

Data Governance Framework: Case study in one Brazilian Organization

André Montoia Barata (Universidade de São Paulo, São Paulo, Brasil) –
andre.barata@usp.br

Edmir Parada Vasques Prado (Universidade de São Paulo, São Paulo, Brasil) –
eprado@usp.br

Resumo

Contar com uma área de TI alinhada às estratégias do negócio está cada vez mais se tornando um diferencial competitivo para as organizações. A Governança de TI (GTI) tem um papel fundamental na realização deste alinhamento, potencializando os processos de TI juntamente com os objetivos do negócio. Alinhar a TI aos negócios da organização é fundamental, porém é necessário também garantir o alinhamento da GTI com a Governança de dados (GD). A GD é responsável pela governança dos dados da organização e proporciona a informação para a tomada de decisões estratégicas para o negócio. Possuir uma GTI alinhada a GD propicia um melhor desempenho para as organizações, que precisam de informações corretas em tempo hábil para a tomada de decisão. Para colaborar com este alinhamento existem os *Frameworks* de boas práticas de gestão, que auxiliam as organizações a implantar esta governança. Este trabalho tem por objetivo identificar os processos e *frameworks* de GD implantados em uma organização brasileira e comparar os benefícios obtidos na implantação com os propostos pela literatura. Para isso será realizado um estudo de caso em uma organização brasileira que implantou processos dos *frameworks* de GD.

Palavras-Chave: Governança de Dados, Governança de TI, Sistemas de Informação.

Abstract

To have an IT aligned with business strategies of organizations is increasingly becoming a competitive advantage for large organizations. IT Governance (ITG) has this key role in achieving this alignment leveraging IT processes with business objectives. Thus, as align IT with the organization's business is fundamental as alignment of GTI with Data Governance (DG). The DG is responsible to managing the organization's data and provides information to making strategic decisions for the business. Have aligned DG with ITG is a better performance for organizations that need the right information in the right time. To aid in this alignment are the frameworks of good management practices that help organizations implement this governance. This study aims to identify the processes and frameworks of DG implemented in a Brazilian organization and compare the benefits achieved in the implementation with the proposed in the literature. For this analyse will be carried out case studies in one Brazilian organization that implemented these DG frameworks.

Key-words: Data Governance, IT Governance, Information Systems.

1 INTRODUÇÃO

As informações estão cada vez mais presentes no dia-a-dia das organizações e fazem parte das decisões estratégicas. Saber analisar, processar e extrair o conhecimento das informações organizacionais é fundamental para a tomada de decisão correta.

O alinhamento estratégico entre TI e negócio não ocorre de maneira simples. Muitas vezes a TI se torna o vilão das organizações devido à falta de alinhamento com seus objetivos. Além disso, muitas organizações deixam a TI como uma área isolada que apenas é utilizada como suporte para as demais áreas da organização (Riekstin, 2012). Assim como alinhar a Governança da TI (GTI) à Governança Corporativa é importante para a transparência e eficiência do negócio, também é fundamental o alinhamento e integração da GTI com a Governança de Dados (GD). A GD é um complemento da GTI com um foco específico em gestão de dados.

As organizações contemporâneas estão cada vez mais se movendo em direção a GD. Isso porque, elas estão reconhecendo os dados como um ativo da organização suportando a tomada de decisões no negócio. Dentro desse cenário, a GD se torna importante porque ela define políticas e procedimentos para assegurar uma gestão de dados proativa e eficiente.

Diversos autores (Cheong& Chang, 2007;Haider& Haider, 2013) que pesquisam sobre GD na literatura relatam sobre a importância de possuir uma GD com processos e práticas bem definidas e um *framework* estruturado gerando benefícios e melhorias para as organizações. Entretanto a literatura não relata se os benefícios propostos com a implantação desses *frameworks* e processos de GD são realmente alcançados após a sua implantação. Além disso, os *frameworks* existentes na literatura não atendem as necessidades e demandas das pequenas e médias empresas.

Verifica-se uma lacuna na literatura em identificar quais os reais benefícios alcançados com a implantação dos processos de GD e se os mesmos são os propostos pela literatura. Dentro desse contexto, o objetivo geral deste estudo consiste em identificar e descrever os processos de GD implantados em organizações brasileira e comparar os benefícios obtidos com aqueles propostos pela literatura. A partir desse objetivo geral, foram definidos os seguintes objetivos específicos: (1) descrever os processos de GD implantados pelas organizações brasileiras e a razão da sua escolha; e (2) descrever os benefícios obtidos com a implantação dos processos de GD e compara-los com os relatados na literatura.

Na seção dois deste artigo está descrito o referencial teórico, construído a partir da literatura, sobre GD e seus *frameworks*. Na seção três estão apresentados os procedimentos metodológicos utilizados na pesquisa, na seção quatro quais foram os resultados obtidos e por fim na seção cinco as considerações finais do trabalho.

2 GOVERNANÇA DE DADOS

A capacidade que uma organização possui em proteger seus dados, revesti-los de qualidade e produzir informações confiáveis, precisas, acessíveis e disponíveis no momento correto é um dos principais fatores determinantes para o valor das organizações modernas (Fernandes& Abreu,2012).

A GD tem como responsabilidade o controle e o gerenciamento dos dados da organização, sendo a responsável por todo controle dos dados e a transformação dos mesmos em informações. A GD é um tema novo que vem ganhando destaque e importância nas

organizações, principalmente em organizações onde existe uma grande massa de dados e estes são de extrema importância para a tomada de decisões.

A GD é um procedimento de tomada de decisões e responsabilidades para com os processos relacionados aos dados, baseando-se em políticas, normas e restrições. O foco de atuação da GD pode variar de organização para organização, começando em privacidade de dados e indo até a qualidade dos mesmos, passando por diversas áreas do ciclo do dado. Para possuir uma GD estruturada e eficiente é preciso que as organizações definam suas necessidades de gestão de dados, bem como os objetivos a serem atingidos, e a partir deste ponto, delimitem o escopo de atuação da GD (Fernandes & Abreu, 2012).

Segundo MOSLEY *et al.* (2009) GD é o exercício da autoridade, do controle e da tomada de decisão sobre os ativos de dados, ou seja, a GD é responsável pelo gerenciamento dos ativos de dados, ela é responsável pelo planejamento, supervisão, controle e uso dos dados. A GD divide-se em dois grandes grupos de atividades: (1) planejamento do gerenciamento de dados; e (2) controle do gerenciamento de dados.

Por meio destas atividades de GD é possível estruturar o gerenciamento de dados de maneira eficiente. Entretanto não basta seguir as atividades e os objetivos relatados, é necessário à utilização de *frameworks* que auxiliem no desenvolvimento e implementação de políticas de GD, para que esta seja bem sucedida e gere valor tanto para os dados da organização quanto para as decisões estratégicas futuras que serão tomadas.

2.1 Trabalhos Correlatos

Uma revisão sistemática sobre GD foi realizada a fim de obter trabalhos correlatos com o tema desta pesquisa, a seguinte estratégia foi utilizada para a revisão sistemática: (1) busca em quatro bases de dados: IEEE, ACM, SCOPUS e PROQUEST; (2) as palavras-chave utilizadas foram “*Data Governance*” e “*Frameworks*”; (3) os critérios de inclusão consideraram os artigos a partir de 2004, disponíveis em versões eletrônicas e com o texto completo e relacionados a processos e *frameworks* de GD, bem como implantação dos mesmos; (4) os critérios de exclusão consideraram artigos que não sejam estudos primários, que não estejam relacionados à GD e que não possuam o texto completo para leitura; e (5) os critérios de qualidade consideraram artigos publicados em periódicos, anais de eventos e revistas renomadas da área de sistemas de informação e com revisão por pares ou revisão por especialistas.

Como resultados foram selecionados 17 artigos, os quais estão apresentados em dois grupos:

a) **Trabalhos sobre *Frameworks* e modelos para Governança de Dados.** Foram encontrados dez artigos que tratam especificamente de *frameworks* ou modelos para GD:

- (1) Haider e Haider (2013) analisam os *frameworks* de GD e de avaliação de maturidade de GD. Os *frameworks* analisados pelos autores são: Dataflux, Governança de Dados da Gartner, Governança de Dados da IBM, DAMA DMBOK. A partir da análise desses *frameworks* e suas funcionalidades e tratativas para a GD nas organizações, os autores desenvolveram um *framework* de GD para o gerenciamento de ativos.
- (2) Khatri e Brown (2010) desenvolveram um *framework* para a GD baseado em domínios de decisões, proporcionando uma GD estratégica e efetiva para as

organizações. Este *framework* baseia-se em cinco domínios de decisões: Princípios de Dados, Qualidade de dados, Metadados, Acesso a Dados e Ciclo de Vida do Dado.

