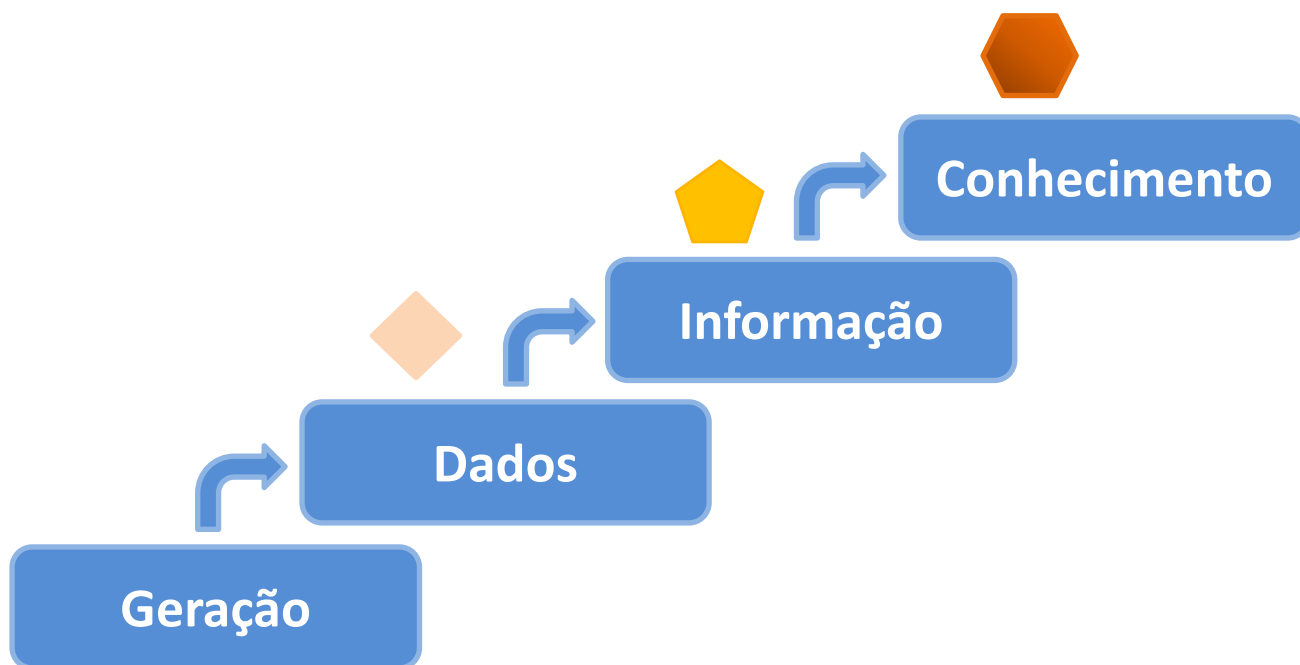




# GOVERNANÇA DE DADOS

## Evolução dos dados da geração ao conhecimento



Cada etapa ao lado gera uma transformação.

A maioria das instituições chega ao degrau de informações.

Quando se avança um pouco mais, aumentando a qualidade dos dados e tratando-os como um dos principais ativos da empresa, tem-se o que chamamos de geração de conhecimento.

Neste patamar, toda e qualquer informação é gerenciada com relevância e entrega valor de forma plena.

# Maturidade Governança de Dados

O sistema DAMA DMBok2 reúne as melhores práticas de governança de dados a partir de pesquisas com mais de 100 empresas e divide o processo em 10 verticais interligadas, conforme se pode ver na imagem ao lado.

Entender o nível de maturidade em governança no qual se encontra uma empresa é o primeiro passo para se traçar um plano de implementação deste processo.

Para isso será feito um diagnóstico da CCR, buscando um melhor entendimento de seu nível atual em relação as melhores práticas de governança, seguindo a metodologia DAMA DMBok.

