obtaining the data for the analysis but does not yet have a timeline of when they will get access to the data. The head of product development has to ponder whether to take a timely decision based on limited information or delay his analysis to rely on comprehensive data.

A data governance program would have helped this situation by treating the sales data as an asset, making it more accessible via a data warehouse, for example, and putting in place streamlined processes for obtaining access to the data.

### Regulatory compliance

The Proceeds of Crime (Money Laundering) and Terrorist Financing Regulations requires specified reporting entities to verify the identity of their clients (aka the Know Your Client requirement). When this requirement was implemented, many organizations suddenly had a significant exercise on their hands to achieve compliance.

As an example, the existing customer data held by many organizations was often incomplete, out of date and of relatively poor quality. In some organizations the problem was made worse by fragmented systems operating with different data standards. This presented many organizations with a huge challenge to achieve compliance within the regulatory deadlines.

Overview

Introduction

Process

Application

Key learnings

Resources

Having a data governance program does not provide organizations with a crystal ball of what will be required in the future. However, a data governance program may have helped these organizations by having enforced data quality standards and policies, which would have made achieving regulatory compliance that much easier.

### *Calculation errors*

Organizations sometimes have the same type of data stored in different systems in different ways. For example, one system might store interest rates as a straightforward percentage (3.99 per cent), another might store it in basis points (399 basis points) and yet another system might store the value as a fraction to four decimal places (0.0399). It's easy to see how this could lead to confusion and error particularly if the data is passed through different systems within the same organization. This problem might also exhibit itself purely by having systems store the interest rate to differing decimal places or perform calculations to different degrees of accuracy.

One of the mandates of a data governance program is to specify how a particular type of data, in this case interest rates, is stored within the organization. In this way users can be certain they are using data in their calculations correctly.

## What is RAISE and how does it apply to data governance?

CPA Canada's RAISE philosophy is founded on the premise that companies need to be Resilient, Adaptive and Innovative to be Sustainable Enterprises. In this world of exponentially expanding data volumes, organizations should look at data governance as an important underpinning of being resilient, adaptive and innovative.

Resilient organizations need the capacity to respond to the unexpected. A data governance program can enable organizations to trust their data at a moment's notice rather than having to spend days or weeks validating it before it can be used. This can enable them to quickly assess the impact of an unexpected event and develop a timely recovery strategy.

Adaptive organizations need to be able to sense what is going on in their business and monitor for changes that occur over time. A data governance program can help businesses safely and securely collect and store data that can provide signals as to the most subtle of changes, enabling the organization to adapt in a timely manner.

Innovative organizations need to take risks and challenge the status quo. Data can help assess and monitor risk, enabling organizations to innovate while at the same time being aware of how their risk profile is evolving. Equally, data can provide insights to help generate and nurture revolutionary ideas to disrupt and transform business.

Implementing a data governance program enables organizations to become more resilient, adaptive and innovative, which in turn will make them more sustainable.

Overview

Introduction

Process

Application

Key learnings

Resources

## How CPAs add value

CPAs have tremendous exposure to organizational data beyond financial information. With the increasing digitalization of business operations, all kinds of organizational data supports and feeds into a business's financial reporting and business forecasting. Combine this with the competencies and responsibilities encompassed in the CPA designation and CPAs are well positioned to lead or be significantly involved in a company's data governance journey.

Many organizations have yet to capitalize on their data as a key business asset. In many cases this is likely due to a lack of appreciation for the value that can be derived from data. CPAs understand this value and can introduce data-based opportunities to the leadership team.

A data governance program will often involve multiple projects targeting different business initiatives and different data domains. CPAs can apply their expertise to make investment decisions about these projects.

A key component of any project or program is measuring its progress and taking adjusting actions as required. CPAs' experience working with a range of key performance indicators (KPIs) means they are well placed to design relevant data KPIs and be involved in the ongoing assessment of a data governance program.

The competencies embodied in the CPA designation – data analysis and interpretation skills, deep business knowledge, public interest responsibilities, assessment capabilities, strategic decision-making – present the opportunity for CPAs to become chief data officers (CDOs), a senior executive position responsible for the data assets of an organization.

# Process

There are four steps to developing a data governance program.

