

Governança de Dados em Organizações Brasileiras

Alternate Title: Data Governance in Brazilian Organizations

André Montoia Barata
Universidade de São Paulo
Rua Arlindo Bétio, Nº 1.000
Ermelino Matarazzo - SP
andre.barata@usp.br

Edmir Parada Vasques Prado
Universidade de São Paulo
Rua Arlindo Bétio, Nº 1.000
Ermelino Matarazzo - SP
eprado@usp.br

RESUMO

Organizações estão cada vez mais buscando dados íntegros e com qualidade para auxiliarem na tomada de decisões estratégicas e geração de valor. Neste contexto a Governança de Dados (GD) proporciona processos e práticas que auxiliam no gerenciamento e manutenção dos dados. Muitos são os *frameworks* para a implantação de processos de GD e os benefícios que estes podem proporcionar, entretanto poucas são as implantações relatadas na literatura. Este trabalho tem por objetivo identificar os processos e *frameworks* de GD implantados em organizações brasileiras e comparar os benefícios obtidos na implantação com os propostos pela literatura. Para isso serão realizados estudos de casos em organizações brasileiras que implantaram ou estão em processo de implantação dos *frameworks* de GD.

Palavras-Chave

Governança de Dados, *Frameworks* de Gestão, Sistemas de Informação.

ABSTRACT

Organizations are increasingly looking for data integrity and quality to assist in strategic making decision and value creation. In this context Data Governance (DG) provide processes and practices that assist in the management and maintenance data. There are many frameworks to implementation DG process and benefits they may provide, however there are few implementation reported in the literature. This study aims to identify the DG process and frameworks implemented in Brazilian organizations and compare the benefits in implementation with those proposed by literature. For this will be carried out case studies in Brazilian organizations that implemented or are implementing DG frameworks.

Categories and Subject Descriptors

K.6.1 [MANAGEMENT OF COMPUTING AND INFORMATION SYSTEMS]: Project and People Management – *Management techniques, Strategic information systems planning, Systems analysis and design.*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference SBSI 2015, May 26–29, 2015, Goiânia, GO, Brasil.
Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

General Terms

Management.

Keywords

Data Governance, Management Frameworks, System Information.

1. INTRODUÇÃO

As organizações contemporâneas estão cada vez mais se movendo em direção a Governança de Dados (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 [4,10] 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 brasileiras 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.

2. 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: (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 [10] analisaram os *frameworks* de GD e de avaliação de maturidade de GD; (2) Khatri e Brown [11] desenvolveram um *framework* para a GD baseado em domínios de decisões; (3) Begg e Cairra [1] exploraram o dilema da GD em PME; (4) Otto [14] desenvolveu um modelo morfológico baseado na análise teórica da GD para as organizações; (5) Cheong e Chang [4] pesquisaram sobre o relacionamento entre Governança de TI e GD; (6) Soares, Deutsch, Hanna e Malik [19] descreveram o *framework* da IBM para análise da maturidade da Governança de Big Data; (7) Soares [18] descreveu um *framework* para governança de *Big Data*; (8) Weber, Otto e Osterle [21] relataram um estudo de caso de análise da GD em seis organizações internacionais; (9) Chapple [3] descreveu a importância de possuir uma GD estruturada, com boas práticas e processos de qualidade implantados e em funcionamento; (10) Olama *et al.* [13] descreveram uma metodologia baseada nas melhores práticas do mercado para mensurar o conjunto de dados das organizações.

b) **Trabalhos sobre processos e práticas para Governança de Dados.** Foram encontrados sete artigos de práticas de GD: (1) Felici, Koulouris e Pearson [8] apresentaram um modelo conceitual para a prestação de contas pelas organizações que utilizam serviços em nuvem suportando a GD; (2) Freitas *et al.* [9], analisaram o processo de qualidade dos dados; (3) Tallon [20] analisou práticas e processos de GD que podem auxiliar na geração de valor e minimizar riscos no uso de aplicações de *Big Data*; (4) Bruening e Waterman [2] analisaram o uso de metadados (*Data Tagging*) para auxiliar os *frameworks* e modelos de GD; (5) Coleman, Hughes e Perry [5] analisaram a dificuldade de agências e departamentos de organizações em compartilhar informações e utilizar metadados de maneira eficiente; (6) Rosenbaum [16] descreveu sobre o significado do *Data Steward* e a GD para a área da saúde; (7) Otto [15] realizou um estudo de caso sobre como a GD está estruturada e organizada nas organizações.