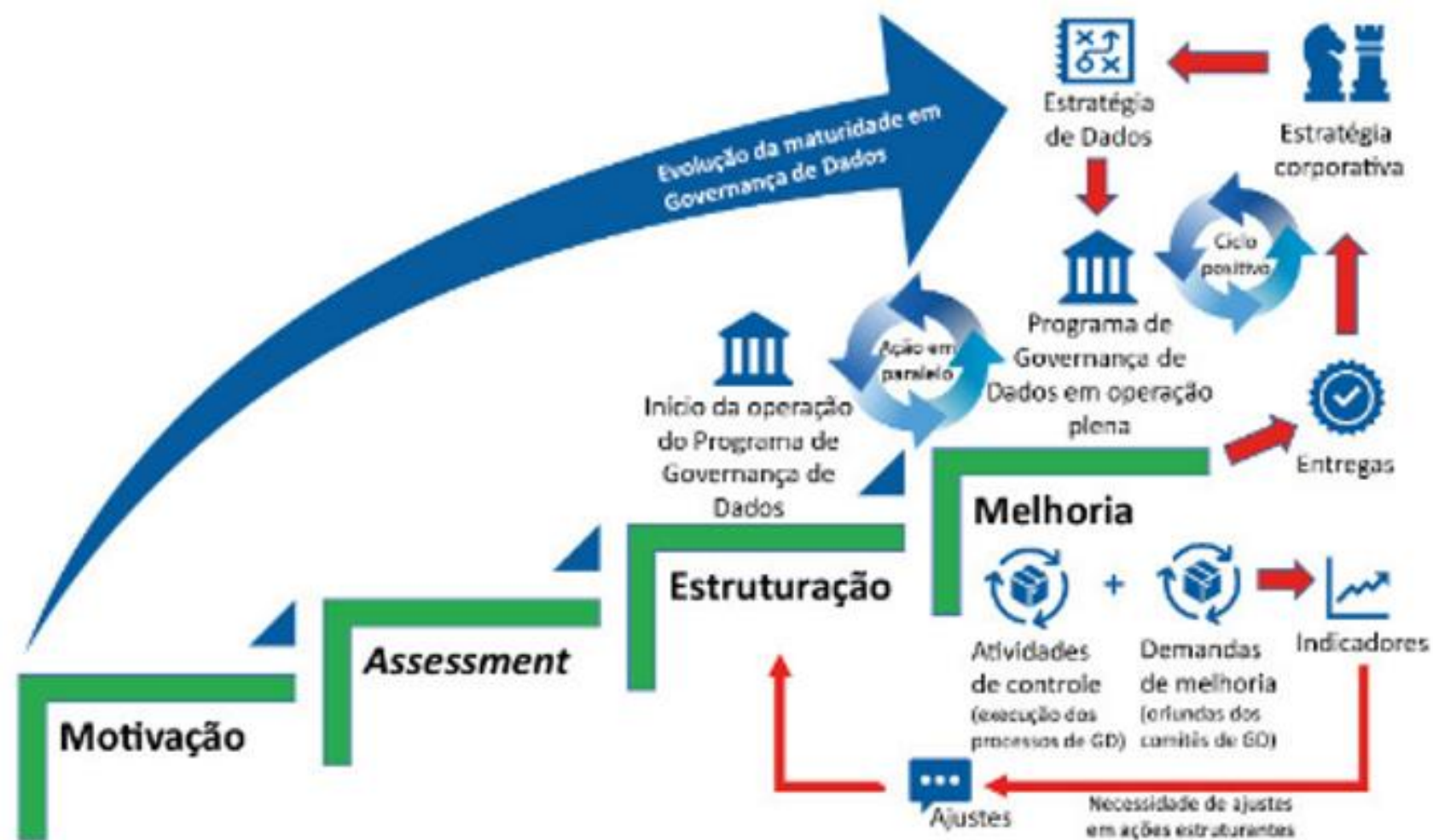


**DAMA-DMBOK2 Data Management Framework**

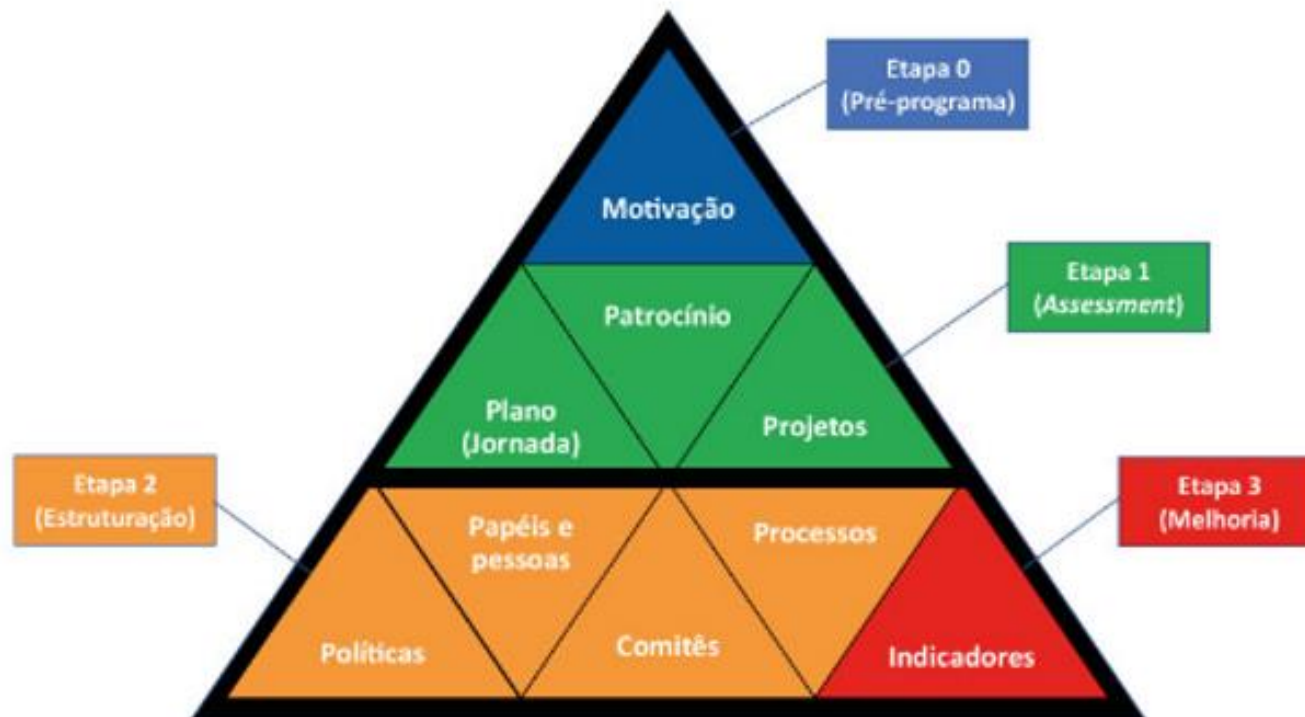
Copyright © 2017 by DAMA International



# Etapas da Governança De Dados



# Componentes básicos de um programa de Governança de Dados



## 0.1-Motivação:

Quais são os ganhos obtidos com o programa? Sem elas o assunto ou programa será taxado como "modismo".