- (3) Begg e Caira (2012) exploram o dilema da GD em PME. As PME possuem particularidades e deficiências para o tratamento dos seus dados, as quais muitas vezes desconhecem as necessidades e possibilidades existentes para tratativas destes dados.
- (4) Otto (2011a) desenvolveu um modelo morfológico baseado na análise teórica da GD para as organizações. Para a construção desse modelo o autor realizou uma análise na literatura sobre pesquisas e publicações relacionadas à GD e realizou um levantamento do material relacionado ao estado da arte sobre GD.
- (5) Cheong e Chang (2007) pesquisam sobre o relacionamento da GTI e a GD e como organizações podem possuir um gerenciamento de dados efetivo sem uma política de GD bem estruturada. Os autores concluem que o gerenciamento de dados e qualidade dos dados de uma organização não é eficiente sem uma boa política de GD.
- (6) Soares, Deutsch, Hanna e Malik (2012) descrevem o *framework* da IBM para análise da maturidade da Governança de Big Data. Os autores relatam a popularização do conceito de *Big Data* e como as organizações devem trabalhar esse conceito para a obtenção de vantagens competitivas e aumento dos lucros, utilizando *Big Data* de maneira correta para proporcionar as informações certas e gerar tomadas de decisões consistentes e produtivas.
- (7) Soares (2012) descreve um *framework* para governança de *Big Data*, o qual é baseado em três dimensões: Tipo de *Big Data*, Disciplinas da GD e Indústrias e Funções. Segundo o autor o *Big Data* é classificado em cinco tipos: Dados Web e de mídias sociais, Dados Máquina-para-Máquina, Transações de grande quantidade de dados, Dados Biométricos e Dados Gerados por Humanos.
- (8) Weber, Otto e Osterle (2009) relatam o estudo de caso de análise da GD em seis organizações internacionais de diferentes indústrias e analisam como essas organizações estruturam sua GD.
- (9) Chapple (2013) descreve a importância de possuir uma GD estruturada com boas práticas e processos de qualidade implantados e em funcionamento. Entretanto o autor relata as dificuldades e desafios encontrados pelas organizações para estruturar sua GD e os grandes problemas enfrentados para sua implantação.
- (10) Olamaet *al.* (2014) descrevem uma metodologia baseada nas melhores práticas do mercado para mensurar o conjunto de dados das organizações baseado no nível de integração desses dados. Os autores propõem um modelo para quantificar o nível de integração de dados de uma organização, por meio dos 10 processos chaves do DMBOK.

b) **Trabalhos sobre processos e práticas para Governança de Dados.** Foram encontrados sete artigos de práticas de GD:

- (1) Felici, Theofrastos e Pearson (2013) apresentam um modelo conceitual para a prestação de contas pelas organizações que utilizam serviços em nuvem suportando a GD. Este modelo consiste em atributos, práticas e mecanismos para a prestação de

contas de serviços em nuvem, bem como permite explicar a interação dos atores e organizações com os ambientes na nuvem.

- (2) Freitas *et al.* (2013), analisam uns dos processos mais importantes para a GD, a qualidade dos dados. Os autores planejam atividades de controle dos dados a fim de aprimorarem o cadastro de clientes de uma instituição financeira, e de aumentar a qualidade e eficiência dos relatórios de sistemas de BI.
- (3) Tallon (2013) analisa práticas e processos de GD que podem auxiliar na geração de valor e minimização dos riscos da utilização de aplicações de *Big Data*. Para que uma organização utilize da grande massa de dados que ela possui, por meio de aplicações de *Big Data*, e proporcionem tomadas de decisões corretas de forma a maximizar os lucros e minimizar os riscos, práticas e processos de GD devem ser bem definidos e estruturados.
- (4) Bruening e Waterman (2010) analisam a utilização de metadados (*Data Tagging*) para auxiliar os *frameworks* e modelos de GD, os quais proporcionam um aumento da eficácia da GD, uma diminuição de riscos e custos para a organização e uma melhoria no desempenho na análise dos dados para a tomada de decisões estratégicas.
- (5) Coleman, Hughes e Perry (2009) analisam a dificuldade de agências e departamentos de organizações em compartilhar informações e utilizar metadados de maneira eficiente. Como conclusão, os autores relatam a importância da utilização de normas e regimes de metadados para possibilitar um aumento na eficiência do compartilhamento de dados da organização.
- (6) Rosenbaum (2010) descreve sobre o significado do *Data Steward* e a GD para a área da saúde e como possuir um modelo de GD eficiente para esta área. A autora descreve o *Data Steward* como o responsável pela condução de todos os processos relacionados à GD e utilização dos dados de maneira eficiente e correta.
- (7) Otto (2011b) realiza um estudo de caso em duas empresas da área de telecomunicações a fim de analisar como a GD está estruturada e organizada nestas empresas. O autor constata que a configuração da GD é determinada por fatores internos e externos às empresas.

2.2 Frameworks de Governança de Dados

A partir da revisão sistemática da literatura verifica-se a existência de vários *frameworks* citados por diversos autores. Entretanto, a grande maioria deles consiste em modelos que são adaptados para atender as necessidades da pesquisa em questão, ou seja, os autores estudados buscam por *frameworks* na literatura e adaptam os processos destes a fim de atender as necessidades específicas da pesquisa que está sendo desenvolvida. Como conclusão da revisão da literatura, foram selecionados sete *frameworks* de GD para serem estudados e para tornarem a base desta pesquisa, os quais estão descritos a seguir.

2.2.1 DAMA DMBOK

A DAMA (*Data Management Association*) é uma organização internacional sem fins lucrativos, constituída por profissionais e técnicos dedicados a promover os conceitos e melhores práticas da Gestão da informação e GD. A DAMA é responsável pelo DMBOK (*Data Management BodyofKnowlegde*) um “corpo de conhecimento” sobre gestão de

dados que foi desenvolvido em 2009 com a participação de mais de 120 profissionais de todo o mundo. Este proporciona uma visão geral sobre gerenciamento de dados, por meio da definição de padrões, terminologias e melhores práticas, entretanto sem o detalhe de técnicas e métodos específicos (DAMAIternational, 2012).

Segundo Mosley *et al.* (2009), existem dez processos-chave para o gerenciamento de dados:

- (1) **Governança de Dados.** A GD é o exercício da autoridade e controle (planejamento, monitoramento e cumprimento) sobre a gestão de ativos de dados.
- (2) **Gerenciamento da Arquitetura de Dados.** Consiste na definição do que os dados organizacionais precisam e desenvolver os desenhos técnicos para proporcionar a definição estruturada.
- (3) **Desenvolvimento de Dados:** Consiste em desenhar, implementar e manter soluções que atendam às necessidades dos dados da organização.
- (4) **Gerenciamento de Operações Database.** Consiste em planejar, controlar e suportar as estruturas dos ativos de dados durante seu ciclo de vida, desde a criação e aquisição do dado até o arquivamento e eliminação do dado.
- (5) **Gerenciamento de Segurança dos Dados.** Consiste em planejar, desenvolver, e executar políticas e procedimento de segurança a fim de promover autenticação, autorização, acesso controlado e auditoria de dados e informações.
- (6) **Gerenciamento de Dados Mestres e Referência.** Consiste em planejar, implementar e controlar atividades a fim de garantir a consistência de dados mestres e referências (DMR).
- (7) **Gerenciamento de Data Warehousing e Business Intelligence.** Consiste em planejar, implementar e controlar processos para prover dados de suporte à tomada de decisões e apoio a colaboradores envolvidos na produção de relatórios de conhecimentos, consultas e análises.
- (8) **Gerenciamento de Documentação e Conteúdo.** Consiste em planejar, implementar, e controlar atividades para armazenar, proteger e acessar dados encontrados em arquivos eletrônicos e registros físicos.
- (9) **Gerenciamento de Metadados.** Consiste em planejar, implementar e controlar atividades para garantir um fácil acesso aos metadados integrados e de qualidade.
- (10) **Gerenciamento da Qualidade dos Dados.** Consiste em planejar, implementar e controlar atividades que apliquem técnicas de gestão de qualidade de dados para medir, avaliar, melhorar e garantir a adequação dos dados ao seu uso.

2.2.2 DATAFLUX

O *framework* DataFlux consiste em um modelo para análise da maturidade da GD das organizações, auxiliando-as a entender o nível de gerenciamento de dados atual e proporcionar um caminho para o crescimento no futuro. O *framework* propõe um progresso por meio de quatro fases: Indisciplinado, Reativo, Proativo e Governado. Organizações precisam identificar em qual fase elas se encontram e o porquê elas estão

nessa fase. O *framework* auxilia as organizações a avaliar qual a melhor fase para o negócio que elas atuam e como e quando elas devem avançar para a próxima fase. (DATAFLUX CORPORATION, 2007). A Figura 1 mostra o *framework* para a maturidade da GD e o uso de aplicações corporativas comuns para cada uma das fases.

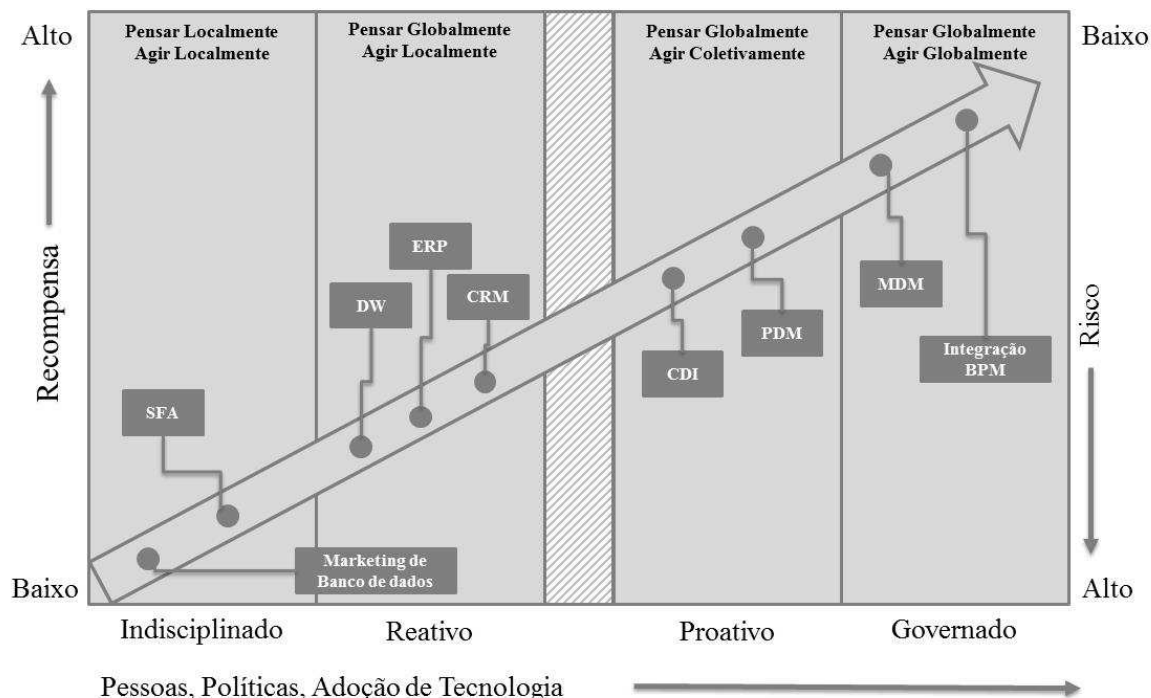


Figura 1 - Framework Dataflux
Fonte: traduzido de Dataflux Corporation (2007)

No fim da implantação dos processos a organização terá implantado uma visão única da organização proporcionando um gerenciamento de dados mestres eficaz, bem como integração de dados de alta qualidade com a gestão dos processos de negócio. Neste nível a organização possuirá rotinas automáticas de controle da qualidade de dados, as quais não necessitarão da intervenção do ser humano (Dataflux Corporation, 2007).