**Step 1**

### Determine where to start
The purpose of this step is to define the strategic benefits and tactical outcomes that the program is intended to produce. This vital first step ensures that the program has clearly established goals.

**Step 2**

### Identify stakeholders and their roles
This step involves identifying the roles that will be required, who will fill them and from where in the organization those individuals will come. This important step makes it clear to everyone who owns what data and where ongoing data responsibilities will reside.

**Step 3**

### Prepare a roadmap
This step is about painting a picture for all interested parties as to what the initiative will look like from initiation through to achieving the program's strategic and tactical objectives.

**Step 4**

### Design the program
This step focuses on the identification and design of the elements required by the program, including the program and people structure, policies and procedures, data rules and technology.

**Next Steps**

Once the above steps have been completed, the next steps will be to create an implementation plan, execute on that plan and put in place the appropriate measures of success.

Overview

Introduction

Process

Application

Key learnings

Resources

# Applying the topic to your organization

**Step 1**

## Determine where to start

The purpose of the first step is twofold:

- Describe the strategic benefits the data governance program is expected to bring to the entity. These will usually be tied to the goals of the underlying business initiative.

- Describe the tactical outcomes the data governance program is expected to deliver and what these will enable the entity to do differently.

Being able to describe the benefits of the program will help in selling it to executive leadership as well as provide a guiding light for the team members who will be involved in building and implementing the program. The ability to articulate the practical outcomes of a program can help to generate excitement and support for the program especially among those stakeholders who will directly benefit.

### Identify the need for a data governance program

The needs and strategic objectives of the business will drive the need for a data governance program. Most business initiatives require or involve data in some shape or form. If an initiative has data associated with it, then it is likely that a data governance program will be beneficial. The maturity of the organization, and size and scope of the initiative will determine the extent of the data governance program.

While data governance programs can be implemented at the enterprise, department or project level – or some combination of these – it is more manageable and practical to start with a focused initiative. Once the program is piloted in this way and proves successful, the learnings can be applied and scaled across the organization.

### Define objectives and intended outcomes

Having determined the need for a data governance program, the next step is to define and align specific program objectives with those of the strategic initiative it supports.

The first example below focuses on a specific aspect of the business whereas the second example touches on data that is pervasive across the business. This will be mirrored in the breadth of the associated objectives.

Overview

Introduction

Process

Application

Key learnings

Resources

*Self-serve investment management system example*

A small financial institution is initiating a project to provide customers with online access to manage their financial investments. The project will involve implementing a new web-based investment management system. Data about the entity's customers and investments will need to be converted from the entity's legacy systems and loaded into the new system. The data might be spread across more than one legacy system and could be of poor quality and low integrity. A data governance program could be used to improve data quality, streamline the conversion of data into the new system and provide ongoing governance over that data.

In this example, the goal of the data governance program will be to ensure that clients using the online system have timely access to complete and accurate data about their investment portfolios. This goal would be supported by:

- designing and implementing the necessary organization, policies, processes, standards and technology to maintain high-quality client investment data on a forward-looking basis

- getting the existing client data into the new program, which includes bringing it into compliance with the forward-looking policies and standards

The benefits of improving data are often wide-ranging because data touches so many aspects of the business. In our example, as well as enabling online self-service for investment clients, a program that maintains high-quality client data might also benefit the following areas:

- customer service – reducing time-consuming calls about incorrect client data will allow agents to focus on providing more fulfilling higher value services

- tax and year-end reporting – higher-quality data will generally reduce the degree of manual intervention and corrections required enabling cost-effective, timely reporting

- customer satisfaction – enabling customers to go online and transact whenever they wish knowing that the data will be correct and up to date will leave clients feeling in control of their investments

*Retail chain management reporting example*

An organization has a strategic goal to improve the management reporting across a retail chain that has been built through acquisition. This project would require standardizing the data about each store (e.g., implementing a consistent store identifier and capturing a consistent set of attributes for each store). This data would be considered enterprise data as it would be used across the business including sales, fleet management, human resources and inventory management. Given the scope of the data, this type of project would positively impact multiple initiatives across the organization, not just the initiative that triggered the program. A data governance program could be used to standardize the entity's store data and provide ongoing governance over that data.