**1.1-Patrocínio:** Quem apoia o projeto? VP, diretoria, gestores?

**1.2-Projeto:** DataPlat por ser estruturante

**1.3-Plano (Jornada):** Roadmap e Framework para implantação

**2.1-Políticas:** Diretrizes que devem ser aplicadas na gestão de dados

**2.2-Papéis e pessoas:** Arquitetos, Data Steward, Data Owner e quais pessoas a ocuparem esses papéis.

**2.3-Comitês:** São estruturas de apoio formado por profissionais de Governança de Dados, Negócio e Tecnologia. Esse grupo se reúne de forma periódica para resolução de conflitos e indicadores para melhoria atuando no operacional, tático e estratégico.

**2.4-Processos:** Os processos padronizam as formas de atuação nas atividades Gov de Dados. Quem está fazendo o que e para um determinado objetivo esperado.

**3.1-Indicadores de Melhoria Contínua:** Instrumentos de qualidade que permitem avaliar o desempenho do programa. São obtidos de forma periódica, com base em informações e análise sobre os ativos de dados.

Onde estamos?

Assessment

# Processo de diagnóstico

Questionário com 113 itens relacionados às 10 dimensões de um bom processo de governança de dados.

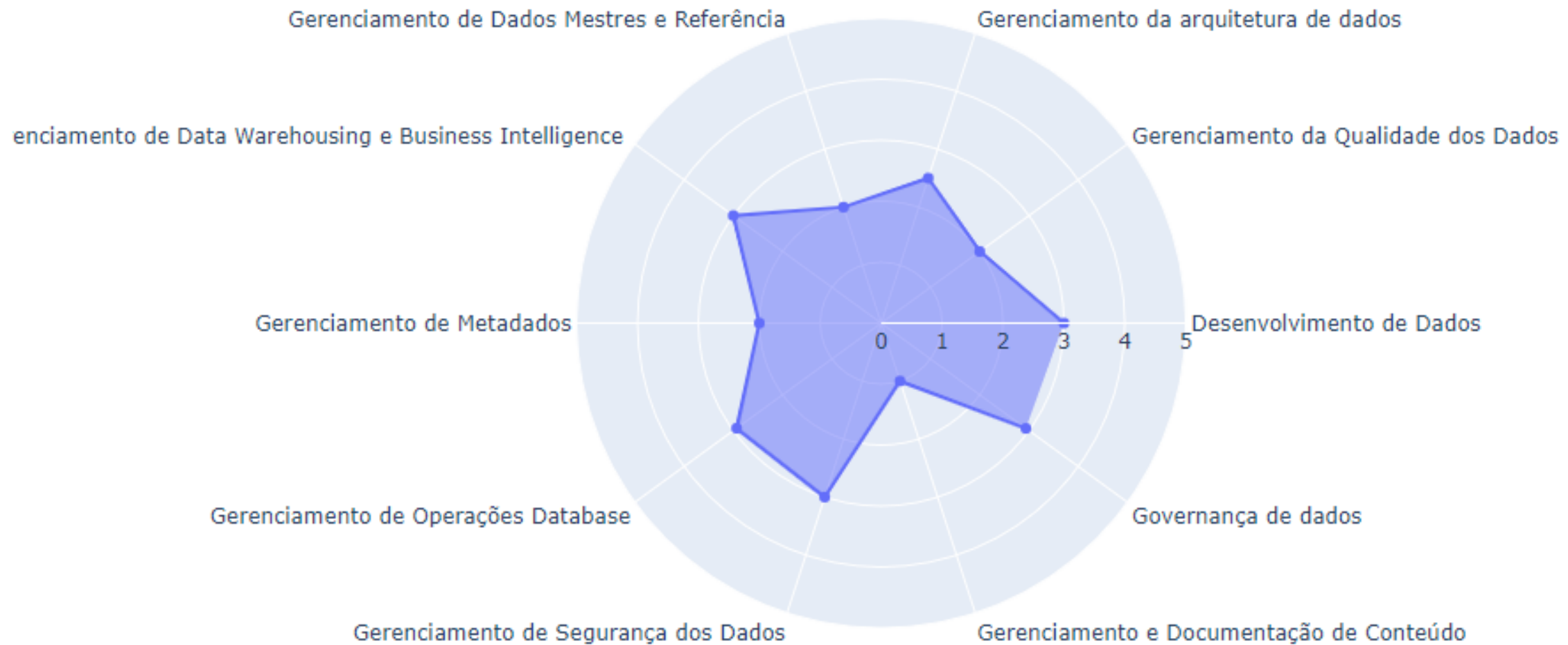
Ao lado observa-se parte do questionário a ser utilizado.

O questionário deverá respondido com notas de 1 a 5 e os resultados são exemplificados a seguir.

