
Data Governance

Eating the elephant one bite at a time



A step-by-step guide for implementing Data Governance in a pragmatic fashion with indicative timelines, recommended artefacts and practical tips for success.

Contents

Introduction	3
Data Governance Implementation	
Phase 1	6
Phase 2	14
Phase 3	23
Concluding Remarks	29



Introduction

Data Governance is recognized as an important capability for organisations of all sizes to drive better business decisions. Its also a critical capability to ensure data is managed in a safe and compliant way.

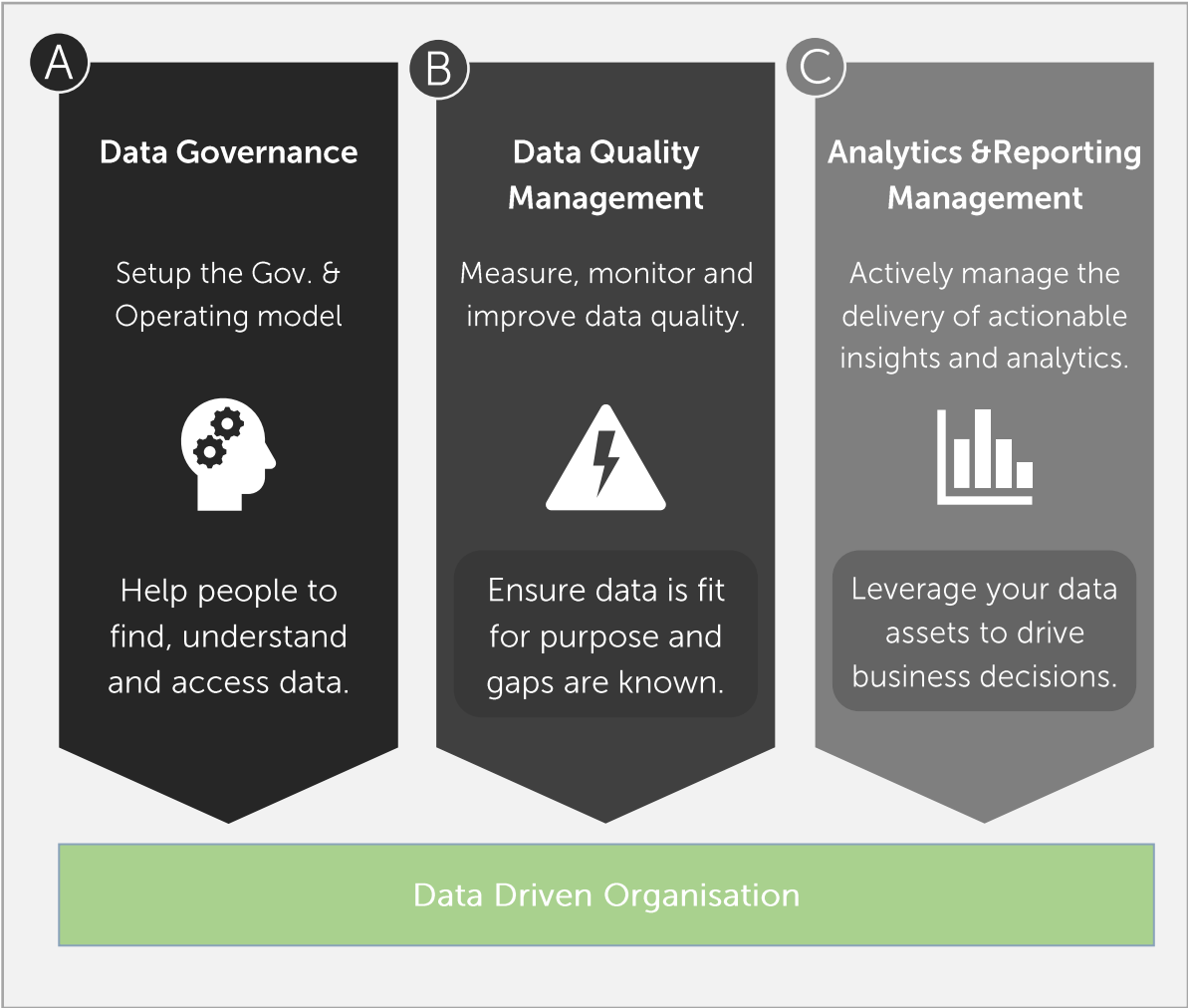
However, data governance is a broad and often unwieldy undertaking. Many organisations do not have vast resources dedicated to data governance efforts and its often left up to a **lone wolf** (or a small pack) to drive data governance uplift across the organisation. With a small resource footprint its easy to become overwhelmed and bogged down in best practice guidelines without knowing how to develop practical capabilities.

You need a pragmatic implementation plan to be successful. You need to be crystal clear on what you want to do and how you are going to tackle the implementation. You need an approach that allows you and your team to be effective by leveraging the brainpower of your organisation to scale. You need to “**eat the elephant one step at a time**”. This document is intended to serve as a blueprint for a pragmatic Data Governance implementation. It’s a step-by-step guide split across 3 key phases and the guide includes indicative timelines and practical tips for success.

3 Pillars of effective Data Management

The approach in this paper takes a view that Data Governance is one of 3 pillars that need to be in place for an organisation to be truly data driven. The first pillar is all about Data Governance and is the focus of this document. It covers the implementation of your Data Governance processes and operating model along with the roles and responsibilities.

Next is pillar B, which focuses on the measurement and monitoring of data quality, along with managing issues and fixing the root cause of data problems. Finally pillar C focuses on analytics and reporting. Pillars B and C will be covered in greater detail in a later document.