In this example, the initial objective was to meet management reporting requirements. However, due to the enterprise-wide nature of store data, the data governance program objectives could be stated much more broadly:

- to enable business initiatives to be implemented efficiently and effectively by providing consistent, accurate and timely store data when needed

In defining the intended outcomes, the various uses of store data across the entity should be considered to help communicate the extensive benefits the program has the potential to deliver. This could involve surveying the various departments across the business to both understand existing initiatives and spark ideas for new ones that could benefit from access to such data.

This might include business insights initiatives for the sales team or automation of fleet routing for the logistics department. Associated outcomes could be:

- improved sales thanks to the generation of insightful store-on-store comparatives

- reduced fuel costs because of more efficient store replenishment deliveries using automated routing software

- new high-potential store locations based on spatial modelling tools

A data governance program could contribute towards increasing sales, reducing costs and helping the business grow.

## Identify stakeholders

**Step 2**

Who will be directly and indirectly impacted by the initiative? Given the extent to which data flows within and between organizations, it is highly likely there will be various external as well as internal stakeholders.

In terms of project management's RACI matrix, "responsible" and "accountable" parties will generally be internal to the organization, whereas "consulted" and "informed" parties are just as likely to be external to the organization as internal.

### User stakeholders

Consider the objectives of data governance when identifying the relevant stakeholders. For example, two common objectives are to make it easier to exchange data within and outside the organization, and to ensure that all users understand the data so that it is used and interpreted appropriately. This example would lead to the identification of both internal and external users of the data, so they could be appropriately consulted and/or informed during the implementation of the data governance program. Engaging these stakeholders at this time would help to prevent potentially costly issues in the future.

Data ownership is another important consideration when identifying relevant stakeholders.

Overview

Introduction

Process

Application

Key learnings

Resources

Returning to the first of our business examples, changes to the data within the investment management system will impact customers and may impact third-party service providers such as investment brokers, fund managers and custodians. Each of these stakeholders will have different concerns about any impending changes depending on how they interact with the data in question. In the case of customers, it is important to consider their role as owners of their personal data.

Identifying stakeholders can be more challenging when there is not a comprehensive understanding of who uses what data and for what purpose. In our second example, where the initiative involves store identifiers and other store-level data (i.e., enterprise data), it may involve much more work and communication to understand who is using this data and to what end. It is common in many organizations for "off-the-radar" business initiatives to take existing data and make use of it for purposes that were not envisioned when the data was originally collected.

Another type of "stakeholder" to consider are those who do not currently have a stake in the data but who might wish to be made aware of changes that could benefit them in the future. For example, this could be a part of the business that doesn't currently use the data because it is not of sufficient quality or scope but want to help shape the data governance program to ensure going forward they will be able to use the data being collected to achieve its business objectives. Again, this is a more challenging set of stakeholders to uncover depending on the size of the organization.

Other types of users such as executive management and those who perform business functions that clearly involve the data in question will be much easier to identify.

### Direct stakeholders

Direct stakeholders are more likely to have a formal role within a data governance program. The roles and titles associated with data governance will vary depending upon the maturity of an organization's data journey. However, the key functions performed by these roles are relevant regardless of maturity level:

- the executive sponsor provides senior level sponsorship for the initiative

- the data owner is accountable for the data in each area or domain

- the data steward is responsible for the data in each area or domain

- the data manager/administrator manages the data on a day-to-day basis

While the executive sponsor role will be the same as for other types of initiatives, the other roles are more specific to the data governance domain.

**Data owners** are the key decision makers with respect to the data. When necessary, they will consult with other data owners within the organization to ensure the impact of any decisions are carefully considered and do not adversely impact other data domains within the organization.

Overview

Introduction

Process

Application

Key learnings

Resources

**Data stewards** perform a more hands-on role with responsibility for overseeing the management of the data on behalf of the data owner. Stewards are responsible for the processes by which the data is collected, maintained and provided to users. A key part of this role is looking after the quality of the data and ensuring it is fit for purpose.

**Data managers** support the data steward(s) and perform the hands-on work of collecting and maintaining the data including assessing and remediating data-quality issues.