1	Grupos	No	Atividade	Nota
2	Governança de dados	1a	Entender a necessidade estratégica dos dados. Estas atividades estão relacionadas a como garantir a qualidade dos dados, quais áreas serão envolvidas no processo, qual o melhor caminho a se seguir, como obter os dados, qual a área de negócio que vai se	2
3	Governança de dados	1b	Desenvolver e manter a estratégia dos dados. Iniciar as ações para obtenção dos dados e aplicação das estratégias definidas na atividade anterior.	2
4	Governança de dados	1c	Estabelecer unidades organizacionais e papéis para as atividades de dados. São os responsáveis pelos dados na organização, são eles, <i>data stewards</i> , DBA, gerente de dados, comitê de governança de dados.	4
5	Governança de dados	1d	Identificar e nomear os <i>Data Stewards</i> . São os responsáveis pelo controle e uso dos dados, estes devem possuir um amplo domínio e conhecimento	4
6	Governança de dados	1e	Estabelecer estruturas organizacionais para governança de dados e para os <i>data stewards</i> . Visão detalhada sobre os responsáveis sobre a gestão dos dados, por exemplo, membros do comitê de governança de dados, principais <i>data stewards</i> , principais DBA,	2
7	Governança de dados	1f	Desenvolver e aprovar políticas, padrões e procedimentos de gestão e governança de dados. Detalhamento dos 3 Ps da governança de dados: Políticas, que são as regras definidas e aprovadas pelos comitês de gestão; Padrões, que são os descritivos dos nomes, documentos, layouts, entre outros; Procedimentos, passo-a-passo de como	2
			Revisar e aprovar a arquitetura de dados.	

# Diagnóstico Governança

## Exemplo de resultados

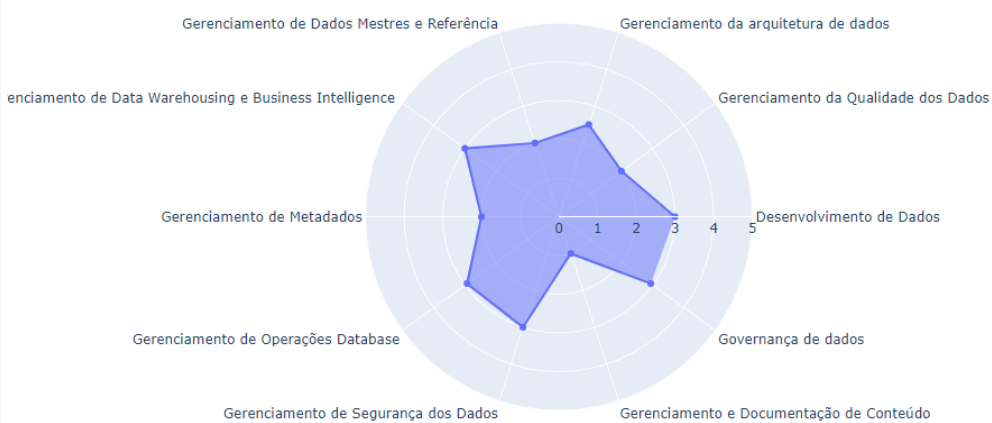




# Diagnóstico Governança

## resultados

	Nota	std
Grupos		
Desenvolvimento de Dados	3.000000	0.000000
Gerenciamento da Qualidade dos Dados	2.000000	0.000000
Gerenciamento da arquitetura de dados	2.500000	0.534522
Gerenciamento de Dados Mestres e Referência	2.000000	0.000000
Gerenciamento de Data Warehousing e Business Intelligence	3.000000	0.000000
Gerenciamento de Metadados	2.000000	0.000000
Gerenciamento de Operações Database	2.933333	0.258199
Gerenciamento de Segurança dos Dados	3.000000	0.000000
Gerenciamento e Documentação de Conteúdo	1.000000	0.000000
Governança de dados	2.937500	0.250000



### Insights

- Gerenciamento e Documentação de Conteúdo apresentam maior fragilidade.

### Legenda

- 0 -> Não sei do que isso se trata
- 1 -> Não temos consciência da importância desta atividade e a mesma não foi iniciada
- 2 -> Temos consciência da importância desta atividade mas a mesma não foi iniciada
- 3 -> Atividade já iniciada mas foi interrompida
- 4 -> Atividade iniciada e em execução
- 5 -> Atividade finalizada