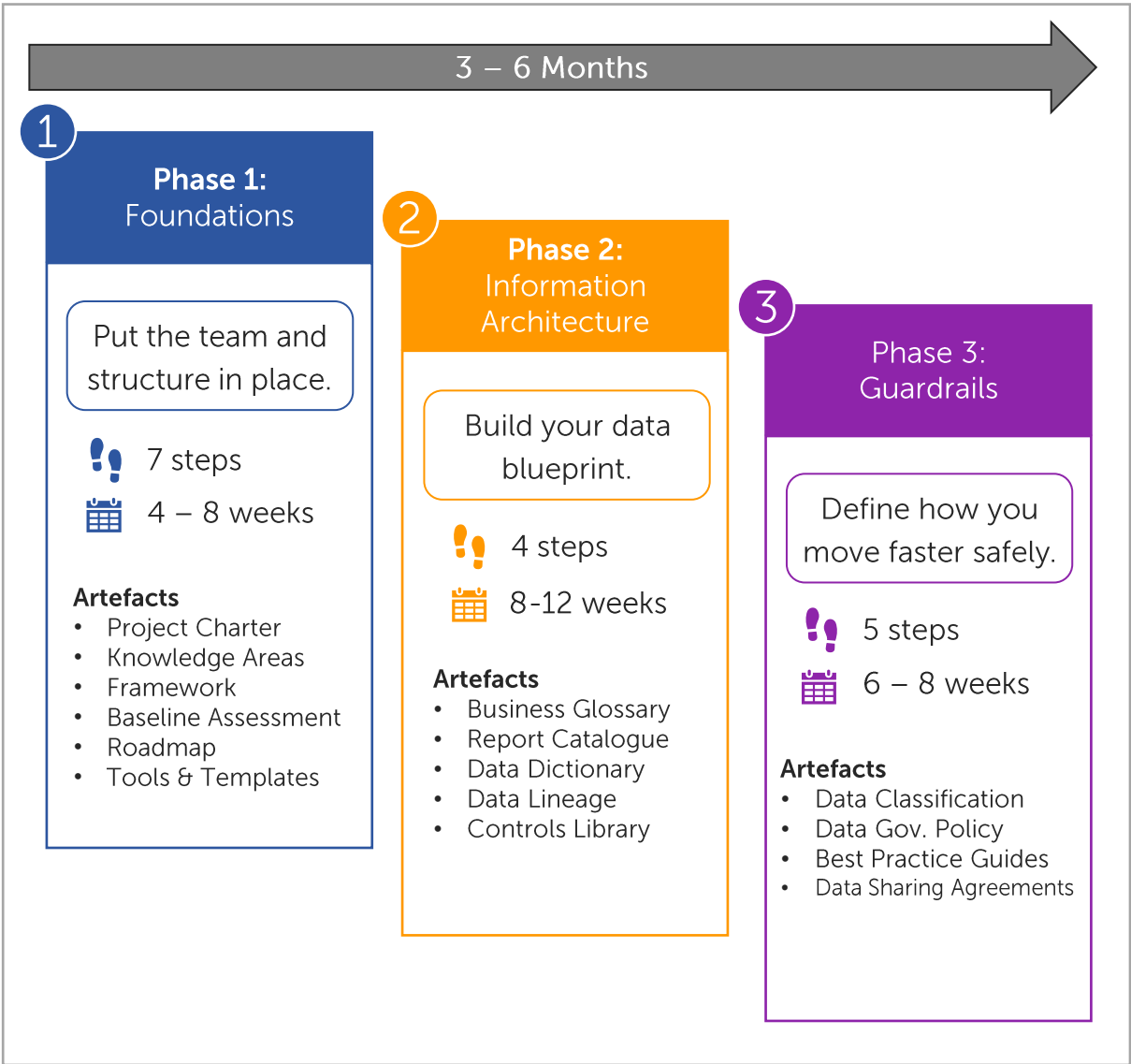
Infornite: 3 Pillars of effective Data Management

Pillar A: Data Governance

This document focuses on Pillar A which is all about Data Governance. The document describes an implementation approach that includes 16 steps split across 3 distinct phases. The objective is to break the initiative into manageable chunks (e.g. to eat the elephant one step at a time 🐘).

Each step has a specific objective and is supported by artefacts that must be completed. We also include practical guidance and common challenges to avoid based on our experience.

Typically, an implementation following this approach will take between 3 to 6 months from end to end. When you work with a partner (like InforNite) they will often provide sample artefacts, templates and training to enable your teams and accelerate your implementation.



InforNite: Phases & Timeline of a Data Governance Implementation

Phase 1: Foundations : Putting the team and structure in place

Good data governance is not really about data; its about people. To succeed, you need to bring together a diverse group of people and equip them with the training and tools they need to develop and execute data governance processes. You must also paint a clear picture of what you want to achieve and craft a vision that people understand and can follow.

Fostering this sense of team and clarity of mission is what Phase 1 of the Data Governance implementation looks at. It's the foundational phase. It covers 7 steps from setting up your charter right through to putting the tools and templates in place.

1 Develop your charter

A charter is a short document that describes what you are seeking to achieve through your Data Governance initiative. It must describe why your organisation is doing this, who your sponsor is and who is responsible for driving the initiative (that's probably you). Your charter must also describe the key reasons you are doing this in the form of short requirement statements such as:

To support our business in making decisions we want to publish a searchable catalogue of all reports that are available across the business today and guidelines for how to access them.

To ensure we are protecting our customers personal information we must document where we store sensitive data today. We must put in place a clear process to ensure we manage who can access this data and ensure people understand how to treat it safely.

Your charter will be a valuable artefact moving forwards. More importantly; it forces you to clearly articulate your reason for doing this and the people that are accountable for making this happen.

Your charter should be no more than 2 pages (double sided if you must). It really shouldn't take more than a couple of days to draft, and you should be able to agree the final document with your sponsor and key team in a couple of review meetings (don't boil the ocean).



You have a short project charter (no more than 2 pages)



You have a committed senior leader that is sponsoring the initiative



You are clear about who is running the initiative

2 Define Knowledge Areas & Ownership

Knowledge Area (also called Data Domains) are a way to group the data across your organisation into high level buckets. As examples, you might have Knowledge Area for customer data, finance data and sales data. Typically, Knowledge Areas will align to your organizational structure or lines of business. They serve a few key purposes for your initiative:

Scope: In the spirit of “eating the elephant one bite at a time”; Knowledge Areas allow you to split your overall scope along these boundaries. Developing the Data Governance capability across each Knowledge Area becomes an initiative of its own and you can prioritize and track the work in each Knowledge Area independently.

Ownership: Knowledge Areas are a mechanism to assign ownership. Typically, a senior leader will be the assigned data owner for each Knowledge Area (e.g., your Head of Marketing will be the data owner for the Marketing Knowledge Area).

Stewardship: Since data owners are assigned to Knowledge Areas it makes sense that Data Stewards are also assigned to Knowledge Areas. This creates alignment between the different buckets of data, the accountable owner and the data steward that is empowered by the data owner to do the work for them.

Categorization: When you publish a business glossary and data catalogues later; your Knowledge Areas are a great way to categorize and search across this information.



Do **not** spend a great deal of time defining your Knowledge Areas. You will not create a perfect and complete set of Knowledge Areas at this stage. Start with something high level that aligns to your organizational structure and evolve these over time. When you start documenting reports and data elements you will need to add Knowledge Areas. After the first few iterations your Knowledge Areas will reach a steady state.