In smaller organizations and initiatives, the roles of data steward and data manager might be performed by the same person. In larger organizations, it is more likely that a data management function will exist, usually within the IT department due to the more technical nature of the required skill set.

For example, in a university, the president/vice chancellor may be the executive sponsor. They may appoint vice-presidents responsible for the student, human resources and finance domains as the data owners. In turn, each data owner would appoint a data steward to oversee the data in each domain (e.g., this might be the owners of the student information system, the human resources information system and the finance information system.) And finally, each data steward would oversee one or more individuals whose day-to-day job would be to maintain the quality of the data, amongst other things, within the respective domain. As noted above, in larger institutions these individuals may sit within the IT department with an indirect reporting line to data stewards.

As an organization matures, additional roles and organizational structures may be required. A CDO may be appointed to preside over a data governance office (DGO), be accountable for institutional data and coordinate the data owners. The CDO and data owners might form a data governance council that meets to decide upon institutional and cross-domain aspects of data governance.

In addition to the above functions, many members of an organization will typically be data users who would need access to data to perform their roles and be expected to comply with the data governance requirements.

## Accountabilities and responsibilities

In organizations with more established data governance, accountability would typically rest with the CDO. However, for organizations starting at a smaller scale this responsibility might sit with the business owner for the initiative or the business sponsor if the data governance initiative is part of a larger project.

In the investment management example, some of the project roles for the system implementation could be combined with those of the data governance program. For example, the executive sponsor for the systems implementation might also sponsor the data governance program, and the business owner of the investment management system might also be the investment management data owner.

Overview

Introduction

Process

Application

Key learnings

Resources

Accountabilities for the retail chain example are not necessarily as obvious due to the enterprise nature of the data. Accountability could temporarily sit with the department that initiated the project and then later transfer to the appropriate business area; or it could immediately sit with those who would take those roles in the longer term.

In this example, the data governance initiative could be sponsored by an executive in the finance area, where the project was initiated, or equally someone in the retail operations part of the business where ownership of store data might ultimately reside. Similarly, the data owner for the purposes of the initiative could come from either area. Where possible the data owner role should reside with someone who has significant familiarity with the relevant data, which might help determine whether the role sits in the finance or retail areas. As noted earlier, in an organization with a more mature data governance function the ownership of enterprise data would reside with the CDO.

Where the stewardship and management roles reside will also depend on the nature of the initiative and the data involved. Familiarity with the data is important but equally so is the need for relevant skill sets. In more mature organizations, the CDO may oversee a DGO staffed with data stewards and managers who specialize in data governance.

## Step 3

### Prepare a roadmap

Executive sponsors and program participants will want and need to understand how a data governance program will play out. They need to have an expectation of where the program is going, how it will change the way the organization will operate in the future and ultimately what outcomes they can expect from it.

Once the program's objectives and participants have been determined, it is important to build a roadmap that brings the project to life for the key decision makers and participants. This high-level narrative should describe how participants will be involved, when the main deliverables will be produced and how the organization will start to reap the benefits of the program. The purpose of this roadmap will be to sell the initiative to the decision makers and at the same time engage the participants involved in the initiative.

With streams of data, in all its forms, flowing around us every minute of every day it can be challenging to grasp how governance might be implemented over an organization's data. This is analogous to trying to understand how governance might be put in place over the quality of all water, not just drinking water. Many users of data, particularly those at more senior levels, have little idea of the amount of pre-processing that took place before the data appeared in their reports. This would be akin to what goes on at the water treatment plant before safe drinking water is fed through our pipes. Now extend that to the water in rivers, lakes and oceans and the enormity of the challenge comes into focus. Similarly with data, some users will find it challenging to comprehend the enormity of the task at hand while others who do understand the enormity may find it difficult to see how an organization could harness such a pervasive asset.

Overview

Introduction

Process

Application

Key learnings

Resources

This challenge can be exacerbated by the terminology that often shrouds technical domains. To avoid confusion and misinterpretations and make it easier for those not familiar with the terminology, establish and maintain an easily accessible glossary of technical terms specific to your organization's initiative.