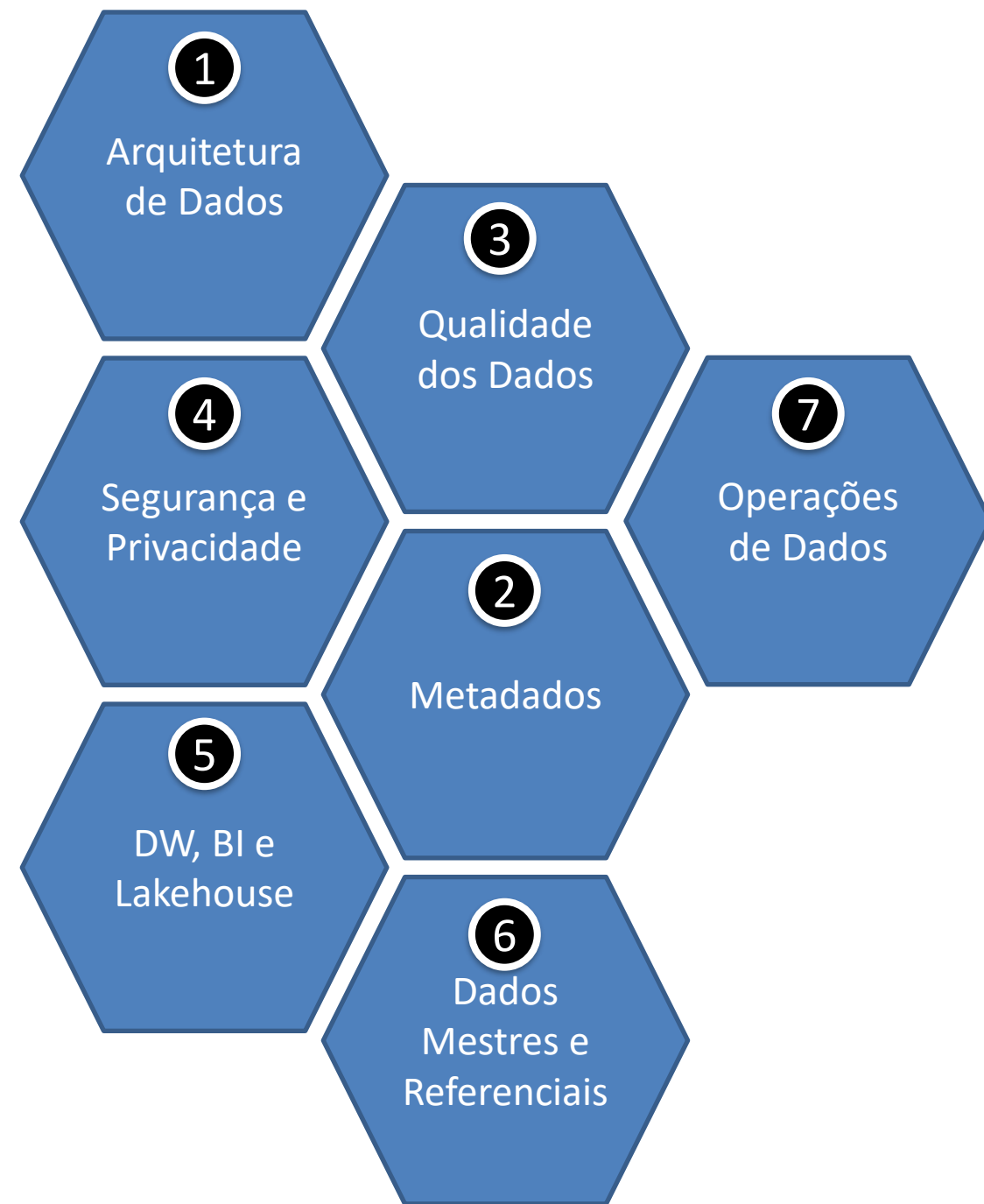
**Para onde Iremos?**

**Frameworking  
e Cronograma**

# Governança de Dados



Fonte: Dama (2012)

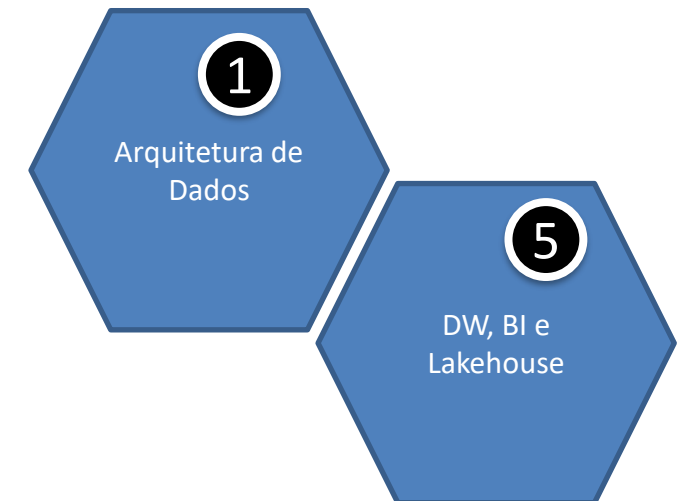


# Cronograma Governança

## Arquitetura de Dados

Trata-se de entender quais os requisitos do projeto de dados, ou seja, entender quais os dados que são necessários, de onde eles vêm e por onde terão de passar até chegar ao dashboard. É definir o caminho deste dado.

A Arquitetura de dados é um conjunto integrado de artefatos de especificação utilizada para definir os requisitos de dados, orientar a integração e controle dos ativos de dados e alinhar os investimentos de dados com a estratégia de negócios. É também um conjunto integrado de diagramas em diferentes níveis de abstração. A arquitetura de dados inclui nomes formais de dados, definições de dados abrangentes, estruturas de dados eficazes, regras precisas de integridade de dados e documentação de dados robusta (DMBOK, 2012).



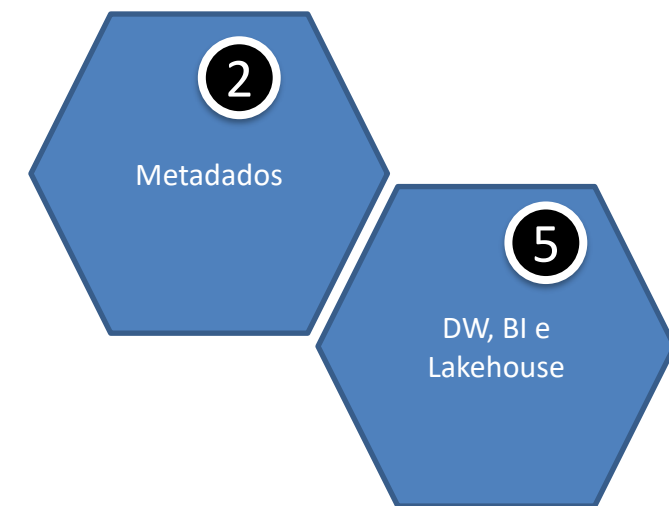
# Cronograma Governança

## Metadado

São informações que complementam um dado original. Por exemplo: Uma tabela de aluno (dado), informações sobre os dados contidos na tabela (Metadado). Portanto, o metadado tem a mesma importância do dado e o gerenciamento de metadados segue as mesmas atividades do gerenciamento de dados.