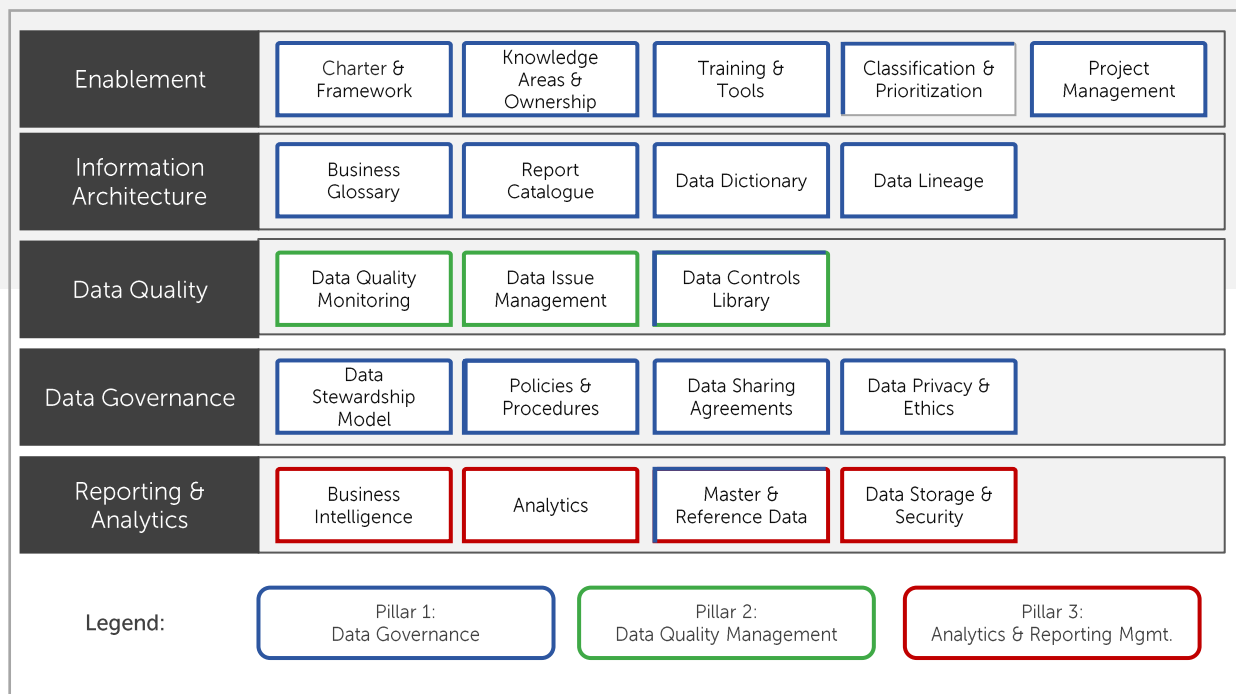
- ✓ You have an initial set of between 5 and 15 Knowledge Areas
- ✓ You have drafted roles & responsibilities for your owners / stewards.
- ✓ You have engaged the data owners for Knowledge Area that you plan to start working on – they have committed to support the initiative.
- ✓ You have identified and engaged data stewards for Knowledge Areas you plan to work on – they have capacity to work on the initiative.

3 Adopt a Framework

A Data Governance framework describes all the things you should be doing and the capabilities you need (in an ideal world). There are numerous data governance frameworks that you can adopt. We often work with the EDM Council's DCAM Framework which is a great choice. You can also develop your own. Either way its good to have one in place:

- It's a valuable reference when planning what you need to do.
- It helps to define the scope of your data governance initiative.
- You can use it to explain your target state to your stakeholders.
- It's a valuable artefact to train your data stewards.

Don't overthink which framework you adopt. Your success will be determined by how well you execute in building capabilities rather than the framework you pick. Spend a little time to review different frameworks and pick one that resonates with you. Feel free to customize the framework but avoid removing parts as you might not realize their value until later. The diagram below represents a simple framework we commonly adopt as its comprehensive and relatively simple. The colors indicate the applicable pillars in the Infornte approach from Page 4.



Infornte: Data Governance Framework

One (often unrecognized) benefit of a framework is that it allows you to call out areas that you are not working on due to resource constraints or business priorities. You can recognize these areas in your comprehensive framework whilst highlighting that they are not on your current roadmap.

4 Train your data owners and data stewards

At this stage you will have identified at least some of your Knowledge Areas and have a few data stewards engaged. The next step is to develop and deliver a persona-based training program to your data owners and data stewards. At a minimal this should cover the framework, the roles & responsibilities and the key artefacts your data stewards need to develop.

There are many training programs and certifications that you can purchase if you have budget including a range of training options that we offer at Infnite. You can also develop your own training programs from the vast range of material freely available online to create an effective training curriculum.

As a Data Governance leader, your success depends on effectively engaging your data owners and data stewards. Training is a critical step to engaging and supporting them. Many of your data stewards will be taking on additional responsibility to be involved in this initiative - its your responsibility to ensure they are trained and empowered to do their best work.



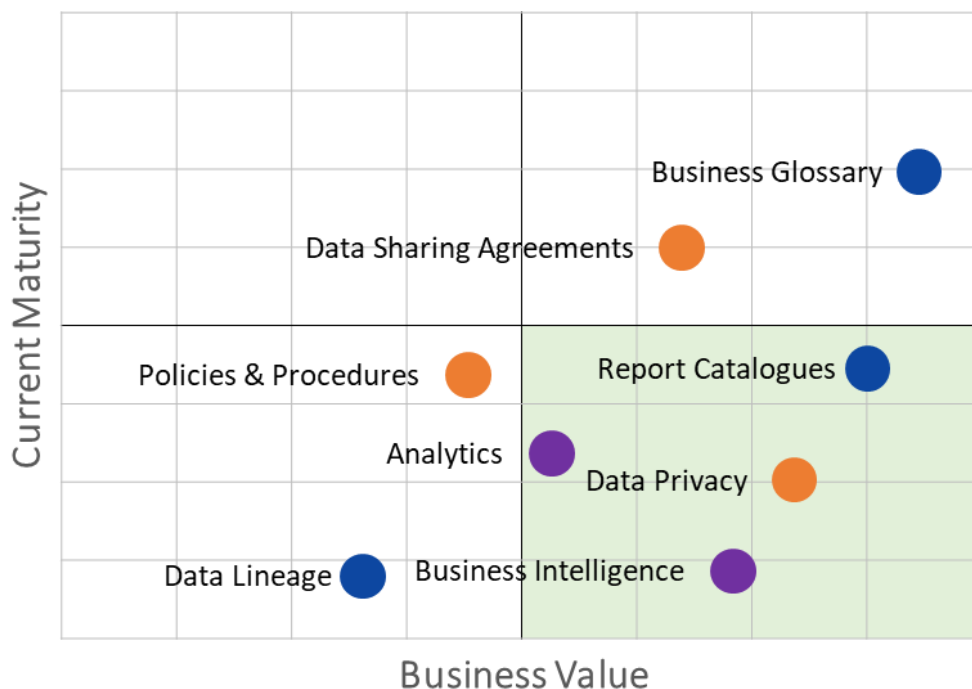
Training is one aspect of the implementation that cannot be time-boxed into a single step. As you start to do more data governance activities you will need to extend the training you offer. You don't need all your training material ready up front, but it is important that you have a repeatable structured curriculum that builds up as you progress on our journey.

5 Assess your current capability

At this stage you have a set of data stewards that understand your framework and have a good understanding of what you are trying to achieve. Now is a good time to use the framework to assess your current capability and identify your strengths and weaknesses.

An assessment doesn't always have to be a major exercise. Its about bringing your data stewards and data owners together to discuss and score your current capability across each aspect of the framework. If you have the budget and time, it can be useful to engage a partner like InforNite that can help facilitate these sessions and provide you with a comparison against your industry.