Para cada fase do Dataflux, os seguintes componentes devem ser definidos e estruturados:

- (1) Pessoas. Envolvidos e suas contribuições.
- (2) Políticas. Atividades e regras de negócios para atingir a GD.
- (3) Tecnologias. Investimentos e tecnologia necessários.
- (4) Riscos e Recompensas. Riscos que a organização encontra na fase atual e o ganho obtido com o avanço para a próxima fase.
- (5) Avançando para a próxima fase. Ações necessárias para o avanço para a próxima fase

2.2.3 Enterprise Information Management (EIM)

O *framework* EIM foi elaborado pela Gartner®, possui foco na análise da maturidade da GD e auxilia as organizações a aumentarem seu nível na cadeia do *framework*. Ele é formado por seis estágios, conforme apresentado na Figura 2: Inconsciente, Consciente,

Reativo, Proativo, Gerenciado e Efetivo. O nível zero (Inconsciente) é quando as atividades do EIM estão se iniciando e o nível cinco (Efetivo) é quando o EIM tornou-se totalmente finalizado e estruturado.

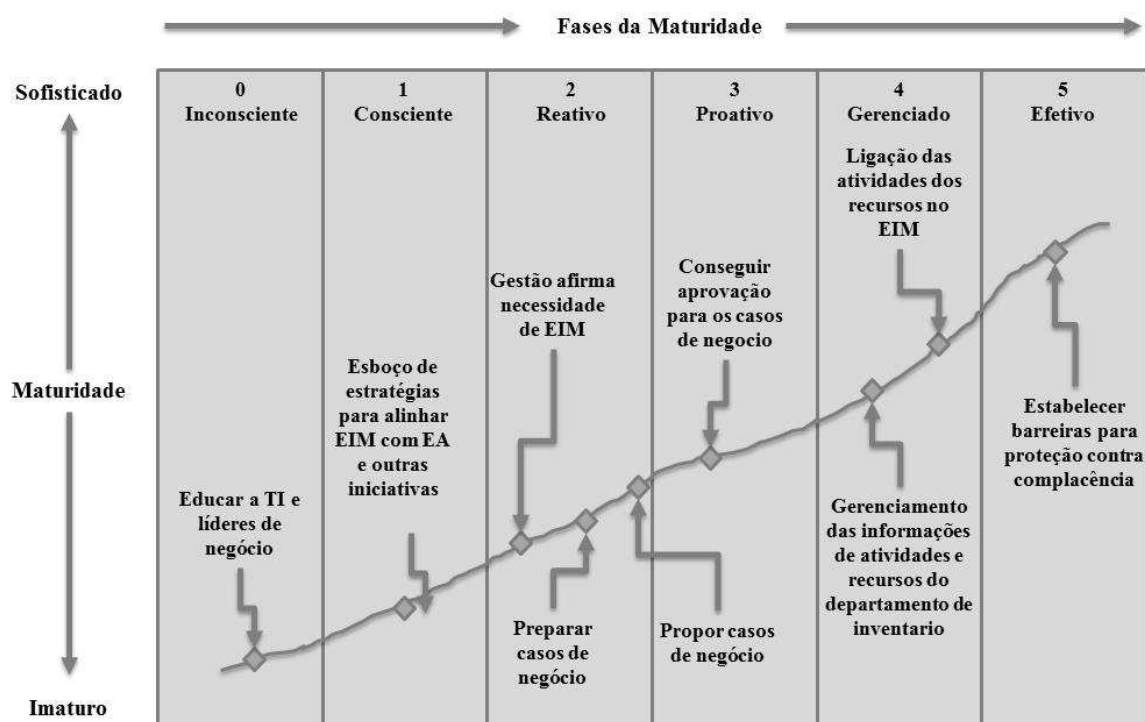


Figura 2 - Enterprise Information Management (EIM)

Fonte: Traduzido de Newman e Logan (2008)

As principais características de cada nível desse *framework* e os passos necessários para que a organização alcance o próximo nível estão descritos a seguir:

- (0) **Inconsciente (Nível 0).** A organização corre risco significativo de informações subgeridas, tal como falhas de conformidade, mau serviço ao cliente e baixa produtividade.
- (1) **Consciente (Nível 1).** A organização alcança alguma consciência sobre a gestão da informação.
- (2) **Reativo (Nível 2).** Líderes do negócio e de TI reagem favoravelmente à demanda por consistência, acurácia e rápida informação entre as unidades de negócio. Eles tomam medidas corretivas para resolver necessidades imediatas.
- (3) **Proativo (Nível 3).** A organização percebe a necessidade de informações para melhorar o desempenho da empresa, e portanto move-se a gestão para o projeto EIM. Informações suportam otimizações de processos.
- (4) **Gerenciado (Nível 4).** A organização percebe que a informação é crucial para o negócio. A organização possui partes significativas implementadas do EIM, incluindo a infraestrutura de informações consistentes.

- (5) **Efetivo (Nível 5).** A organização explora as informações em toda a cadeia de suprimentos, com acordo de níveis de serviços bem definidos e constantemente revisados.

2.2.4 Governança Big Data

O *framework* de Governança Big Data foi elaborado pela IBM ® e possui três dimensões para a governança da Big Data, conforme ilustra a Figura 3.

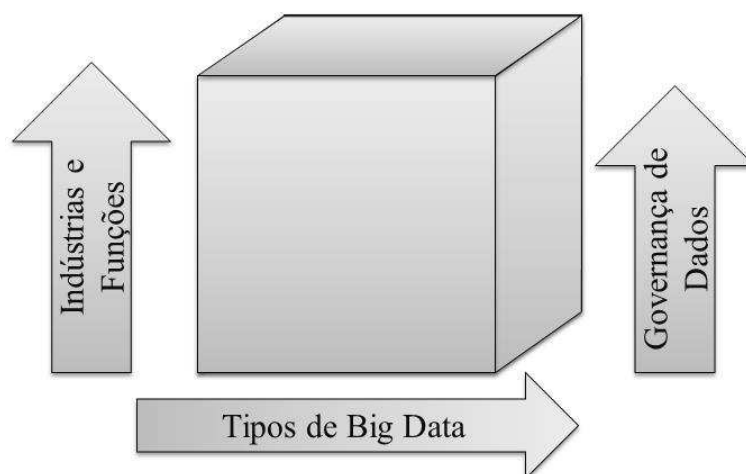


Figura 3 - Framework IBM três dimensões Governança Big Data
Fonte: traduzido de Soares (2012)

As três dimensões deste *framework* consistem em (Soares, 2012):

- (1) **Tipo de Big Data.** Big Data pode ser classificada em cinco tipos: dados web e de mídias sociais; dados máquina-para-máquina; transações de grande quantidade de dados; dados biométricos; e dados gerados por humanos.
- (2) **Disciplinas da GD.** As tradicionais disciplinas da GD são Metadados, Qualidade dos Dados, Integração de Dados Mestres, Gestão do ciclo de vida da informação, Organização, Privacidade e Processo de Integração de Negócios.
- (3) **Indústrias e Funções.** A análise do Big Data é dirigida por meio de casos de uso específicos para indústria e funções, tais como, Marketing, Serviço ao Cliente, Segurança da Informação, ou TI.

Para cada indústria e função existem estudos de casos específicos, os quais possuem um tipo de Big Data, disciplinas de GD e uma solução específica para a resolução do problema. Por exemplo, para a indústria de telecomunicações os tipos de Big Data são dados web e mídias sociais e as disciplinas de GD utilizadas são privacidade e integração de dados mestres. Para a área de segurança da informação o tipo de Big Data é dado máquina-para-máquina e a disciplina utilizada é metadados. Portanto, para cada indústria e função específica existem dimensões do *framework* que atendam às necessidades do negócio (Soares, 2012).

2.2.5 Maturidade da Governança Dados

O *framework* de Maturidade da Governança Dados também foi elaborado pela IBM ®, possui onze categorias e está ilustrado na Figura 4. Para cada categoria existe um bloco de

questões que possibilitam a análise da maturidade de GD da organização baseada na categoria que está sendo analisada (Soares,Deutsch, Hanna&Malik, 2012). As principais questões por categorias podem ser observadas a seguir.

- (1) **Resultados dos negócios:** identificação dos *stakeholders* chaves no negócio para o programa de governança de *Big Data*; e quantificação dos benefícios financeiros da governança de *Big Data*.
- (2) **Estrutura organizacional e Consciência:** definição de escopo para a *Big Data* que será aplicado na organização; priorização dos tipos de *Big Datas* que precisam ser governados; e comunicação clara e troca de informação entre o time que é responsável pelo repositório de *Big Data* e o time tradicional.