Os meta-dados descrevem a estrutura e significados a respeito de dados e, assim contribuem para que seu uso seja eficiente ou ineficiente, oferecendo contexto aos dados relacionados, ou seja informações que gerem conhecimento (TURBAN; ET. ALL, 2009).

**Comentário:** O *Purview* é a ferramenta que está facilitando o processo de mapeamento dos dados. Ela realiza um scan em diversas fontes catalogando colunas e seus tipos de dados de forma automática. Inicialmente, está sendo mapeado o datalake, alvo do projeto no que se refere a governança, mas poderá ser estendido para as demais fontes.





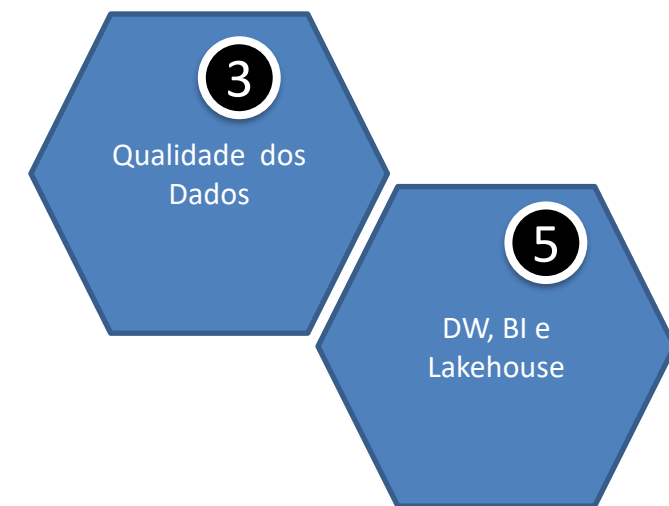
# Cronograma Governança

## Qualidade dos Dados

Planejar e projetar o saneamento do dado, provendo qualidade para que este possa gerar informações confiáveis para suportar a tomada de decisão. O objetivo é planejar, implementar e controlar atividades que apliquem técnicas de gerência de qualidade de dados para medir, avaliar, melhorar e garantir a adequação dos dados ao seu uso pretendido.

O Gerenciamento da Qualidade dos Dados inclui o planejamento e a implementação de técnicas de gerenciamento da qualidade para medir, avaliar e melhorar a adequação dos dados para uso em uma organização.

**Comentário:** A forma de como os dados estão sendo armazenadas no Datalake (Bronze, Silver e Gold) incluindo o Synapse contribui para Gerenciamento da Qualidade dos Dados. Um bom exemplo é a disposição e agrupamento por domínio, seja por sistema como na Landing ou por entidade na Silver, além de enriquecimento e regras de negócio nas demais áreas. As regras serão provadas para garantir que as informações oriundas dos dados qualificados sejam confiáveis.

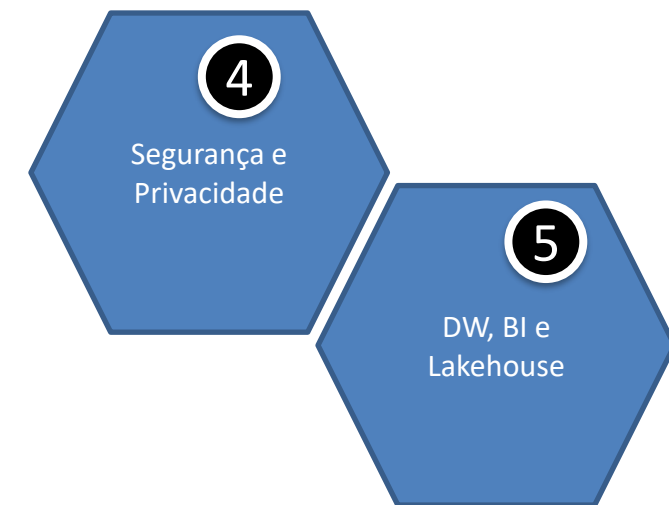


# Cronograma Governança

## Segurança e Privacidade

Consiste no planejamento, desenvolvimento e execução de políticas e procedimentos para assegurar a devida autenticação, autorização, acesso e auditoria nos ativos de dados e informações. (DMBOK, 2012).

**Comentário:** O *Purview* é a ferramenta que está facilitando o processo de mapeamento dos dados com algumas sugestões automatizadas de dados sensíveis que precisarão ser ofuscados ou mascarados(Ex.: Colunas com o nome CPF e RG) conforme política de acesso. O catálogo enriquecido manualmente permitirá saber quem é o dono do dado e seus dados de contato. Um outro documento permitirá saber os papéis e pessoas que terão acesso aos dados disponibilizados no datalake e nas camadas de consumo como o Synapse e Power BI.