When doing this step, you should perform a 2-dimensional assessment – this means you assess both the strength of your capability and its importance to achieving your business goals. This recognizes that some capabilities will be more or less important for different parts of your organization. This assessment can feed your roadmap planning where you focus on capabilities that have low maturity but high business value.



At this stage we have a lot of the important prerequisite pieces in place. We have prioritized some Knowledge Areas, developed a framework, identified and trained some data stewards and assessed the maturity and importance of different capabilities within the Knowledge Areas.

The next step is to create a short charter and roadmap for each prioritized Knowledge Area. Just like our overall charter (from step 1) these documents should be no more than a couple of pages and the roadmap really doesn't need to be anything more than a set of key requirements that we want to start with.

The artefacts themselves are useful but what's more important is that approval of the artefacts demonstrate a concrete commitment from your stakeholders.

- ✓ The Data Owner must commit to and sponsor the charter and roadmap for their Knowledge Area.
- ✓ The Data Stewards must take ownership for the delivery of the roadmap within their Knowledge Area.

This last point is really important; **"Data Stewards must take ownership for the delivery"**. This comes back to "eating the elephant one bite at a time". You want to scale Data Governance across your business, and you are putting in place the frameworks, training, templates and tools to enable your teams. However, you must make the Data Owners and their delegated Data Stewards own the initiative. If you are trying to own and drive each initiative from a small set of central resources, you will not be able to scale effectively.

Now, this doesn't mean you abscond the Data Governance Team from having responsibility. Rather you should think about this as a program and project model. The Data Governance program is the responsibility of the Data Governance team, but each Knowledge Area should operate as a project driven by the Data Owner and Data Stewards.

7

Develop templates or tools

In the next phase you will start building out your Information Management artefacts (glossaries, data dictionaries, report catalogues and control libraries). To do this effectively you need to have a tool and templates in place. The tool should support the data stewards and make their lives easier. It must also ensure that all the information you put together is searchable and accessible across the business.

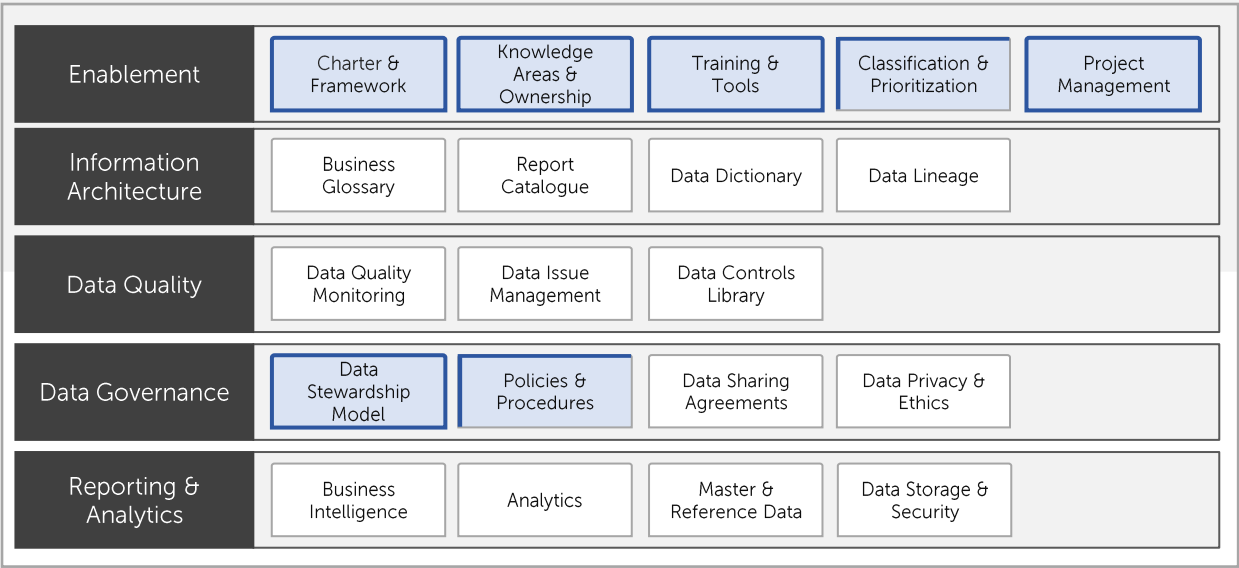
Its possible to get started using Microsoft Excel or other non-specialized tools. However, its really hard to scale and maintain Data Governance without a dedicated tool. Your data stewards will become frustrated, and you will have misalignment, duplication and rework. The cost of an enterprise-wide license for a modern data governance tools like Infornite starts from \$15,000 per year. This is within reach of most budgets and will justify the investment through increased productivity very quickly.

The tool also allows you to publish all the work you have done back out to the business and demonstrate the value that the data governance program is doing.

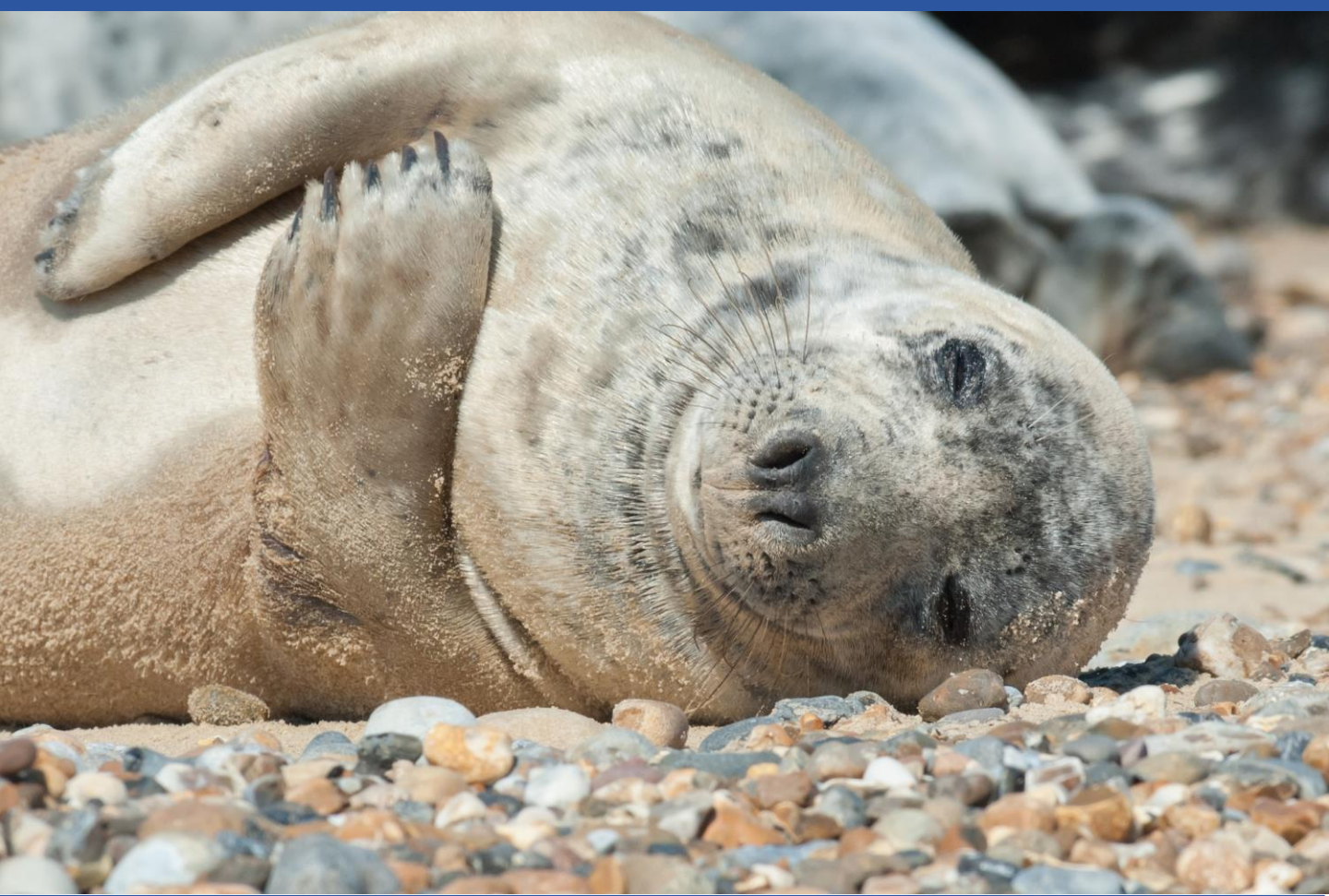
.....