3. 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.

3.1 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 Body of Knowledge*) 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 [6].

Segundo Mosley *et al.* [12], 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.

3.2 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 [7].

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 [7]. 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.

3.3 Enterprise Information Management

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: 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.

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 sub geridas, 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.

3.4 Governança Big Data

O *framework* de Governança *Big Data* foi elaborado pela IBM® e possui três dimensões para a governança da *Big Data* [18]:

(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 utilizada 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 [18].

3.5 Maturidade da Governança Dados

O *framework* de Maturidade da Governança Dados também foi elaborado pela IBM® e possui onze categorias. 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 [19]. 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 Data* 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.
- (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 [19].

3.6 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. 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 [10].

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 da lacuna entre a maturidade da GD desejada e aquela na qual a organização se encontra [10].

3.7 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 [11]:

(1) Princípios de Dados. Clarificar o papel dos dados como um ativo. (Funções potenciais e responsabilidades: proprietário dos dados; responsável pela custódia, transporte e armazenamento dos dados); responsável pelo conteúdo associação e regras de negócios dos dados; fornecedores de dados; consumidores de dados; e principal departamento responsável pelas políticas de dados.

(2) Qualidade de Dados. Estabelecer requisitos para a utilização dos dados. Funções potenciais e responsabilidades: proprietário dos dados; 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: proprietário dos dados; 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.

3.8 Framework de Notre Dame

O *framework* de GD da universidade de Notre Dame 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.

Segundo Chapple [3], cada pilar desse *framework* desenvolve as seguintes metas e objetivos:

(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.

4. MÉTODO DE PESQUISA

Esta pesquisa é do tipo exploratório e qualitativo, pois tem a finalidade de obter uma imersão inicial sobre o assunto sem a utilização de medições numéricas [17].

Como estratégia de pesquisa foi utilizada o estudo de caso para avaliar a implantação dos processos de GD. Um estudo de caso pode ser único ou múltiplo, de acordo com a abordagem e o direcionamento utilizado. No contexto desta pesquisa, a estratégia do estudo de caso múltiplo foi utilizada. Baseado nessa estratégia, os estudos foram realizados em duas organizações que implantaram processos ou *frameworks* de GD [22].

Foram selecionadas duas organizações de diferentes ramos de atuação e que implantaram a GD. A organização A é do ramo automobilístico e está implantando processos de GD para estruturar seus dados mestres de clientes, fornecedores, materiais e serviços. A organização B é do ramo portuário e logístico e também implantou processos de GD para estruturar os dados mestres de clientes, fornecedores, materiais e serviços.

Para cada organização foram entrevistadas duas pessoas que estiveram diretamente envolvidas na implantação dos processos de GD: (1) Gerente do projeto da consultoria, responsável pela implantação do *framework* na organização; e (2) Analista de dados da consultoria, que possui conhecimentos em GD e que participou da implantação do *framework*. As entrevistas foram realizadas no segundo semestre de 2014.

A partir dessas entrevistas foi possível comparar diferentes pontos de vista em níveis e organizações distintas, em relação à implantação de um mesmo processo de GD.

5. ANÁLISE DOS RESULTADOS

Foi desenvolvido um instrumento para a realização do estudo de caso e as entrevistas com os envolvidos. O instrumento avalia os 11 processos mais relevantes encontrados na literatura, ou seja, os processos mais utilizados e pesquisados pelos autores, bem

como os benefícios que a implantação desses processos traz. Foram questionados aos entrevistados os processos implantados na organização e a porcentagem de implantação atingida, bem como os benefícios atingidos.

5.1 Processos Implantados

A Tabela 1 apresenta os 11 processos mais relevantes na literatura, bem como a porcentagem de implantação de cada processo nas organizações estudadas, segundo a visão dos entrevistados.