# Cronograma Governança

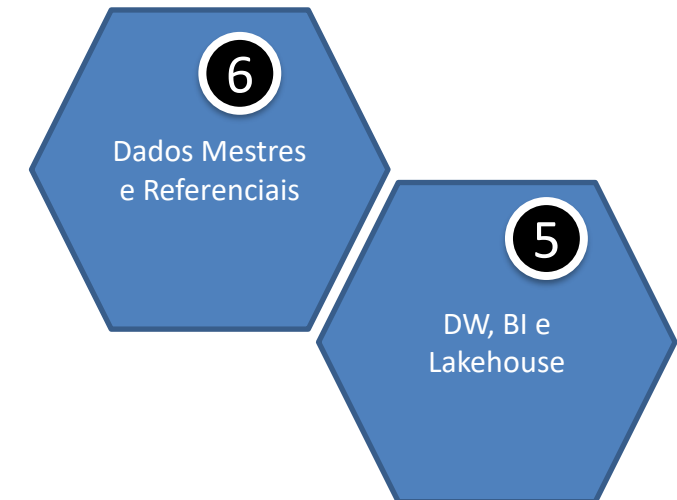
## Dados Mestres

O Gerenciamento de Referência e Dados Mestre inclui reconciliação e manutenção contínuas dos principais dados compartilhados críticos para permitir o uso consistente entre os sistemas da versão mais precisa, oportuna e relevante da verdade sobre entidades essenciais de negócios .

Os dados mestres constituem os principais dados da empresa. São dados sobre entidades de negócios vitais que dão contexto às transações. Esses dados de acordo com DATA Governance Glossary “descrevem as entidades centrais de uma empresa e são utilizados por vários sistemas de processos de negócios e de TI. Exemplos desses dados são as partes, por exemplo, (clientes, funcionários, fornecedores, parceiros), lugares, por exemplo, (locais, territórios de vendas, escritórios), e as coisas (contas, produtos, bens, conjuntos de documentos) ”.

A gestão de dados mestres é o controle sobre os valores de dados mestres para viabilizar o uso contextual, consistentes, compartilhados entre os sistemas, de forma acurada, temporal e relevante para as entidades essenciais do negócio. (DMBOK, 2012).

**Comentário:** Os dados alvos e estratégicos para o negócio são consumidos diretamente da origem. Dessa forma, compartilha-se os dados de forma confiável e segura nas camadas de consumo gerenciadas(Ex.: Gold , Synapse e PowerBI). Outro ponto importante é a centralização dos dados no datalake.



# Cronograma Governança

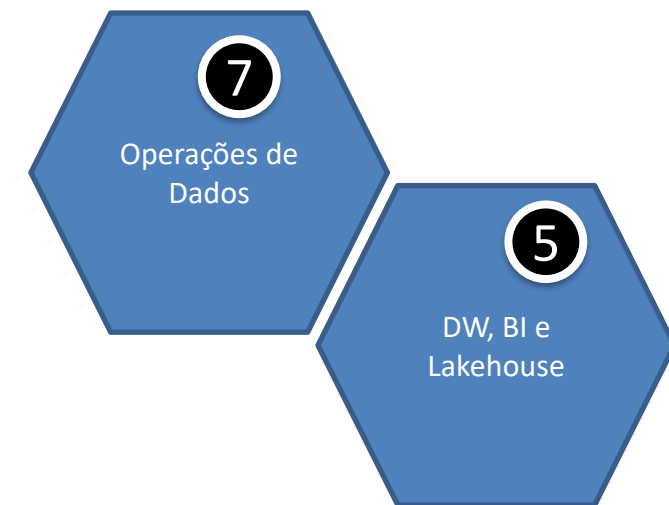
## Operações de Dados

**Operações de Dados:** Trata-se do planejamento, controle, manutenção e suporte ao ativo dado, durante todo o seu ciclo de vida, ou seja, desde sua aquisição, passando por sua exibição, até a eliminação desse dado. Vale ressaltar que um planejamento de Recuperação de desastre é de extrema importância para qualquer projeto.

No ciclo da sua vida, o dado pode ser extraído, exportado, importado, migrado, validado, editado, atualizado, limpo, transformado, convertido, integrado, segregado, agregado, referenciado, revisado, relatado, analisado, garimpado, salvo, recuperado, arquivado e restaurado antes de eventualmente ser eliminado (DMBOK, 2012).

Para o DMBOK (2012) o ciclo de vida dos dados está diretamente ligado ao ciclo de vida dos sistemas, como observa-se na adaptação de Rêgo (2013). .