Figura 4 - Framework de Maturidade da Governança Dados

Fonte: traduzido de Soares, Deutsch, Hanna e Malik (2012)

- (3) **Stewardship:** definição da matriz de papéis e responsabilidades para os elementos de dados críticos; e formalização dos papéis e responsabilidades dos *Data Stewardship* com os Recursos Humanos.
- (4) **Gestão de riscos de dados:** determinação dos *stakeholders* chave para a governança de *Big Data*; estabelecimento da ligação entre Governança de *Big Data* e Gestão de Risco; e existência de um planejamento realista da continuidade do negócio.
- (5) **Política:** documentação sobre as políticas para Governança de *Big Data*; possibilidade de inspeção para execução; e tradução das políticas em um conjunto de controle operacionais.
- (6) **Gestão da qualidade de dados:** consenso sobre os problemas associados à qualidade de *Big Data* em que o valor do dado pode ou não pode ser alto ou obvio; qualidade de dados está sendo inspecionada de uma forma repetível e documentada; e identificação das dimensões da qualidade dos dados que podem ser menos aplicadas a *Big Data*.
- (7) **Gestão do ciclo de vida da informação:** determinação do volume de armazenamento para *Big Data* e a taxa anual de crescimento; determinação do custo do

armazenamento do *Big Data* e da taxa anual de crescimento; e o programa de retenção deve incluir as citações legais que orientam a retenção de *Big Data* por país, estado e província.

- (8) **Gestão da segurança e privacidade:** deve existir um entendimento dos regulamentos de privacidade, dos termos de uso dos dados das mídias sociais e das políticas definidas para o uso aceitável dos dados pelos clientes; e deve existir criptografia nos ambientes de produção para qualquer *Big Data* sensível.
- (9) **Arquitetura de dados:** determinação das aplicações que devem ser movidas para as infraestruturas das plataformas de *Big Data*; determinação das aplicações que devem ficar de fora das infraestruturas das plataformas de *Big Data*; e considerar o impacto dos dados mestres e dados referenciais no *Big Data*.
- (10) **Classificação e metadados:** existência de uma terminologia para o negócio, incluindo termos chaves relacionados ao *Big Data*; identificação dos *Data Stewards* para gerenciar os termos chaves do *Big Data*; determinação da frequência com que os metadados são atualizados ou mantidos em sincronia com o negócio ou a TI; e verificação da existência de uma captura dos metadados chaves operacionais para identificar situações em que o *Big Data* não é carregado.
- (11) **Auditoria de registro de informações e relatórios:** verificação da existência de administradores de banco de dados, empreiteiros e outros terceiros que possuem acesso sem criptografia a *Big Data* sensíveis, tais como, dados de geolocalização, detalhes de chamadas telefônicas, medidores de leitura inteligentes e alegações de saúde.

As organizações precisam governar a *Big Data* como qualquer outro dado organizacional e regras e políticas devem ser seguidas e implementadas para que resultados possam ser alcançados (Soares,Deutsch, Hanna&Malik, 2012).

2.2.6 Gestão de Ativos

A gestão de ativos nas organizações precisa de uma visão em longo prazo para gerar, processar e gerenciar os dados permitindo uma melhoria contínua em soluções que suportem o ciclo de vida dos ativos. Portanto, para gerenciar o ciclo de vida do ativo existe a necessidade de um aprendizado de como planejar, executar e gerenciar os ativos de dados da organização. Baseado nessas constatações, a Figura 5 ilustra o *framework* de GD para a gestão de ativos. O ciclo de vida do ativo tem sete perspectivas: competitividade, desenho, operações, suporte, *stakeholders*, eficiência do ciclo de vida e perspectiva de aprendizagem. Este *framework* incorpora aspectos como qualidade de dados, integração, padronização, interoperabilidade e gerenciamento de risco, por meio de conexões entre diferentes perspectivas (Haider& Haider, 2013).

A seta circular no meio da Figura 5 significa que o *framework* permite o ciclo de melhoria contínua com base na aprendizagem generativa da organização. Este *framework* orienta como funções de GD devem ser implementadas e ao mesmo tempo pode avaliar a eficácia dos dados existentes nas funções de governança. Assim os dados tornam-se um facilitador estratégico na gestão do ciclo de vida dos ativos e permite a aprendizagem generativa. O *framework* também fornece uma análise do *gap* entre a maturidade da GD desejada e a atual que a organização se encontra (Haider& Haider, 2013).

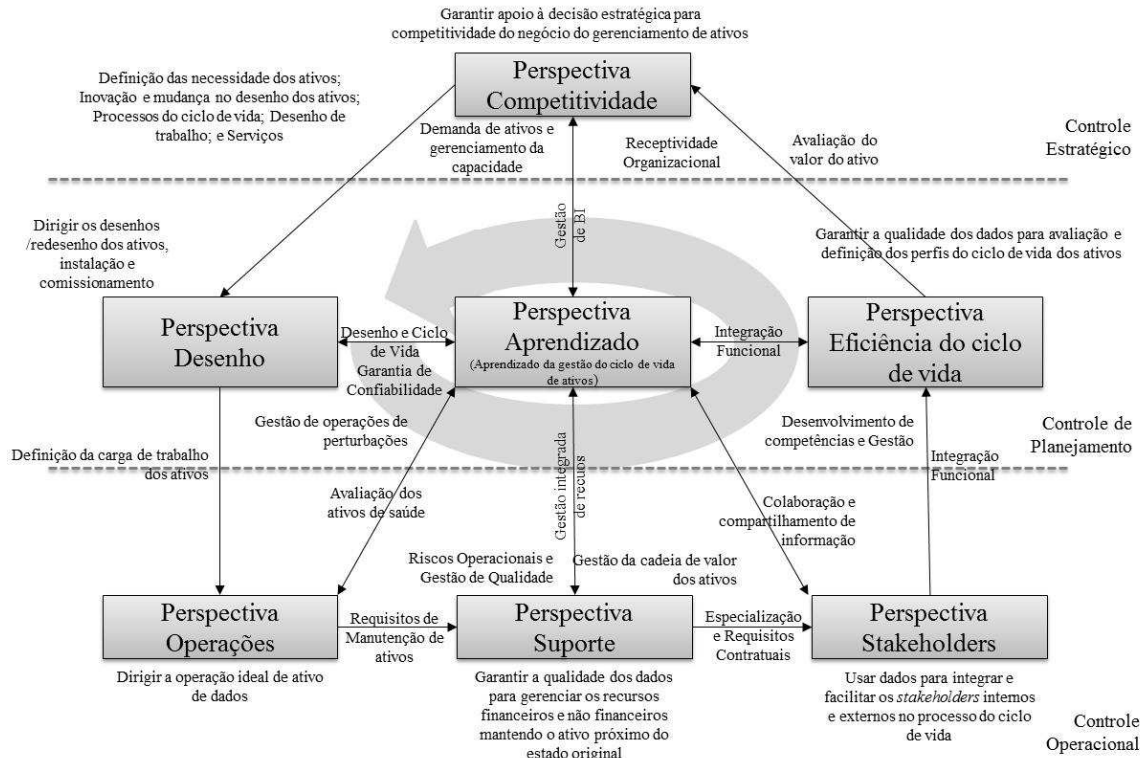


Figura 5 - Framework de Governança de Dados para Gestão de Ativos
Fonte: traduzido de Haider e Haider (2013).

A Figura 6 ilustra a implementação do framework nas organizações. Primeiramente as organizações precisam dos processos do negócio dentro de cada perspectiva descrita no framework. O próximo passo é desenvolver uma linha de base para o controle das iniciativas de GD dentro da organização. Esta linha de base será baseada nos processos do framework DAMA DMBOK. Esses processos de GD devem ser aplicados para cada um dos processos de negócio identificados para os domínios das perspectivas de gestão de ativos (Haider & Haider, 2013).

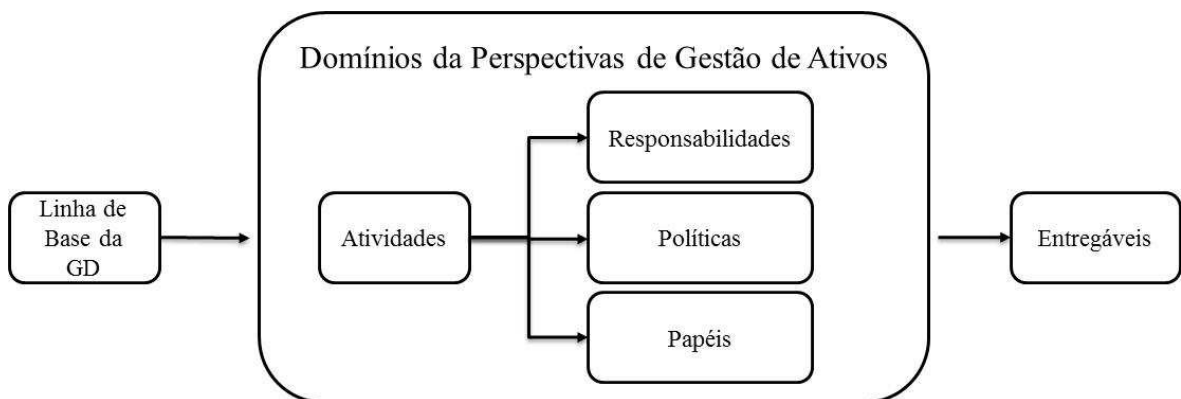


Figura 6 - Implementação nas organizações
Fonte: traduzido de Haider e Haider (2013)

Introdução, sustentação e maturidade da GD na gestão de ativos dependem de três dimensões: processos de ciclo de vida do ativo, elementos funcionais da GD e fatores críticos que contribuem para competência organizacional e facilitam a GD, tendo em vista estes pontos é essencial à avaliação da eficácia das iniciativas da GD. A Figura 7 ilustra o

cubo da GD que deve ser aplicado para cada uma das sete perspectivas do *framework*. Para cada processo, funções de GD serão avaliadas de acordo com as diferentes dimensões das competências organizacionais (Haider& Haider, 2013).