Having a roadmap that explains in plain language what is going to happen over the course of the initiative is essential to getting all participants on board and working towards a common goal. The roadmap may also contain diagrams that can help both decision makers and other participants visualize the process, further helping them to understand what they are approving and getting involved in. Finally, the roadmap might include interim goals, to recognize and build on medium-term successes and gains.

Once the initiative is approved and underway, participants can continue to use the roadmap to see what progress has been made, where the initiative is at and what lies ahead.

## How does the roadmap tie back to the objectives of the initiative?

In the investment management example, the data governance roadmap will be heavily intertwined with the plans for the main systems-implementation project. The roadmap might include the following high-level tasks:

- identify the key stakeholders in investment management data and their roles

- understand the project's (and organization's) data requirements as they relate to investment management

- understand what investment data the organization already has, where it is stored and its quality level

- identify what processes, if any, already exist surrounding the collection and maintenance of investment management data

- identify what technologies, if any, already exist to collect, maintain and manage investment management data (e.g., data quality software)

- assess the existing people, processes and technologies against leading practices ("a gap analysis") and determine the actions required to close the gap

- in consultation with identified stakeholders:
  — define data quality rules
  — define data policies (e.g., for the usage of investment management data including privacy and ethical aspects)
  — define processes for the ongoing maintenance and management of investment data
  — determine additional technologies required to support the new processes

- define and implement governance processes (e.g., to provide assurance that policies are being followed and exceptions identified by processes are addressed)

Some of these tasks, such as understanding the organization's existing investment management data, will overlap with those required by the main project. Often a project's primary focus is to convert the data efficiently and effectively into the new system (a one-time

Overview

Introduction

Process

Application

Key learnings

Resources

event) whereas those tasked with data governance are looking at the data as an asset that requires longer-term considerations (an ongoing process).

While similar in many ways, the tasks for the retail chain example would differ due to the data having an enterprise-wide scope and the initiative's focus being to ensure standardized data is available for all stores. For example, one additional task related to store data might be to source new data and fill any gaps in the existing data.

While data governance programs will have some differences to other types of governance programs, the general stages are similar:

- identify key stakeholders
- collect information on the current state
- perform a gap analysis between the current state and leading practices
- determine measures to close the gap
- define business rules and policies
- implement governance processes to ensure policies and procedures are followed

## Step 4

### Design the program

Once the initiative has approval to proceed, the next step is to design the data governance program. One way to look at this is in terms of people, processes and technology:

- what people and organizational structures will be needed?
- what processes will need to be executed?
- what technology will be needed to support these processes?

### People

One of the key imperatives of a data governance program is to create communication channels to help ensure that the appropriate people are:

- involved in data-related decision-making
- consulted about potential changes to data that are relevant to them
- informed about the outcome of data-related decisions and other changes to data that may impact them

In addition to identifying the appropriate people, it will also be necessary to get them on board with the initiative and determine what means of communication will be used.

The data owner is responsible for making decisions over a certain data domain. In our investment management example, this might be the business owner responsible for the investment management system. In our retail example, where store data is used across many parts of the organization, the responsibility is less obvious. In a more mature organization with a CDO, responsibility might sit with them. However, for an organization just starting out

with data governance it might make more sense for it to sit with the business leader for retail operations or a finance lead or possibly a combination of the two.

Data stewards perform the role of data domain experts and have responsibility for ensuring that data governance policies and procedures are implemented. They would be consulted on any potential changes to the data, effectively acting as proxies for the data owner.

Initiatives that involve more than one data domain will have multiple data owners and stewards. In the investment management example, ownership of the customer data, the investment data, the customer transaction data and the associated finance data might sit with different parts of the business, each with their own respective data owner and data expert (steward).

In this example, the data lead for the main project might establish a data stewards council with stewards from each area that meets regularly throughout the course of the project. The data owners would be invited to the meeting when decisions need to be made. Once the project is over, the data stewards council would ideally persist as part of the organization's ongoing data governance program. The role of the chair would then transition from the project's data lead to the CDO.

In many situations, there may not be a data domain expert as such, and one will need to be appointed for the purposes of the initiative. In the retail example, this could well be the case, particularly if store data had not previously been thought of as an organizational asset. In this case, one could see an opportunity for a member of the finance reporting team to get involved; someone familiar with manipulating and analyzing data for financial reporting purposes.