**Comentário:** Um desenho de fluxo de dados será criado no Bizagi com os papés envolvidos da área de Governança.



# Exemplo de Cronograma Governança

DW, BI e Lakehouse	DW, BI e Lakehouse	DW, BI e Lakehouse	DW, BI e Lakehouse	DW, BI e Lakehouse	DW, BI e Lakehouse	DW, BI e Lakehouse
Arquitetura de Dados	Arquitetura de Dados	Arquitetura de Dados	Arquitetura de Dados	Arquitetura de Dados	Arquitetura de Dados	Arquitetura de Dados
	* Definição de ferramenta de linhagem	* Definição de ferramentas para consumo (Ex.:Sql Database para SandBox e outros)	* Linhagem dos dados com Purview (Databricks)	* Definição dos padrões de dados para dev, HML e Prod	* Melhoria contínua	* Melhoria contínua
	Meta Dados - Data Mapping	Meta Dados - Data Mapping	Meta Dados - Data Mapping	Meta Dados - Data Mapping	Meta Dados - Data Mapping	Meta Dados - Data Mapping
		* Acompanhamento DataD5:D6 Mapping	* Acompanhamento Data Mapping	* Acompanhamento Data Mapping	* Acompanhamento Data Mapping	* Acompanhamento Data Mapping
		* Scan em Synapse	* Reexecução de Scan	Scan em PowerBI	* Scan em outras fontes externas	* Reexecução de Scan
	* Setup e treinamento Purview	* Purview - Definição de estrutura do Glossário (Padrão nomemclatura)	* Implantação Glossário	* Treinamento de Uso do Glossário	* Validação de uso do Glossário	
	Seg. Privacidade - Mascaramento	Seg. Privacidade - Mascaramento	Seg. Privacidade - Mascaramento	Seg. Privacidade - Mascaramento	Seg. Privacidade - Mascaramento	Seg. Privacidade - Mascaramento
	* Discussão inicial sobre Mask	* Definição anonimização	* Validação anonimização	* Apoio à anonimização	* Apoio à anonimização	* Apoio à anonimização
		* Definição de acesso x roles x ferramentas	* Definição de acesso x roles x ferramentas	* Implantação de acesso x roles x ferramentas	* Validação de acesso x roles x ferramentas	* Validação de acesso x roles x ferramentas
			Qualidade dos dados * Melhoria Contínua	Qualidade dos dados * Melhoria Contínua	Qualidade dos dados * Melhoria Contínua	Qualidade dos dados * Melhoria Contínua
	Operação dos Dados - Comunidade e Governança	Operação dos Dados - Comunidade e Governança	Operação dos Dados - Comunidade e Governança	Operação dos Dados - Comunidade e Governança	Operação dos Dados - Comunidade e Governança	Operação dos Dados - Comunidade e Governança
	* Definição EGD - Time	* Definição do fluxo de trabalho EGD	* implantação EGD	* Evolução EGD	* Evolução EGD	
	Dados Mestres	Dados Mestres	Dados Mestres	Dados Mestres	Dados Mestres	Dados Mestres
	* Definição de Responsável	* Revisão de fonte de dados	* Melhoria contínua	* Melhoria contínua	* Melhoria contínua	



# LGPD

# Cronograma Governança - LGPD

## Consequência da má Governança

### Dados do início de 2021

Litigiosidade na atual circunstância

...



Mesmo sem a entrada em vigor da LGPD, temos mais de 50 ações que tratam sobre a má governança de dados.



Houve decisão do STJ condenando a parte reclamada a pagar indenização por DANOS MORAIS IN RE IPSA.



Há tese no Ministério Público no sentido que a má governança de dados pessoais pode gerar Improbidade Administrativa.

# Cronograma Governança - LGPD

## Consequência da má Governança

### LGPD

Sanções: Agosto de 2021

\*\*\*

- Advertência, com prazo para medidas corretivas.
- Multa simples: 2% sobre faturamento (excluídos tributos) ou R\$ 50 milhões de reais por infração.
- Multa diária.
- Publicização da infração.
- Bloqueio de dados pessoais.
- Eliminação de dados pessoais.
- Suspensão parcial do funcionamento do banco de dados.
- Suspensão do exercício da atividade de tratamento dos dados pessoais.
- Proibição parcial ou total do exercício de atividades relacionadas a tratamento de dados.



## Cronograma Governança - LGPD

### Consequência da má Governança

Na Europa, já vigora o regulamento de Dados desde 25/05/2018. Desde então:

- Foram protocolados mais de **114 mil** reclamações individuais
- As multas somam mais de **R\$ 1,5 bilhão**;
- Foram notificados mais de **89 mil** incidentes
- As multas aplicadas incluem vazamento de dados, falhas de adoção de medidas seguras de
- Prevenção de dados, acesso indevido a dados sensíveis, venda não autorizada para terceiros
- Mais de 80% das penalidades referem-se a má gestão dos Dados e não a vazamento deles

## Cronograma Governança - LGPD



Fonte: Dama (2012)

O desenho representa os 5 pilares do DAMA que convergem diretamente com a adequação a Lei Geral de Proteção de Dados (**LGPD**). A seguir 5 passos que precisam ser tratados com time de segurança, jurídico e outras áreas pertinentes.



## 1 – Data Mapping

**Onde** – Onde estão sendo utilizado ou transferidos?

**Quem** – Quem responde por isso?

**Como?** Como o tratamento dos dados está sendo feito

**Por que?** Com que objetivos e bases Legais

Áreas ou processos de negócio



## Governança Dados, TI (Segurança e Outros) e Negócio



### 2 – Bases de uso legais

- Consentimento por documentos legíveis e claros, porém pode ser revogado (Mais forte do ponto de vista legal)
- Legítimo interesse do Controlador ou terceiro



## Governança Dados, Jurídico e Negócio



### 3 – Legítimo interesse

- Buscar o equilíbrio de comerciais e as atividades da CCR com relação ao dado trabalhado
  - Elencar esses interesses.
- Teste de proporcionalidade envolvendo Objetos e Necessidade
- Propósito x Proporção



## Governança de Dados, Jurídico e Negócio



### 4 – Descoberta física dos dados

- Complementação do passo 1
- Confirmação dos dados e metadados nas fontes sugeridas
  - Elencar esses interesses.
- Purview auxiliará no processo de auto Data Discovery
- Discovery, Metadados, Qualidade, Catálogo dos Dados



**Governança Dados, TI (Segurança e Outros) e  
Negócio**



### 5 – Avaliação de Impacto e Privacidade dos Dados DPIA (Data Privacy Impact Assessment)

- Avaliar impactos de novas iniciativas: Novo projeto, produto ou demanda (Uso interno e externo)
- Visão contínua de Governança de Dados.



### Governança de Dados, TI (Segurança e Outros) e Negócio





**Obrigado**  
[radixeng.com.br](http://radixeng.com.br)