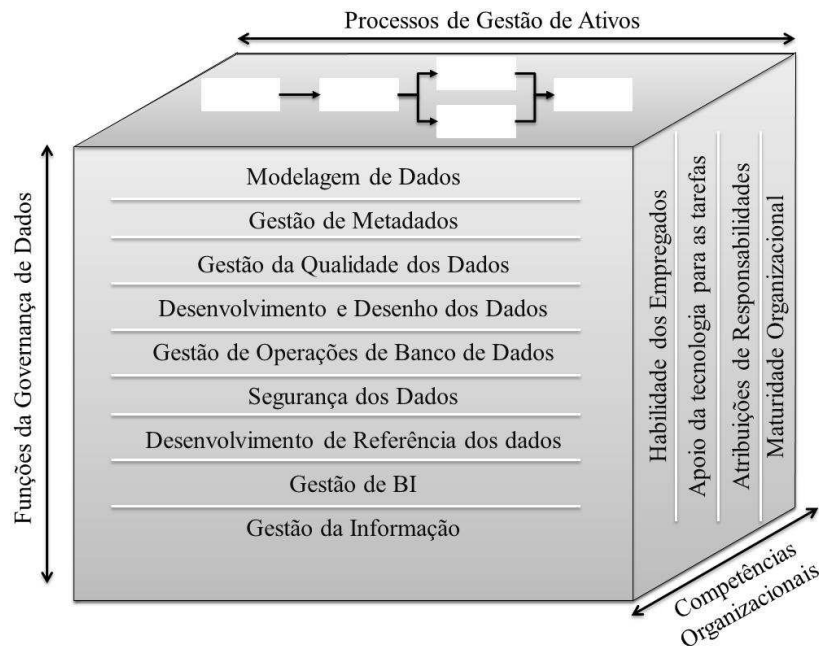


Figura 7 - Cubo da Governança de Dados
Fonte: traduzido de Haider e Haider (2013)

Ao realizar essas avaliações pode-se mensurar a lacuna entre o estado atual e o ideal para a GD na organização. Esta análise pode funcionar como um roteiro para sustentação, maturidade e melhoria contínua das funções GD na organização.

2.2.7 Governança de Domínio de Decisões de Dados

O *framework* de governança de domínios de decisões de dados é baseado em cinco domínios da GD que se relacionam entre si, conforme ilustra a Figura 8.

Princípios de dados		
Qualidade de Dados	Metadados	Ciclo de Vida do Dado
	Acesso aos Dados	

Figura 8 - Domínios de Decisões da Governança de Dados
Fonte: traduzido de Khatri e Brown (2010)

A partir de cada domínio de decisão e suas características, perguntas são elaboradas a fim de obter tipos de decisões a serem tomadas para cada domínio.

- (1) **Princípios de Dados.** Clarificar o papel dos dados como um ativo. **Funções potenciais e responsabilidades:** *Data Owner* (Proprietário dos Dados); *Data Custodian* (Responsável pela custódia, transporte e armazenamento dos dados); *Data Steward* (Responsável pelo conteúdo associação e regras de negócios dos dados);

Data Supplier (Fornecedores de dados); *Data Consumer* (Consumidores de dados); e *Data Council* (Principal departamento responsável pelas políticas de dados).

- (2) **Qualidade de Dados.** Estabelecer requisitos para a utilização dos dados. **Funções potenciais e responsabilidades:** *Data Owner*; especialista no assunto; Gerente de Qualidade de Dados; e Analista de Qualidade de Dados.
- (3) **Metadados.** Estabelecer a semântica ou o “conteúdo” dos dados, de modo que ele seja interpretável pelos usuários. **Funções potenciais e responsabilidades:** arquiteto de dados corporativos; modelador de dados corporativos; engenheiro de modelagem de dados; arquiteto de dados; e comitê de arquitetura empresarial.
- (4) **Acesso aos Dados.** Especificar os requisitos para acesso aos dados. **Funções potenciais e responsabilidades:** *Data Owner*; beneficiário dos dados; gestor de segurança de dados; agente de segurança de dados; analista técnico de segurança de dados; e comitê de arquitetura de desenvolvimento empresarial.
- (5) **Ciclo de Vida do Dado.** Determinar a definição, produção, retenção e a retirada dos dados. **Funções potenciais e responsabilidades:** arquiteto de dados corporativos; e gestor da cadeia de informação.

2.2.8 Framework de NotreDame

O *framework* de GD da universidade de NotreDame baseia-se em duas bases principais: acesso aos dados e tecnologias, que proporcionam a sustentação do *framework*. Possui também cinco pilares – qualidade e consistência, políticas e padrões, segurança e privacidade, *compliance* e retenção e arquivamento – que interligam as duas bases. O *framework* está representado na Figura 9.

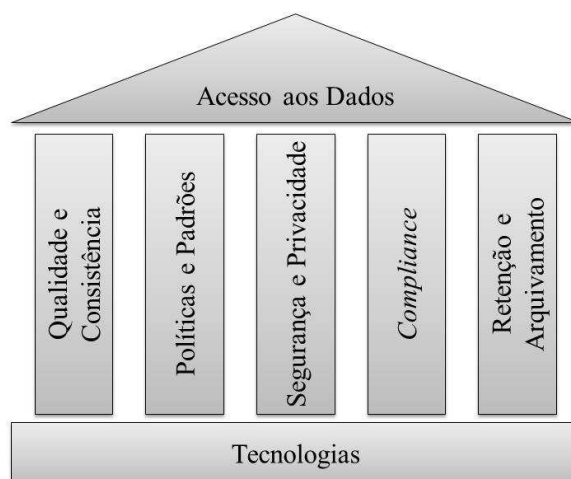


Figura 9 - Framework de Governança de Dados de NotreDame

Fonte: traduzido de Chapple (2013)

Segundo Chapple (2013), cada pilar desse *framework* desenvolve metas e objetivos que estão descritos a seguir:

- (1) **Acesso aos Dados.** Tem o objetivo promover aos usuários a habilidade de acesso aos dados que eles precisam para o negócio e um tempo hábil para o gerenciamento destes dados.

- (2) **Tecnologias.** Proporcionam as ferramentas de suporte para o desenvolvimento de práticas de dados juntamente com os processos do negócio.
- (3) **Qualidade e Consistência.** Garante que os dados utilizados pelos *stakeholders* da organização vieram de fontes confiáveis e integras e que todos os envolvidos possam interpretar os dados da mesma forma.
- (4) **Políticas e Padrões.** Proporciona a política de GD bem definida para a organização a fim de suportar as práticas e os padrões a serem desenvolvidos nos outros pilares do *framework*.
- (5) **Segurança e Privacidade.** Controles e práticas são desenvolvidos e implementados a fim de garantir a segurança e a privacidade adequada aos dados que são utilizados pela organização.
- (6) **Compliance.** Proporciona a garantia que os dados sensíveis estão seguindo leis e normas e regulamentações impostas pelo governo.
- (7) **Retenção e Arquivamento.** Proporciona práticas eficazes e eficientes a fim de garantir a preservação dos dados e informações para futuras gerações.

3 MÉTODO DE PESQUISA

Esta pesquisa é do tipo exploratório e qualitativo. Qualitativo, porque visa descobrir e refinar um tema de pesquisa por meio de coleta de dados, com descrições e observações e sem a utilização de medições numéricas. Exploratório, porque tem a finalidade de familiarizar-se com o tema de pesquisa e obter a imersão inicial explorando o assunto abordado, além do esclarecimento de um conceito de pesquisa e a criação de novas hipóteses (Sampieri, Collado&Lucio, 2006).

Como estratégia de pesquisa será utilizado o estudo de caso. Usa-se o estudo de caso quando se deseja analisar e avaliar os resultados de um determinado acontecimento, por exemplo, realiza-se um estudo de caso para avaliar a implantação dos processos de GD, coletando dados e informações a fim de aplica-los em um modelo, obter os resultados da análise e proporcionar as conclusões do estudo. Um estudo de caso pode ser único ou múltiplo, de acordo com a abordagem e o direcionamento utilizado. No contexto da pesquisa, a estratégia do estudo de caso único será utilizada, a fim de realizar uma avaliação inicial sobre a implantação da GD em uma organização brasileira de grande porte. Esta estratégia será utilizada para a realização da imersão inicial sobre o assunto e a análise da implantação de processos/*framework* de GD em uma organização. (Yin, 2009).

Toda pesquisa possui um projeto de pesquisa, ou seja, uma sequência lógica que conecta os dados empíricos aos objetivos iniciais do estudo e as conclusões a serem obtidas. O estudo de caso possui cinco componentes de um projeto de pesquisa: (Yin, 2009):

1. **Questões do Estudo:** Estudo de caso visa responder questões de pesquisa do tipo “Como” e “Por que”, questões que estão em concordância com os objetivos e as questões dessa pesquisa que pretende analisar como uma organização brasileira implantou os processos de GD, os benefícios que estes processos trouxeram, e o porquê da escolha dos processos implantados.

2. **Proposições do estudo:** São as proposições que devem ser examinadas a partir das questões e objetivos da pesquisa, por exemplo, identificar os processos de GD implantados em uma organização brasileira e a razão da sua escolha.
3. **Unidade de Análise:** Consiste em identificar o “caso” do estudo de caso, nesta pesquisa o “caso” é a organização brasileira que implantou processos de GD e participou do estudo de caso.
4. **Lógica que une os dados às proposições:** Consiste na análise dos dados que serão coletados no estudo de caso, tais como: combinação de padrão, construção de explanação, análise de séries temporais, modelos lógicos e síntese de casos cruzados.
5. **Crítérios para interpretar as constatações:** Este componente constitui as técnicas que serão utilizadas para interpretação das análises dos dados, estatísticas são muito utilizadas nesse componente, além de gráficos comparativos, tabelas, entre outros.

A estruturação desses cinco componentes proporciona à construção de uma teoria preliminar a coleta dos dados, a qual auxiliará na coleta, análise e conclusões que serão obtidas (Yin, 2009).