Phase 1: Foundations : Data Governance Components

You have now covered the 7 steps of phase 1. The diagram below illustrates the components of the Infornite Data Governance framework that are fully or partially delivered during this phase.



Infornite: Data Governance Framework: Foundational Components



Foundations Summary

At this point you should be feeling pretty good that your Data Governance implementation is off to a great start. You now have a lot of the pre-requisite capabilities in place including:

- 1 You have a clear charter and committed sponsors and data owners.
- 2 You have scoped your program into manageable Knowledge Areas.
- 3 You have developed and adopted a robust data governance framework.
- 4 You have equipped your stewards with the training they need to succeed.
- 5 You have assessed your current capability to identify gaps and set priorities.
- 6 Your data stewards own a roadmap within each of their Knowledge Areas.
- 7 You have appropriate tools and templates to support the execution.



Information Architecture: Building our blueprint

Now that you have the right foundations in place with a team of data stewards and the right tools; you can move to Phase 2 of the Data Governance implementation. This phase focuses on Information Architecture. This is where you start to build your Business Glossary, Data Catalog and Controls Library.

These artefacts build a blueprint of your data landscape. They will help the entire organisation to become more data driven by enabling people to **find**, **understand** and **access** the data they need. They also provide a mechanism to define and assign ownership of controls across the data lifecycle – this leads to safer management of your data and alignment with regulatory guidelines.

During this phase, the goal is to build initial artefacts for high priority data within a Knowledge Area. You will also put in place the operating rhythm so that the stewardship work can continue after the initial implementation is finished. This phase covers 4 key steps which begins with your Business Glossary and moves through to cover Data Lineage and Controls.

A business glossary is a great place to start your Information Architecture journey. The data stewards in each Knowledge Area create a list of the key business terms, acronyms and metrics used in that area. Its likely that much of this information will already exist in dispersed documentation that your stewards can pull together. Ask your stewards to populate the information directly into your data governance tool.

Bear in mind that you are not looking for perfection in terms of how your stewards write definitions or to have a 100% complete business glossary in place. The goal is to create a decent starting point that you can add to and improve over time. A good way to build momentum is to set your data stewards a challenge each week that gives them something specific to focus on.

- Define the top 10 Key Performance Indicators for their area.
- Describe the 20 most used acronyms in their area.
- List the 25 terms that a new starter would need to know.
- List all the entity types (person or organization) that they deal with.



At this early stage you should take the time to build an effective and collaborative operating rhythm amongst your data stewards. Bring the group together regularly to discuss progress and share challenges. Create a channel on your organisations messaging platform (e.g., MS Teams or Slack). Organize a team event or an informal social catch up via video messaging once a fortnight.

As you start to kick off the execution phase you have a one-time chance to build a great collaborative culture so make sure that you take it.



A business glossary is in place for active Knowledge Areas.

A report catalogue is simply a list of the reports that are produced across the different Knowledge Areas. You must ask your data stewards to bring this information together in your data governance tool. If your organisation is using reporting tools like Power BI or Tableau, then you can quickly harvest information about the reports they contain. For each report ask your stewards to capture:

- A description of the report and the insights it can provide.
- The frequency with which the report is created or updated.
- The type of report (e.g., Excel File, Tableau Dashboard, PDF).
- Description of how someone can get access to the report.

This is a good time to start linking the terms and metrics from your business glossary to the reports where they can be found. In this way you are opening the path for users to search for terms and metrics in your glossary and then follow through to understand how they can access this information in reports.



A report catalogue is in place for active Knowledge Areas.

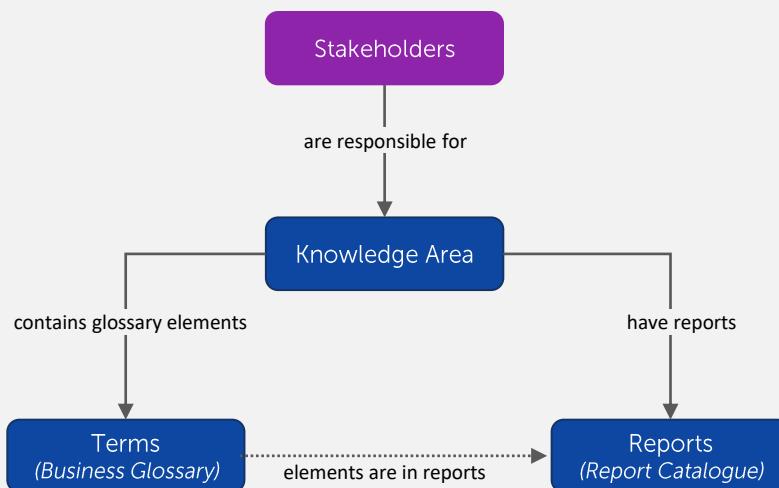


Metamodel

As you build out your data governance artefacts, you may notice an inherent model emerging. The model depicts the things you need to understand as part of your data governance initiative and the connections between them.

At InforNite we call these different things “facets”. So far, we have touched on 4 facets. We discussed *Knowledge Areas* and *Stakeholders* in Phase 1. The third facet are the *Elements* that we will capture in the Business Glossary during step 8. Our fourth facet are the *Reports* covered in step 9 when we build the report catalogue.