Tabela 1. Processos de GD implantados nas organizações

Nº	Processos de GD da literatura	% de implantação	
		A ^(*)	B ^(*)
1	Políticas e Padrões de Dados	65	75
2	Gerenciamento da Arquitetura de Dados	25	63
3	Desenvolvimento de Dados	7	66
4	Gerenciamento de Operações Database	0	0
5	Gerenciamento de Segurança dos Dados	75	22
6	Gerenciamento de Dados Mestres e Referência	57	100
7	Gerenciamento de Data Warehousing (DW) e Business Intelligence (BI)	15	0
8	Gerenciamento de Documentação e Conteúdo	0	50
9	Gerenciamento de Metadados	0	0
10	Gerenciamento da Qualidade dos Dados	78	69
11	Compliance e Auditoria de Dados	65	0

Legenda: A= empresa do setor automobilístico; B = empresa do setor portuário e logístico

A porcentagem de implantação para cada organização foi calculada através da média entre a porcentagem de implantação do processo na visão do gerente de dados da consultoria com a porcentagem de implantação do processo na visão do analista de dados da consultoria. O máximo aceitável de divergência entre a porcentagem de implantação do gerente com a porcentagem de implantação do analista foi de 25%, nos casos em que a divergência foi superior a 25%, buscou-se uma fonte adicional de informação para se chegar à média.

Pode-se observar que dois processos não foram implantados em nenhuma das duas organizações: Gerenciamento de Operações Database (4); e Gerenciamento de Metadados (9). Segundo os entrevistados esses dois processos estavam fora do escopo do projeto de implantação. O processo 4, responsável por definições de tecnologias e banco de dados, muitas vezes, já está implantado nas organizações ou até existem outros projetos específicos para esse processo. O processo 9 exige um nível de

maturidade em GD muito elevado para as organizações que desejam implanta-lo, condição que as organizações estudadas não se encontravam, pelo contrário, elas não possuíam nenhum processo ou procedimento de GD implantado, ou seja, não existia nível de maturidade em GD.

Observa-se que a organização A implantou oito processos de um total de 11, entretanto nenhum deles obteve 100% de implantação. Se considerar um nível de implantação adequado como 50%, o número de implantados diminui para cinco processos, ou seja, dos processos mais relevantes relatados na literatura, apenas 45% foi efetivamente implantado. A organização B implantou sete processos de um total de 11, entretanto o processo Gerenciamento de Dados Mestres e Referência atingiu 100% de implantação e considerando o mesmo nível de 50% de implantação, o número de implantados diminui para seis processos, ou seja, 55% dos processos foram efetivamente implantados. Comprando as duas organizações a média fica em 50% de implantação dos processos relatados na literatura.

As duas organizações possuem semelhança no escopo de implantação da GD, ou seja, o foco das duas são os dados mestres de clientes, fornecedores e materiais, entretanto a motivação para a implantação da GD foi diferente. A principal motivação da organização A foi evitar a perdas tributárias que são causadas pelo cadastro indevido de dados mestres e possibilitar uma recuperação dos créditos de impostos. Já a principal motivação para a organização B foi reestruturação dos dados mestres que estavam distribuídos em diversos sistemas e a implantação de um novo ERP, o SAP. As duas organizações também se assemelham no tamanho, por serem grande organizações com número de funcionários e área de TI bem estruturada, além disto, o projeto de implantação da GD possui escopo parecido com o tempo de implantação de aproximadamente um ano para as duas organizações.

Em relação às dificuldades e limitações encontradas na implantação da GD, as duas organizações possuem convergência, tais como: dificuldade na realização da mudança cultural das organizações, dificuldade de comunicação entre áreas internas da organização que utilizam de um mesmo dado, grande quantidade de processos, envolvidos no cadastro de dados mestres, falta de alinhamento entre os diferentes *stakeholders* envolvidos no projeto, centralização dos conhecimentos em poucos usuários dificultando a passagem desse conhecimento.

As lições aprendidas nos projetos de ambas às organizações também convergem para: envolvimento de pessoas que são os pontos focais e que possibilitarão melhorias na implantação dos processos, mapeamento minucioso de todos os *stakeholders* envolvidos no projeto, bem como explicitar os benefícios que serão proporcionados após a implantação da GD, detalhar a importância do projeto para toda a organização e divulgá-lo amplamente nos meios de comunicação da organização.