Data users will often be the largest and most widespread community of stakeholders coming from both inside and outside the organization. In the investment management example, the user community could include the following groups:

- customers
- customer service agents
- investment management consultants
- investment product specialists
- investment brokers
- investment custodians
- finance analysts
- tax compliance
- regulatory compliance
- regulators

The most suitable method and frequency of communicating with these groups will depend on the nature of the group. In some cases, such as for customers, a periodic update newsletter might be sufficient, whereas it may be necessary to establish formal scheduled meetings with the regulator to keep them apprised of progress and timing.

One key role that has been touched on but not formally discussed is that of the data governance lead; someone needs to be accountable for leading the data governance program. In larger, more mature organizations this role would typically be filled by the CDO with the program being run by a DGO. However, in organizations just starting out someone else will need to take on this role. Where the initiative is linked to a systems implementation project, such as the investment management example, the data lead on the project might take on the role. In the retail example, it is not so clear who should take on the role in the absence of a CDO. It might make sense for the role to be fulfilled by the data owner.

The roles discussed so far have all been management-type roles that sponsor, make decisions, provide governance and generally oversee the program. It will also be necessary to have team members who have the hands-on skills to work with the data and perform tasks such as data quality assessments, data analysis and data manipulation. These tasks generally fall under the umbrella of data management, which may or may not sit within the data governance team.

Individuals with the required skills often reside within the IT department although this doesn't have to be the case. Depending on the nature of the initiative and the organization, such individuals might be co-opted to work on the data governance initiative or provide their expertise directly from their place in the IT department.

## Process

Every initiative will have its own unique process to achieve data governance objectives but the basic steps will be similar.

In the investment management example, the following high-level steps will be required:

- Understand the existing data and what is needed by the new system. Much of this would typically be performed by the systems implementation team although more with an eye to executing the data conversion rather than longer-term governance of the data.

- Develop data quality rules and controls, any new policies (e.g., data usage policy) and any necessary amendments to existing privacy and security policies. The drafts would be reviewed with relevant stakeholders prior to final approval by the data owners.

- Document and publish rules, controls and policies.

- Communicate rules, controls, policies and implementation expectations to stakeholders.

- Implement ongoing assessments to measure progress in achieving data governance objectives.

The diagram below captures the above steps. While it suggests a sequential process there may be some back and forth between the different steps.



Depending on the initiative, there may or may not be a continuous cycle. The first time around the steps may be more linear and informal in nature with more formality coming later as the benefits of data governance start to be realized and permanent processes are implemented.

### Research, analyze and consult

The first step is to gain a detailed understanding of the existing data and plot a path to where the initiative intends to take that data. This includes a hands-on exercise known as "data discovery," which involves analyzing the existing data and assessing its quality.

As well as analyzing the data this step will involve consulting with relevant stakeholders about the existing data as well as the intended future state. This will provide a richer understanding beyond analysis alone. Speaking to users of the data including customers and suppliers can also help to understand expectations of the user community.

### *Propose, discuss and approve*

The next step is to develop the proposed rules, controls and policies needed to help achieve the desired objectives. This may include rules defining what constitutes acceptable data quality, preventative controls to protect data quality going forward and policies to govern who can access the data and how it can be used. The focus of these elements will depend on the nature of the initiative and the desired outcomes.

Once proposals have been developed, they may need to be reviewed with relevant stakeholders before final submission for approval by the data owners.

### *Document rules, controls and policies*

This step may be a separate step to create formally documented rules, controls and policies, or it may have been largely covered in the prior step. Either way, it is important to document what has been agreed upon in sufficient detail so that it can be communicated to relevant stakeholders.

Some examples of policies that support data governance are:

- privacy policy
- security policy
- data usage policy
- data disposal policy

These are not necessarily dedicated to data governance and many organizations may already have them in place. As part of a new initiative, it is important to review existing policies to ensure they encompass any new scenarios.

### *Communicate*

The penultimate step is to make sure that stakeholders know what has been agreed to and what is expected of them. Ideally, certain stakeholders will have been so involved in the process that the communication stage is more about formalizing the rules, controls and policies than the communication of new information.