These facets and the connections between them are commonly referred to as a Metamodel – it’s a useful structure that helps describe the different facets that you care about and how they relate to each other. Your data governance tool will bring these facets to life by allowing you to connect them and make them searchable across the organization



10 Document our data dictionaries

At this stage, your data stewards are starting to get into a rhythm and your business teams are starting to see the early benefits of data governance via the business glossary and report catalogue. The next step is to start building the technical data dictionaries for key systems that have been prioritized in each Knowledge Area.

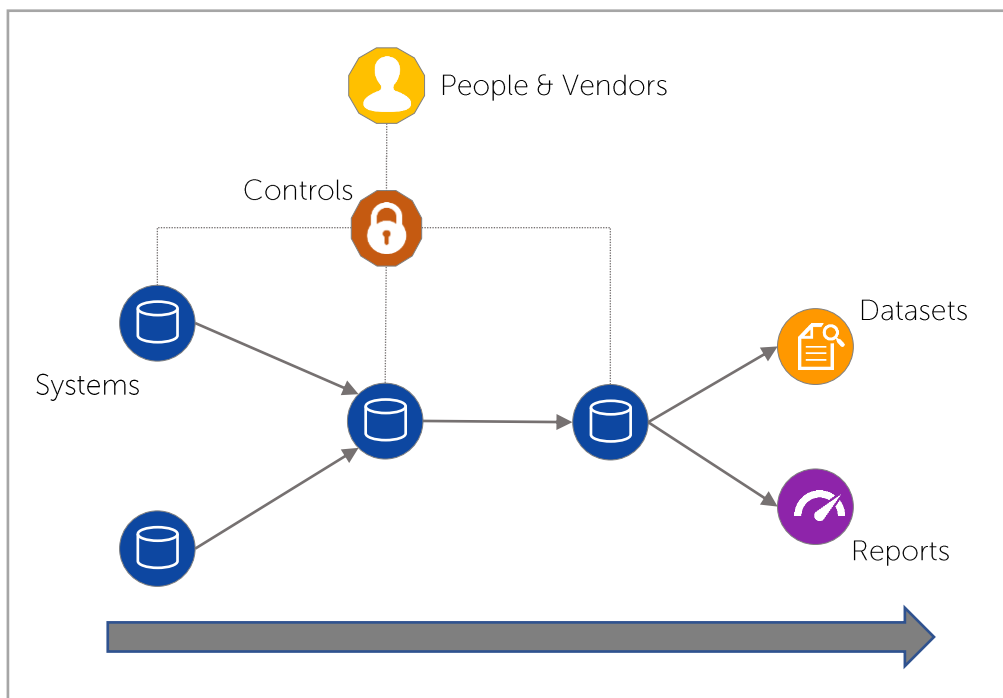
11 Understand where data comes from (Data Lineage & Controls)

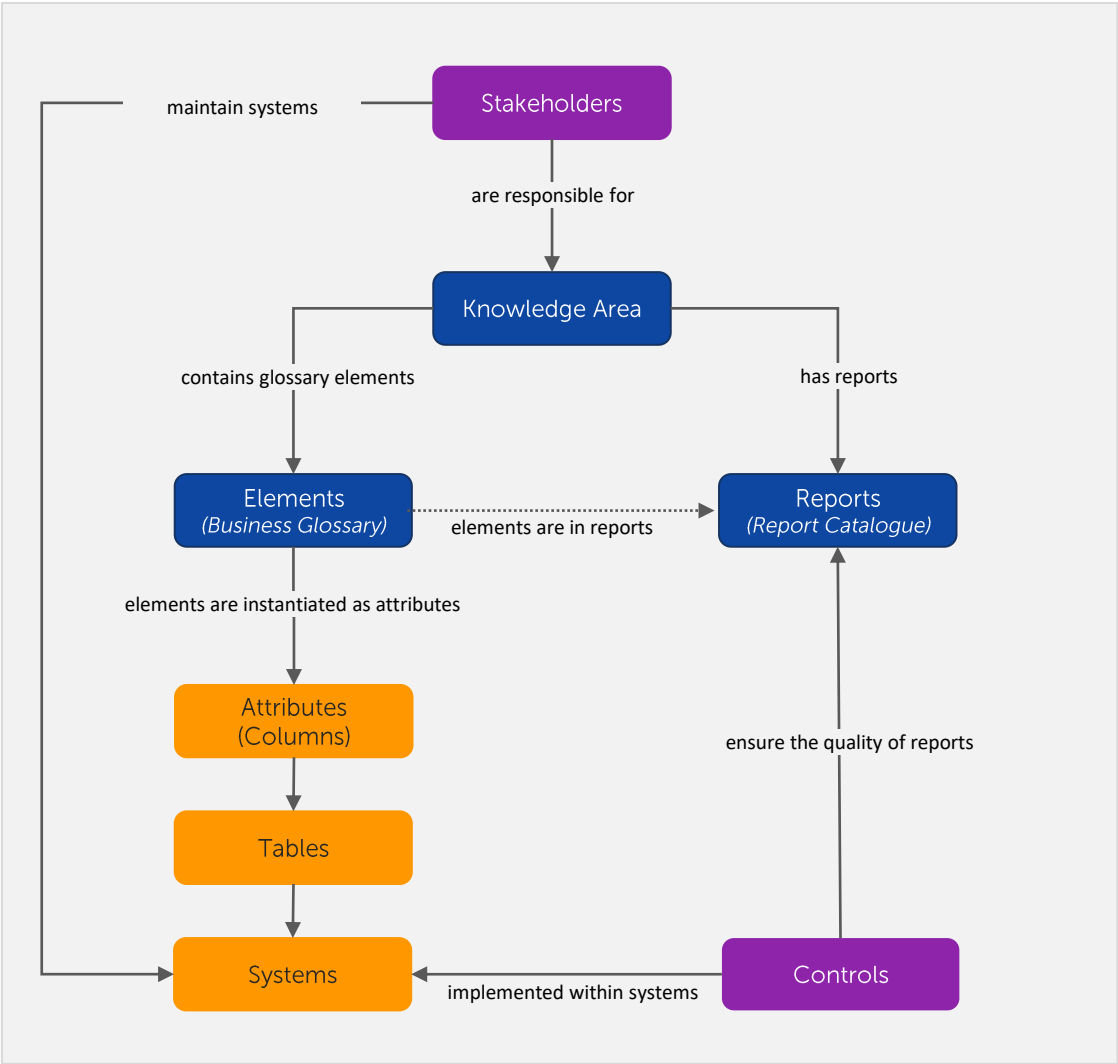
It's useful to document and analyse the data flows from system to system. These data flow diagrams and the transformations that happen along the way are commonly called Data Lineage and can be used for many things:

- Tracing upstream from a report to identify the root cause of a data quality issue so that it can be fixed.
- Understanding the downstream data sources that are impacted when a system or process fails.
- Depicting the different steps that contribute to a final dataset or report so that we can assess if appropriate checks and balances (controls) are applied at each step to give confidence in the report.