Para a realização do estudo de caso desta pesquisa escolheu-se uma organização brasileira do ramo de atuação agrícola que está em fase final da implantação de processos de GD para estruturar os dados mestres de materiais. Para esta organização foram escolhidos os seguintes entrevistados:

1. **Gerente do projeto da consultoria** – Gerente da consultoria responsável pela implantação dos processos/*framework* na organização estudada
2. **Analista de dados da consultoria** – Consultor com conhecimentos em GD que participou da implantação dos processos/*framework*.

As entrevistas serão realizadas com os envolvidos na ordem que foi apresentado anteriormente.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Foi desenvolvido um instrumento para a realização do estudo de caso e as entrevistas com os envolvidos. O instrumento é constituído pelos seguintes tópicos:

- (a) **Processos e Benefícios:** Esta parte do instrumento é composta pelos 11 processos mais relevantes da literatura, ou seja, os processos mais utilizados e pesquisados pelos autores, bem como os benefícios que a implantação desses processos traz. Serão questionados aos entrevistados quais processos foram implantados na organização estudada e a porcentagem de implantação atingida, bem como os benefícios atingidos. Também existe a possibilidade de o processo ter alcançado outro benefício para a organização, o que será levantado durante a entrevista. Após a identificação dos processos implantados e benefícios atingidos, algumas perguntas serão realizadas, tais como: 1) Por que foram escolhidos esses processos para a implantação?; 2) Como foi realizada a implantação desses processos?; 3) Qual foi a estratégia utilizada para a implantação dos processos?; 4) Quais foram as dificuldades e limitações encontradas na implantação?; 5) Quais foram as lições aprendidas com a implantação?. Estas perguntas visam identificar como foi realizada a implantação dos processos, o que motivou esta implantação, dificuldades encontradas e lições aprendidas com o projeto.

- (b) Organização:** Nesta etapa do instrumento serão realizadas perguntas para o entendimento da organização que foi implantado os processos/*framework* de GD, tais como, número de funcionários, ramo de atuação, faturamento, produtos comercializados, tamanho da área de TI, patrocinadores do projeto, áreas impactadas e beneficiadas com o projeto, principais envolvidos no projeto e a abrangência do projeto de implantação.
- (c) Consultoria:** A última parte do instrumento consiste em perguntas relacionadas à consultoria responsável pela implantação dos processos/*frameworks* de GD. O objetivo desta etapa é levantar informações tais como, início do projeto, tamanho da equipe que implantou o projeto, ordem de grandeza do projeto, escopo de implantação, tempo de implantação, nível dos recursos envolvidos, metodologia, gestão adotada no projeto e *framework* adotado no projeto.

Este instrumento foi aplicado de forma individual com o Gerente da consultoria responsável pela implantação da GD e com o Analista de dados que participou da implantação. Os resultados obtidos serão descritos a seguir:

4.1 Entrevista com o Gerente de Dados

A Tabela 1 demonstra os 11 processos mais relevantes na literatura, bem como a porcentagem de implantação de cada processo na organização estudada, segundo a visão do entrevistado. Também estão detalhados os benefícios previstos na literatura para cada processo, confrontados com os benefícios alcançados com a implantação do processo, também segundo a visão do entrevistado.

Tabela 1 – Processos x Benefícios Entrevista Gerente de Dados

Nº	Processos	% Implantação	Benefícios Previstos	Benefícios Alcançados
1	Políticas e Padrões de Dados	50%	1) Políticas e padrões interligados a necessidade do negócio	X
			2) Políticas de GD bem definidas e estruturadas	
			3) Melhora no controle dos processos de negócio	X
				Otimização da área de suprimentos
2	Gerenciamento da Arquitetura de Dados	75%	1) Melhoria no desempenho da análise dos dados	X
3	Desenvolvimento de Dados	0%	1) Melhora na performance, manutenção e utilização dos dados	
4	Gerenciamento de Operações Database	0%	1) Dados preservados e efetivamente arquivados	
5	Gerenciamento de Segurança dos Dados	0%	1) Confiabilidade nos dados que estão seguros e privados	
			2) Acesso aos dados com segurança e integridade	
			3) Riscos relacionados aos dados e informações minimizados	
6	Gerenciamento de Dados Mestres e Referência	100%	1) Controle/Redução dos custos relacionados a problemas com dados	X
			2) Local centralizados para obtenção e utilização dos dados mestres da organização	X
7	Gerenciamento de Data Warehousing (DW) e Business Intelligence (BI)	0%	1) Melhora na performance e eficiência dos relatórios BI	
8	Gerenciamento de Documentação e Conteúdo	25%	1) Aumentar a produtividade e agilidade no trabalho com dados	X
			2) Taxonomia, padronização, armazenamento e utilização dos dados bem definidos e estruturados	X
9	Gerenciamento de Metadados	0%	1) Dados interpretados de maneira eficaz e eficiente uso das informações	
			2) Melhoria na tomada de decisões estratégicas	
10	Gerenciamento da Qualidade dos Dados	100%	1) Dados com Acurácia, Atualidade, Consistência, Completude e Integridade	X

Nº	Processos	% Implantação	Benefícios Previstos	Benefícios Alcançados
			2) Diminuição de custos relacionados a duplicidade de dados	X
			3) Aumento da satisfação do consumidor dos dados	X
11	Compliance e Auditoria de Dados	75%	1) Dados aderentes a normas legais e regulamentados	X
			2) Redução dos riscos de multas e ações judiciais devido a violação dos dados	X
			3) Controle de Fraudes de dados	

Fonte: próprio autor

Pode-se observar que segundo o Gerente de Dados foram implantados seis processos de um total de 11, destes apenas dois deles chegaram em 100% da implantação (**Gerenciamento de Dados Mestres e Referência e Gerenciamento da Qualidade de Dados**). **Políticas e Padrões de dados** atingiu 50% da implantação e até o final do projeto atingirá os 100%, após a entrega da política de GD e implantação do comitê de Gestão de Dados. **Gerenciamento da Arquitetura de Dados** atingiu 75% de implantação, isso devido ao escopo aplicado no projeto ser apenas os dados mestre de materiais, apesar de representar grande quantidade de dados na organização, existem outros dados mestres além de materiais, como clientes e fornecedores, por exemplo, que não estavam no escopo da GD. **Gerenciamento de Documentação e Conteúdo** atingiu 25% de implantação, isso porque não foi implantado e desenvolvido a parte de inativação e eliminação de dados no ciclo de vida da informação, isso aconteceu por opção da organização em não eliminar dados, por isso o processo não foi implantado em 100%. **Compliance e Auditoria de Dados** atingiu 75% da implantação, foi implantado todos os processos para a aderência de normas legais, entretanto está faltando o controle da auditoria de dados, que são os procedimentos realizados pelo comitê de gestão de dados para o controle dos *logs* de alteração de dados, esta parte será implantada assim que o comitê estiver em funcionamento.

Dentre os processos com 0%, o entrevistado relata que o processo **Desenvolvimento de Dados**, não proporciona benefícios na prática devido à dificuldade de customizações no SAP, por esse motivo não foi implantado. O processo **Gerenciamento de Segurança dos Dados** não foi implantado porque como o escopo é apenas dado mestre de materiais não existia a necessidade de uma aplicação de segurança nesses dados, a qual seria necessária caso fosse um dado mestre de cliente, funcionário, entre outros. O processo **Gerenciamento de Metadados**, exige um nível de maturidade da GD elevado, principalmente nos processos de dados mestres e de qualidade, para que os metadados sejam implantados de forma efetiva e eficiente. Os outros processos com 0% não estavam no escopo de implantação ou até já estavam implantados na organização.

Pode-se observar que para os processos implantados, todos os benefícios relatados na literatura foram atingidos, com exceção do processo **Políticas e Padrões de Dados** que o benefício “Políticas de GD bem definidas e estruturadas” não foi atingido, porque a política ainda não foi entregue, e um novo benefício foi alcançado com esse processo: “Otimização da área de suprimentos”. Para o processo **Compliance e Auditoria de Dados** o benefício “Controle de Fraudes de dados” também não foi atingido, porque o comitê de gestão de dados ainda não foi implantado.

Como a principal motivação para a implantação da GD, a organização buscava por tomada de decisões menos intuitivas e sim mais factíveis, ou seja, tomada de decisões a partir de dados e informações mais confiáveis, para isso existe a necessidade de uma GD. Os processos implantados foram escolhidos pela organização em conjunto com a consultoria e

por ser a solução para a causa raiz do problema, ou seja, proporcionar dados com qualidade e íntegros para a tomada de decisões baseadas em informações confiáveis. Primeiramente foram implantados os processos: **Gerenciamento de Dados Mestres e Referências** e **Gerenciamento da Qualidade dos Dados**, e a partir da implantação desses dois processos bases para a GD foram implantados os outros quatro processos descritos anteriormente. A estratégia utilizada para a implantação foi a seguinte: Inicialmente realiza-se o congelamento da informação e a limpeza e padronização dos dados, em paralelo o desenvolvimento dos processos que serão implantados, bem como a política de GD. Após a padronização dos dados, atualiza-se o ambiente com os novos dados padronizados e inicia-se a tratativa do delta, porque após o congelamento da informação, novos dados surgiram e devem ser tratados como delta. Simultaneamente a tratativa do delta inicia-se a implantação dos processos de GD para o ambiente. Depois de finalizado a implantação dos processos de GD, atualiza-se o ambiente com o delta.

Como principal dificuldade e limitação encontrada na implantação dos processos foi o ERP, que no caso da organização estudada é o SAP. As restrições impostas pelo SAP, tais como dificuldade de customizações de tabelas mestres, e os riscos que a organização enfrentaria por customizar o SAP eram grandes. Outra dificuldade encontrada foi à transferência de conhecimento entre os envolvidos no processo, porque as diversas áreas e pessoas envolvidas possuem cada uma um padrão de qualidade específico, e a transferência do conhecimento novo implantado foi uma grande dificuldade. Como lição aprendida na implantação dos processos consiste na realização de um grande mapeamento de todos os envolvidos e afetados pela implantação, a fim de minimizar os futuros constrangimentos de áreas que não foram mapeadas, porque muitas vezes a própria organização não tem o conhecimento de que determinada área era impactada pelos processos de GD.