### *Measure*

The final step in the program is to design and implement metrics that can be used to determine the progress of the initiative and ultimately its success.

When an organization is piloting data governance for the first time, it will likely look more like a project than a program with the project executing the process outlined above. As the organization moves beyond the piloting stage and begins to adopt data governance on a broader scale, it may then move to create a program. Such a program would likely comprise one or more projects with some coming to a natural conclusion and others reaching a point where the ongoing aspects of data governance are transitioned over to be run by the program's DGO.

Overview

Introduction

Process

Application

Key learnings

Resources

## Technology

Data governance programs involve not just governing the data itself but also managing data about the data or metadata. Along with communication, this aspect of a data governance program can take up a significant percentage of the available resources.

Some examples of metadata and why it is needed include:

| Metadata example | What it is | Why it's important |
|---|---|---|
| Names of data fields | Identifies the categories of data so it can be referenced in queries and programs. | Data is much easier to reference when the names assigned to it use a common naming convention. The same data may also be used in multiple systems so consistent naming is helpful. |
| Data access rights | Describes who can access what data and what they can do to that data (create, read, update, delete). | Access rights provide a means of giving users an appropriate level of access to the data, which is a fundamental aspect of security and supports privacy requirements. |
| Data lineage | Describes how data got to its current state (e.g., its source and what changes have been made to it). | It is important for data users to have a clear understanding of the data they are using. If data has been changed in some way between when it was sourced and when it is used this is important to know. |

These are just a few examples. However, it is easy to see the amount of metadata itself can be quite voluminous.

There are sophisticated tools available on the market to help efficiently manage metadata across multiple systems.  That said, in the early stages of developing a data governance program it may make sense to track this information within the individual systems. Where necessary, spreadsheets can be used for information not typically tracked in systems or where information needs to be brought together across systems.

Overview

Introduction

Process

Application

Key learnings

Resources

# Key learnings

Many organizations have bountiful amounts of data available to them yet much of that data is unusable or cannot be accessed in a timely manner. This guideline helps organizations start down the path of treating their data as a business asset they can use to create value and gain competitive advantage.

The four steps outlined in this guideline will help organizations determine when and where a data governance program might be beneficial, identify the stakeholders who will implement it, prepare a roadmap for where the program will take their organization and plan the elements that will comprise it.

Creating a data governance program requires resources, effort and perseverance. At first, the pervasive nature of data may make the task seem unfeasible.  However, starting with a small scope and tightly defined objectives can help build momentum prior to tackling more organization-wide data challenges. As with other assets, data will require related standards, policies and procedures and ongoing maintenance to retain and grow its value.

Two of the biggest challenges with getting and maintaining control of your data are the ease with which it can be duplicated and the lack of clear ownership due to its pervasive nature. Communication and clearly defined roles and responsibilities are key to getting your data governance program off on the right track.

When managed appropriately, data can become one of your organization's most valuable assets, providing competitive advantage, mitigating business risk and achieving regulatory compliance. One day these previously untamed assets may appear on your balance sheet!

Overview

Introduction

Process

Application

Key learnings

Resources

# Resources

## CPA Canada Management Accounting Guidelines

CPA Canada. (2021). *Cybersecurity from the inside out*.

CPA Canada. (2020). *From data to decisions: A five-step approach to data-driven decision making*.

CPA Canada. (2019). Strategy mapping: A six-step process.

CPA Canada. (2018). Scenario planning: Plotting a course through an uncertain world.

## Other CPA Canada publications

CPA Canada. (2021). Data governance – Is it on your board's radar?

CPA Canada. (2020). Data governance primer: Rising to the challenge of a data-driven world.

CPA Canada. (2020). Data management foundations certificate.

CPA Canada. (2020). Mastering data: A series of articles for CPAs.

## About the author

Darren James is a self-professed data geek who has spent much of his career focused on data analytics and helping organizations manage and look after their data. During his time with a big four professional services firm, he built the firm's data risk practice and led the development of its audit analytics function.

Darren has worked in most industry sectors and functional areas. His experience has ranged from supporting businesses through the data aspects of their systems implementations through to the development of global data analytics tools and platforms.

Overview

Introduction

Process

Application

Key learnings

Resources

**cpacanada.ca/MAGs**