Data Lineage comes in many forms; some organisations create detailed technical lineage where every field is mapped through every step of the journey along with the transformation logic that is applied to each step. For most organisations high level lineage which depicts the key systems and data flows will add the most value whilst minimizing the effort to create and maintain the lineage.

During this step your Data Stewards should work with your technology teams to build a high-level view of the systems and data sources. Each steward will use these diagrams to create data flows for the important reports and datasets in their area. These flows should be enriched with information about the controls that are applied at each step (e.g. reconciliations, data quality checks) and the people that are responsible for managing the systems at each step.







Information Architecture Summary

The Information Architecture phase of the Data Governance implementation is an interesting and challenging one. Its interesting because you are getting your hands dirty and producing meaningful artefacts that can really help your organisation. Its challenging because there is a learning curve for your data stewards, and you need to make be there to support and guide them on the journey.

A key differentiator between this and the foundations phase is that the Information Management phase is not really a phase at all. Its an ongoing process that should transition to a set of BAU stewardship activities. As part of the implementation project the goal is not to document everything but rather to put the processes in place. Help your stewards build the initial set of artefacts and create an approach that can be repeated.

At the end of this phase, you should have:

- 8 A single searchable enterprise business glossary.
- 9 A searchable catalogue of key reports that everyone can access.
- 10 A data dictionary for key assets that can be enriched over time.
- 11 Defined data flows for key reports and datasets along with a methodology to document controls and the people that are responsible for managing them.



Guardrails: How do we move faster safely




At this point you have the foundations in place and your stewardship model is operating. You have also started creating glossaries and dictionaries that are being used by your business to help them **find**, **understand** and **access** data. You are now ready to move to Phase 3 of your Data Governance implementation. This phase is about formalising the policies and standards that your organisation will follow.

In traditional approaches to Data Governance this phase often comes first. The initiative will often start by drafting policies and standards. These are usually grounded in best practices but rarely reflect reality. This is because the team has not yet gotten their hands dirty and don't really understand the scope. Generally, it's hard to create something practical and implementable until you've looked under the covers a little bit.

In this implementation approach we build out the skeleton of the policy and other artefacts as we work through Phases 1 and 2 with a focus on practical implementation. We then formalize these policies and standards during phase 3. However, this approach only works as part of a timeboxed initiative – we don't recommend that running your governance processes indefinitely without creating the formal policies and standards.

Data should be classified so that you can understand its sensitivity and make informed decisions about who can access it. During this step you define a Data Classification Policy. This classification standard will apply to all your data. It will be referenced by your Data Stewards as they enrich your data dictionaries. There are existing standards (including ISO 27001) that you can leverage to put this together.

In addition to the Data Classification Policy, you should also develop Data Access Guidelines during this step. Typically, these guidelines outline the request and approval process for a given piece of data considering the sensitivity of its data classification. A simple approach is to say that all sensitive data requires the approval of the Data Owner (that we identified back in step 2). The Data Access Guidelines should also specify the frequency with which access reviews are required.

-  A documented Data Classification Standard that has been approved by the Data Owners.
-  A documented Data Access Guideline exists and each System Owner is aware of them.
-  You have developed and published training artefacts on the Data Classification Standard and Data Access Guidelines.

In this step you create your Data Governance Policy and guidelines. The policy is essentially the rulebook for how you treat data in your organisation. These are the non-negotiable rules that everyone must comply with. A Data Governance Policy is a powerful artefact when crafted as a set of clear requirement statements. The Data Governance Policy can serve as a guardrail that allows the organisation to be more agile and move faster because there are fewer gray areas and guess-work. It brings clarity to people and projects so that they can operate more effectively.

In addition to the policy you should craft and publish guidelines about how you want people to manage and treat data. Unlike the data policy these guidelines are not hard rules. They will often reflect your aspirations but not necessarily our current maturity. However, they serve an important purpose which is to ensure we don't pollute our policy (the must) with rules that are impractical at this stage (the should).



A clear Data Governance policy exists that is enforceable.



Published data governance guidelines exist.

14 Project Review & Impact Assessments

Your data stewards are (hopefully) working hard to build a metadata blueprint of your data landscape. Now is a good time to put in place formalized reviews of project and change initiatives to ensure that the work they do is aligning to what you stewards have built and the policies you have put in place. To ensure this alignment you should develop a formal project review process along with a checklist of items to review and a set of artefacts that the project is expected to put in place.

A key objective is to ensure that every project contributes to and maintains the Information Architecture work that our data stewards are building. If a project changes a database, then they must update the data catalogue; if a project creates a new report, then it must be reflected in the report catalogue.



Data Impact Assessments are performed for every project.



Every project updates the Information Management blueprint.

15 Data Sharing Agreements & Vendor Management

Data Sharing Agreements are a contract between 2 parties (data producer and data consumer) that describes how they work together on the exchange of data. This includes the data requirement and expectations about service level agreements and data quality. The Data Sharing Agreement also describes what the data can and can't be used for.

Data Sharing Agreements can be a very useful tool between different internal teams and systems. However, they really come into their own to help govern data sharing with external vendors. During this step we should create a standard Data Sharing Agreement template. The template must align with our Data Governance Policies.

The template can also serve as a great baseline against which to review and evaluate existing agreements in the form of contracts that we have in place with vendors.



We have a Data Sharing Agreement template

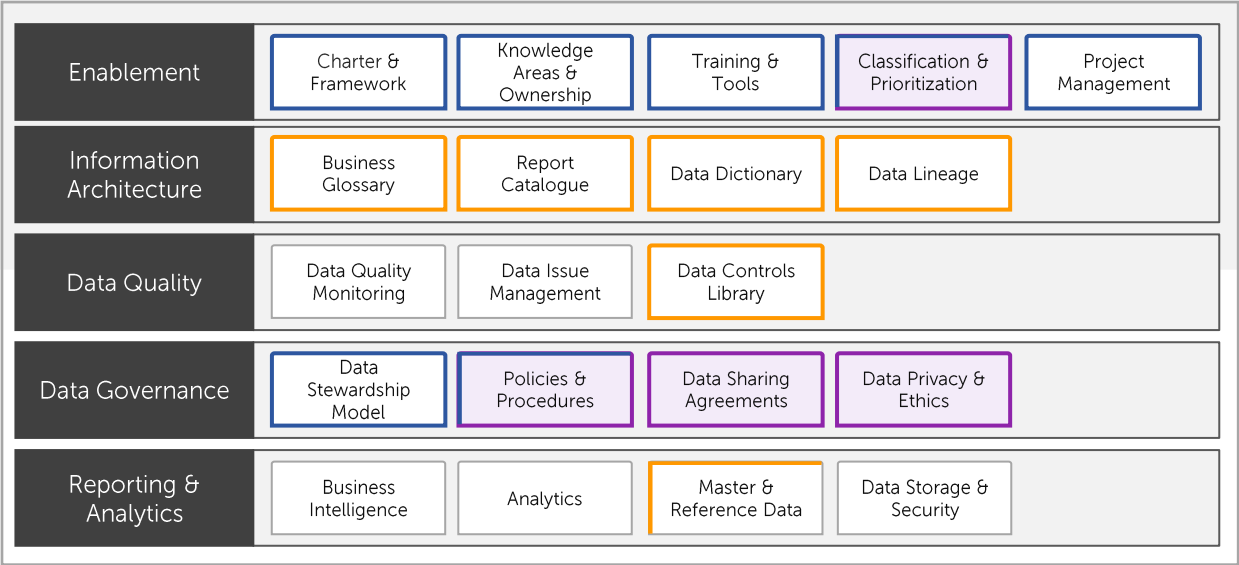
The final step in this phase is to set up a formalized Data Governance Committee. This will often be made up of the Data Owners and senior leaders within the organisation. The role of this group is to approve standards, set priority and drive the strategic direction of the data governance program. A quarterly or bi-monthly forum is appropriate.