4.2 Entrevista com o Analista de Dados

A Tabela 2 demonstra os 11 processos mais relevantes na literatura, bem como a porcentagem de implantação de cada processo na organização estudada, segundo a visão do entrevistado. Também estão detalhados os benefícios previstos na literatura para cada processo, confrontados com os benefícios alcançados com a implantação do processo, também segundo a visão do entrevistado.

Tabela 2 – Processos x Benefícios Entrevista Analista de Dados

Nº	Processos	% Implantação	Benefícios Previstos	Benefícios Alcançados
1	Políticas e Padrões de Dados	100%	1) Políticas e padrões interligados a necessidade do negócio	X
			2) Políticas de GD bem definidas e estruturadas	X
			3) Melhora no controle dos processos de negócio	X
				Agilidade na utilização dos dados, principalmente na área de suprimentos
2	Gerenciamento da Arquitetura de Dados	100%	1) Melhoria no desempenho da análise dos dados	X
3	Desenvolvimento de Dados	100%	1) Melhora na performance, manutenção e utilização dos dados	X
				Possível redução de custos devido a padronização dos dados
4	Gerenciamento de Operações Database	0%	1) Dados preservados e efetivamente arquivados	
5	Gerenciamento de	100%	1) Confiabilidade nos dados que estão	X

Nº	Processos	% Implantação	Benefícios Previstos	Benefícios Alcançados
	Segurança dos Dados *		seguros e privados	
			2) Acesso aos dados com segurança e integridade	X
			3) Riscos relacionados aos dados e informações minimizados	X
6	Gerenciamento de Dados Mestres e Referência	100%	1) Controle/Redução dos custos relacionados a problemas com dados	X
			2) Local centralizados para obtenção e utilização dos dados mestres da organização	X
7	Gerenciamento de <i>Data Warehousing</i> (DW) e <i>Business Intelligence</i> (BI)	0%	1) Melhora na performance e eficiência dos relatórios BI	
8	Gerenciamento de Documentação e Conteúdo	100%	1) Aumentar a produtividade e agilidade no trabalho com dados	X
			2) Taxonomia, padronização, armazenamento e utilização dos dados bem definidos e estruturados	X
9	Gerenciamento de Metadados	0%	1) Dados interpretados de maneira eficaz e eficiente uso das informações	
			2) Melhoria na tomada de decisões estratégicas	
10	Gerenciamento da Qualidade dos Dados	100%	1) Dados com Acurácia, Atualidade, Consistência, Completude e Integridade	
			2) Diminuição de custos relacionados a duplicidade de dados	
			3) Aumento da satisfação do consumidor dos dados	X
				Dados Inteligíveis
11	<i>Compliance</i> e Auditoria de Dados	50%	1) Dados aderentes a normas legais e regulamentados	X
			2) Redução dos riscos de multas e ações judiciais devido a violação dos dados	X
			3) Controle de Fraudes de dados	X

*Todos os benefícios atingidos, com exceção para os itens automotivos.

Fonte: próprio autor

Pode-se observar que segundo o Analista de Dados foram implantados oito processos dos 11 listados, desde apenas **Compliance e Auditoria de Dados** não atingiu 100%. **Compliance e Auditoria de Dados** atingiu 50% da implantação, isso porque foram implantados todos os procedimentos para validações e verificações da aderência as normas legais, entretanto não foram implantados os procedimentos de auditoria de dados, de maneira a prevenir e atender as normas de *compliance*.

Dentre os processos com 0%, o entrevistado relata que o processo **Gerenciamento de Operações Database** e **Gerenciamento de Metadados** estavam fora do escopo da implantação, o processo **Gerenciamento de Data Warehousing (DW) e Business Intelligence (BI)** já estava implantado na organização e a consultoria proporcionou apenas alguns exemplos de relatórios de qualidade que foram incorporados a política de GD.

Pode-se observar que para o processo **Políticas e Padrões de Dados** todos os benefícios foram atingidos e um novo benefício foi alcançado: “Agilidade na utilização dos dados, principalmente na área de suprimentos”. Para o processo **Desenvolvimento de Dados** todos os benefícios foram atingidos e um novo benefício foi alcançado: “Possível redução de custos devido à padronização dos dados”. Para o processo **Gerenciamento de Segurança dos Dados** todos os benefícios foram alcançados, entretanto para os itens automotivos, os benefícios em questão não são aplicáveis, isto porque os itens automotivos são tratados como exceção com campos abertos em tabelas separadas dos demais, por isso a não garantia dos benefícios relatados. Os itens automotivos representam aproximadamente

20% da base de materiais da organização. O processo **Gerenciamento da Qualidade dos Dados** atingiu apenas um benefício dos citados: “Aumento da satisfação do consumidor dos dados”, e um novo benefício foi atingido: “Dados Inteligíveis”. O benefício “Dados com Acurácia, Atualidade, Consistência, Completude e Integridade” e “Diminuição de custos relacionados à duplicidade de dados” não foram atingidos porque segundo o entrevistado os dados foram saneados com base nas informações disponibilizadas pela organização e muitas destas não estão completas e consistentes. Também com relação aos dados automotivos, esses benefícios não são aplicáveis, portanto eles não foram atingidos. Para o processo **Compliance e Auditoria de Dados** todos os benefícios foram atingidos, entretanto a garantia de *compliance* e regulamentações as normas legais são garantidas apenas para os dados inseridos no SAP após a implantação, dados inseridos anteriormente não possuem essa validação. Para os demais processos implantados todos os benefícios foram alcançados.

Como motivação para a implantação da GD consiste na falta de padronização existente nos dados mestres da organização, o que levou a organização a contratar a consultoria para implantar a GD. A escolha dos processos implantados não ocorreu de forma arbitrária, primeiramente iniciou-se a higienização e padronização dos dados e uma política de GD foi desenvolvida, baseada nos processos implantados, a fim de garantir que os dados mantenham-se higienizados e padronizados. Primeiramente foi realizada a extração inicial dos dados do sistema e iniciou-se uma higienização e padronização dos dados a quatro mãos com a organização. Em paralelo iniciou-se o desenvolvimento da política de GD baseada nos processos implantados, após a conclusão destas etapas ocorreu a tratativa dos dados que foram cadastrados após a extração inicial, os quais foram chamados de delta.

Como principais dificuldade e limitações encontradas durante o projeto de implantação da GD consistem em realizar as adequações das metodologias de padronização e higienização a necessidade da organização e obter a qualidade de dados esperada pela organização, a qual muitas vezes não conseguia explicitar essa qualidade. Como lições aprendidas para o projeto consistem no aprendizado do desenvolvimento de uma política de GD para uma organização de grande porte, aperfeiçoamento da metodologia de padronização e higienização para grandes quantidades de dados e implantação de uma política de GD eficiente e eficaz para que a mesma seja utilizada de maneira ágil no dia a dia da organização.

4.3 Comparativo entre as Entrevistas

Pode-se observar a existência de algumas semelhanças e divergências entre a entrevista do gerente de dados com a entrevista do analista de dados. Como semelhança pode-se observar que ambos identificaram um benefício de agilidade e otimização para a área de suprimentos com a implantação do processo **Políticas e Padrões de Dados**. Ambos também confirmaram a implantação de seis processos, porém com algumas porcentagens de implantação diferente, entretanto foram justificadas por ambos e como o projeto está em fase final de implantação e revisão pelo gerente, dois dos processos que o gerente afirmou não estar em 100% de implantação chegará nessa porcentagem até o final da implantação.

Como principais diferenças podem-se citar dois processos que segundo a visão do analista foram implantados: **1)Desenvolvimento de Dados** e **2)Gerenciamento de Segurança dos Dados**, ambos com 100%. Para o Gerente o processo um não possibilita benefícios na prática, devido à dificuldade de customização do SAP, entretanto segundo o Analista este processo foi implantado principalmente devido a realização do *Workflow* de cadastro de

materiais que foi desenvolvido no SAP. Além disso esse processo atingiu um novo benefício, segundo a visão do Analista, “Possível redução de custos devido a padronização dos dados”. Em relação ao processo dois, segundo a visão do Gerente o mesmo não foi realizado porque o escopo de implantação era apenas materiais, o qual não possui a necessidade de implantação de uma gestão de segurança de dados. Já segundo a visão do Analista, esse processo foi implantado, com normas, procedimento e políticas para a segurança dos dados, com exceção dos itens automotivos que possuem campos abertos e não estão 100% seguros.

Estas são as principais divergências encontradas entre as entrevistas, entretanto as semelhanças são muitas, principalmente em relação aos benefícios atingidos, as dificuldades encontradas e as lições aprendidas.

4.4 Dados da Organização

A organização estudada contratou a consultoria para a implantação da GD para os dados mestres de materiais. A Tabela 3 demonstra as informações sobre a organização estudada.

Tabela 3 – Informações sobre a organização estudada

Nº	Organização	
1	Número de Funcionário	De 5.000 a 10.000 Funcionários
2	Faturamento Médio	Acima de 60 milhões
3	Ramo de Atuação	Agricultura, pecuária, produção florestal, pesca e aquicultura
4	Produto(s) Comercializado(s)	Açúcar, Etanol e Energia Elétrica
5	Colaboradores da área de TI	De 20 a 30 funcionários
6	Patrocinador do projeto	Superintendência de Suprimentos
7	Áreas beneficiadas/impactadas	Todas as áreas de produção, Suprimentos, Contabilidade, Fiscal, TI e Inteligência de Mercado
8	Principais departamentos envolvidos	Engenharias, Fiscal, Suprimentos e TI
9	Abrangência do projeto	Matriz e demais filiais

Fonte: próprio autor

O estudo contempla uma organização de grande porte, pode-se observar pelo faturamento médio e número de funcionários, do ramo da agricultura, como foco na produção de açúcar, etanol e energia elétrica. Possui uma área de TI de médio porte e como patrocinador do projeto de GD a área de superintendência de Suprimentos. O projeto proporcionou benefícios para diversas áreas, tais como, produção, suprimentos, contabilidade, fiscal, TI e inteligência de mercado. Os *stakeholders* do projeto foram às áreas de engenharias, fiscal, suprimentos e TI e o projeto teve abrangência na matriz e em sete filiais.

4.5 Dados da Consultoria

Foi contratada uma consultoria especializada para a implantação da GD na organização. Os dados relacionados à consultoria, bem como o projeto de implantação realizado pela mesma, podem ser observados na Tabela 4.

Tabela 4 – Informações sobre a consultoria

Nº	Consultoria	
1	Período do projeto	05/2013 - 02/2015
2	Equipe que implantou o projeto	De 3 a 6 recursos
3	Nível dos recursos envolvidos	10% - Executivos (Sócios, Diretores e Gerentes) 20% - Sêniores 70% - Juniores (Consultores e Trainees)
4	Custo do projeto	Acima 2 milhões
5	Escopo contemplado no projeto	Políticas e Padrões de Dados

N°	Consultoria
	Gerenciamento da Arquitetura de Dados Desenvolvimento de Dados * Gerenciamento de Segurança dos Dados * Gerenciamento de Dados Mestres e Referência Gerenciamento de Documentação e Conteúdo Gerenciamento da Qualidade dos Dados Compliance e Auditoria de Dados
6	Metodologia de gestão de projetos PMI e Scrum
7	Responsabilidade da gestão do projeto Consultoria
8	Metodologia/Framework de implantação da GD DAMA

* Processos implantados apenas segundo a visão do Analista

Fonte: próprio autor

Como pode-se observar o projeto teve duração de aproximadamente dois anos, com finalização prevista para fevereiro de 2015, o qual contou com uma equipe de três a seis recursos que eram constituídos por: 10% de executivos, 20% de sêniores e 70% de juniores. A ordem de grandeza do projeto foi de acima de dois milhões de reais, o qual possuiu o seguinte escopo de implantação: (1) Políticas e Padrões de Dados; (2) Gerenciamento da Arquitetura de Dados; (3) Desenvolvimento de Dados; (4) Gerenciamento de Segurança dos Dados; (5) Gerenciamento de Dados Mestres e Referência; (6) Gerenciamento de Documentação e Conteúdo; (7) Gerenciamento da Qualidade dos Dados; (8) Compliance e Auditoria de Dados. Os processos três e quatro foram implantados apenas segundo a visão do Analista. Como metodologias de gestão de projeto foram utilizados o PMI e o Scrum, com responsabilidade da consultoria, e como metodologia/framework para a implantação foi utilizado o DAMA DMBOK.

5 CONSIDERAÇÕES FINAIS

A GD é um modelo de governança recentemente difundido na literatura e com poucos casos de implantação, porém vem ganhando importância no meio acadêmico e nas organizações. Organizações desejam possuir dados de qualidade que proporcionem tomadas de decisões eficazes e eficientes e que gerem valor, entretanto buscar por processos que se adequem a realidade brasileira e que gerem esses benefícios é o grande desafio encontrado pelas organizações na atualidade.

Este trabalho identificou os benefícios obtidos por uma organização brasileira que implantou a GD e comparou estes benefícios com os relatados na literatura, bem como identificou dificuldades e lições aprendidas na implantação dos processos de GD. Pode-se observar que dos processos relatados na literatura entre 50% e 70% foram realmente implantados, e estes muitas vezes não atingiram o 100% de implantação.

Em relação aos benefícios atingidos, pode-se observar que existem alguns benefícios que não foram relatados na literatura, entretanto a grande maioria dos relatados foram atingidos com a implantação. Outro ponto interessante é a necessidade da maturidade da implantação da GD, ou seja, realizar uma nova análise dos benefícios atingidos após seis meses ou até um ano de implantação, coletando a percepção dos principais afetados por esses benefícios, que são as áreas da organização relatadas nesse estudo.

Outro ponto importante a se ressaltar são as dificuldades encontradas e lições aprendidas, entre as dificuldades o SAP foi um grande inibidor o que muitas vezes dificulta a aplicação da teoria na prática, inviabilizando a implantação de alguns processos. Também as lições aprendidas proporcionam um entendimento sobre as diferentes percepções existentes em

uma organização, bem como os diferentes níveis de exigência e pontos de vistas com relação à qualidade de um dado.

Como trabalhos futuros serão realizados estudos de múltiplos casos para verificar e validar o instrumento desenvolvido, bem como realizar a triangulação de informações entre diferentes organizações que implantaram a GD, coletando percepções de diferentes pontos de vistas e focos de atuação.

REFERÊNCIA BIBLIOGRÁFICA

Begg, Carolyn & Caira, Tom. (2012). Exploring the SME Quandary: Data Governance in Practice in the Small to Medium-Sized Enterprise Sector. *Electronic Journal Information Systems Evaluation*, Paisley, 15(1), p. 3-13.

Bruening, Paula J & Waterman, Krasnow K. (2010). Data tagging for new information governance models. *Security & Privacy, IEEE*, 8(5), p. 64–68, set/out.

Chapple, Mike. (2013). Speaking the same language: Building a data governance program for institutional impact. *Educause Review*, 48(6), p.14-27, nov/dez. 2013.

Cheong, Lai Kuan & Chang, Vanessa. (2007). The need for data governance: a case study. 2007, In: *18th Australasian Conference on Information System*, Toowoomba.

Coleman, David. W., Hughes, Allen A. & Perry, Wayne D. (2009). The role of data governance to relieve information sharing impairments in the federal government. *Computer Science and Information Engineering, WRI World Congress IEEE*, 4, p. 267-271.

DAMA International. (2012). *DAMA: Guia para o Corpo de Conhecimento em Gerenciamento de Dados*. Technics Publications, LLC, NJ, USA.

Dataflux Corporation. (2007). The Data Governance Maturity Model Establishing the People, Policies and Technology That Manage Enterprise Data. Disponível em <<http://www.sas.com/offices/NA/canada/lp/DIDQ/DataFlux.pdf>>. Acesso em: 10 out. 2014.

Freitas, Patrícia Alves *et al.* (2013a). Aspects of data quality that cause impact on business intelligence systems. *Computational Science and Engineering (CSE)*, IEEE 16th International Conference, p. 76–80.

Freitas, Patrícia Alves *et al.* (2013b). Information governance, *Big Data* and data quality. *Computational Science and Engineering (CSE)*, IEEE 16th International Conference, p. 1142–1143.

Fernandes, Aguinaldo Aragon & Abreu, Vladimir Ferraz. *Implantando a Governança de TI, da estratégia à Gestão de Processos e Serviços*. 3ª Edição, São Paulo: Brasport Livros e Multimídia Ltda., 2012, 615p.

Haider, Waqar & Haider, Abrar. (2013). Governance structures for engineering and infrastructure asset management. *Technology Management in the IT-Driven Services (PICMET)*, IEEE Proceedings of PICMET'13, p. 1229–1238.

Khatri, Vijay & Brown, Carol V. (2010). Designing data governance. *Communications of the ACM*, 53(1), p. 148–152, jan.

Leal, Ricardo Pereira Câmara& Saito, Ricardo. (2003). *Finanças corporativas no Brasil. ERA-eletrônica*, São Paulo, 2(2), p. 2-15, jul/dez.

Mosley, Mark *et al.*(2009). *The DAMA Guide to the Data Management Body of Knowledge*, 1st edition, USA, Technics Publications, LLC, 406 p.

Newman, David& Logan, Debra. (2008). *Gartner Introduces the EIM Maturity Model*. Gartner, ID Number: G00160425, USA.

Olama, Mohammed M.*et al.*(2014).A qualitative readiness-requirements assessment model for enterprise big-data infrastructure investment. In: *SPIE Sensing Technology Applications*.

Otto, Boris.(2011a).*A morphology of the organisation of data governance*.University of St. Gallen, Institute of Information Management, Müller-Friedberg.

Otto, Boris. (2011b). Organizing data governance: findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems*, 29(1), p.45–66, jan.

Riekstin, Ana Carolina.(2012).*Modelo de Governança de Tecnologia da Informação de Escritório ao chão de fábrica*. São Paulo: USP, 209p. Tese (Mestrado) – Programa de Pós-Graduação em Engenharia, Escola politécnica, Universidade de São Paulo, São Paulo.

Sampieri, Roberto Hernandez, Collado, Carlos Fernández& Lucio, María del Pilar Baptista.(2006).*Metodologia de Pesquisa*, 3^a edição. São Paulo: McGraw-Hill, 583p.

Soares, Sunil.(2012).*A Framework that Focuses on the “Data” in Big Data Governance*.Jun. 2012, Disponível em < <http://ibmdatamag.com/2012/06/a-Framework-that-focuses-on-the-data-in-big-data-governance/>>. Acessoem: 10 out. 2014.

Soares, Sunil, Deutsch, Tom, Hanna, Stewart& Malik, Piyush. (2012).*Big Data Governance: aFramework to Assess Maturity*. Disponível em < <http://ibmdatamag.com/2012/04/big-data-governance-a-Framework-to-assess-maturity/>>. Acessoem: 10 out. 2014.

Tallon, Paul P. (2013). Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost. *Computer, IEEE*, 46(6), p.32–38.

Weber, Kristin, Otto, Boris&Osterle, Hubert. (2009). One size does not fit all—a contingency approach to data governance. *Journal of Data and Information Quality (JDIQ)*, 1(1), jun.

Yin, Robert K. (2009).*Case Study Research: Design and Methods*, 4th edition. California: SAGE Publications, Inc., 217p.