Most traditional data governance implementations will set up a Data Governance Committee very early on and often as the first step. However, these committees usually have nothing to govern and are a waste of everyone's time. You need to have policies in place and data stewards creating output so that you have a series of activities to govern and decisions that need to be made. You need to work closely with your data owners throughout the implementation, but the Data Governance Committee is something that becomes valuable as you transition from the initial implementation to a Business-as-Usual model.

.....

Phase 3: Guardrails: Data Governance Components

You have now covered the 5 steps of phase 3. The diagram below illustrates the components of the InforNite Data Governance framework that are fully or partially delivered during this phase.



InforNite: Data Governance Framework: Guardrail Components

Guardrails Summary

During this final phase of the Data Governance initiative, we focused on wrapping formality and procedures around the good work that was developed in Phase 1 and 2. In the end of the day Data Governance should bring control and security to our data landscape and often we need to evidence these capabilities to our partners and regulator. To achieve this and create a sustainable capability we need to have formalized and enforceable policies and procedures. The policy and procedures provide the Guardrails that allow your teams to move faster with fewer grey areas.

At the end of this final phase, we have:

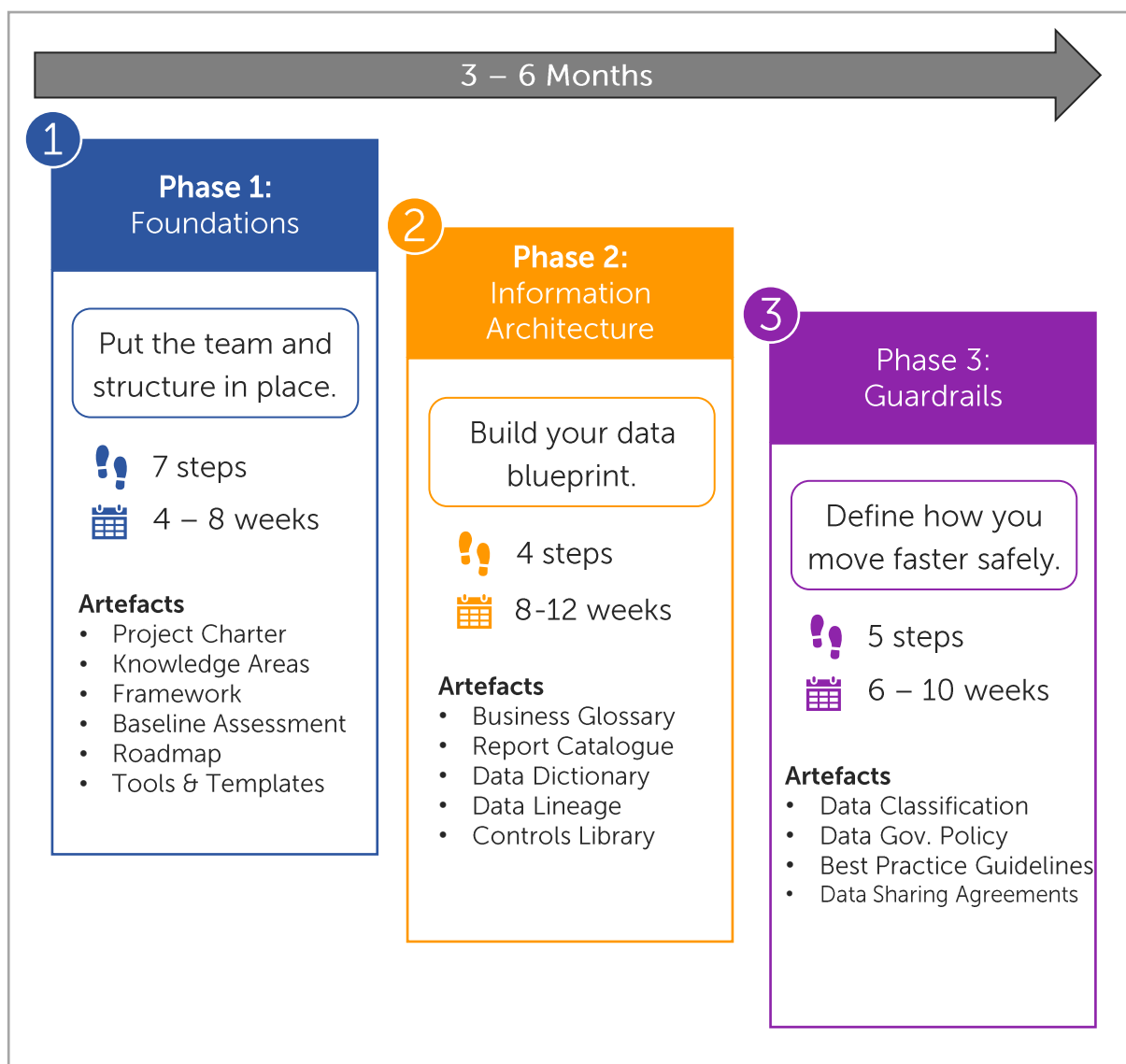
- 12 An approved data classification standard and data access process.
- 13 Governance Policy and Guidelines that focus on enabling teams.
- 14 Projects are engaged and contribute to the information architecture.
- 15 Data Sharing Agreements are developed and in place with key vendors.
- 16 Formalized Data Governance Committee is in place.



Closing Remarks

Data Governance has the power to be a transformational capability for your organisation. It's the keys to unlocking the potential of your data and truly empowering your organisation to make great data driven decisions. It's also a hugely fulfilling discipline to work in when you focus on practical steps that deliver tangible outcomes and solve real problems for your organisation.

However, data governance is broad undertaking and its often left up to a **lone wolf** to drive data governance uplift across the organisation. With a small resource footprint its easy to become overwhelmed so you need to follow a pragmatic approach that allows you "to eat the elephant one step at a time" and demonstrate progress after every bite.



Infornite: Phases & Timeline of a Data Governance Implementation

Contact

Michael O'Sullivan
InforNite | Founder & Director

✉ michael@inforNite.com

☎ + 61 415 652 507

🎥 Book a discovery meeting with us at inforNite.com/meet

🌐 Join our data community at inforNite.com/community