5.2 Benefícios Obtidos

A Tabela 2 apresenta os benefícios obtidos pelas duas organizações, separados pelos 11 processos.

Tabela 2. Benefícios obtidos pelas organizações

Nº	Processos de GD da	% benefícios obtidos
----	--------------------	----------------------

	literatura	A ^(*)	B ^(*)
1	Políticas e Padrões de Dados	100	100
2	Gerenciamento da Arquitetura de Dados	100	100
3	Desenvolvimento de Dados	33	100
4	Gerenciamento de Operações <i>Database</i>	0	0
5	Gerenciamento de Segurança dos Dados	83	100
6	Gerenciamento de Dados Mestres e Referência	100	100
7	Gerenciamento de <i>Data Warehousing</i> (DW) e <i>Business Intelligence</i> (BI)	50	0
8	Gerenciamento de Documentação e Conteúdo	0	67
9	Gerenciamento de Metadados	0	0
10	Gerenciamento da Qualidade dos Dados	83	83
11	<i>Compliance</i> e Auditoria de Dados	83	0

Legenda: A= empresa do setor automobilístico; B = empresa do setor portuário e logístico

Cada processo possui de dois a três benefícios relatados na literatura, os quais foram questionados aos entrevistados a obtenção desses benefícios na implantação do processo. Para o cálculo das porcentagens de benefícios atingidos utilizou-se a seguinte regra:

$$\% \text{ Benefícios obtidos} = [\text{N}^\circ \text{ benefícios atingidos}] / [\text{N}^\circ \text{ de benefícios totais}]$$

Por exemplo, para o processo Gerenciamento de Segurança de Dados existem três benefícios relatados na literatura, para a empresa A um entrevistado confirmou a obtenção de todos os benefícios, já o outro entrevistado confirmou a obtenção apenas de dois objetivos, portanto temos cinco confirmações de benefícios de seis possíveis, ou seja, 83%. Para a empresa B os dois entrevistados confirmaram a obtenção de todos os benefícios propostos pela literatura.

Pode-se observar que existem alguns processos que tiveram 0% de benefícios atingidos, isto porque não ocorreu implantação do processo. Para Gerenciamento de Operações *Database* e Gerenciamento de Metadados não existiu implantação desses processos para nenhuma das duas organizações, portanto não existiram benefícios atingidos. Para Gerenciamento de *Data Warehousing* (DW) e *Business Intelligence* (BI) e *Compliance* e Auditoria de Dados existiu implantação e benefícios relatados apenas na organização A, já para Gerenciamento de Documentação e Conteúdo existiu implantação e benefícios relatados apenas na organização B.

Pode-se observar, e também foi percebido durante as entrevistas, que quanto maior a porcentagem de implantação de um processo, maior é a porcentagem de benefícios atingidos. Na maioria dos processos que tiveram acima de 60% de implantação também

obtiveram 60% ou mais de benefícios atingidos, ou seja, eles são diretamente proporcionais.

Além dos benefícios relatados na literatura e atingidos na implantação dos processos, os entrevistados relataram mais oito benefícios, entre estes três foram relatados para a organização A, quatro foram relatados para a organização B e um foi relatado por ambas às organizações.

6. 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 duas organizações brasileiras que implantaram 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, em média 50% 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.

Outro ponto importante a se ressaltar são as dificuldades encontradas e lições aprendidas. Entre as dificuldades, a mais relevante consiste na falta de alinhamento da visão dos *stakeholders* do projeto, principalmente aqueles de departamentos diferentes que utilizam da mesma informação. Também como a principal lição aprendida para um projeto de implantação da GD consiste no mapeamento minucioso de todas as áreas envolvidas e beneficiadas com o projeto, e explicitar a importância do benefício que a GD irá proporcionar para a organização, bem como a importância da participação de todos os impactados com o projeto.

Como trabalho futuro será realizado mais um estudo de caso em uma terceira organização a fim de verificar e validar o instrumento desenvolvido, bem como realizar a triangulação de informações entre três diferentes organizações que implantaram a GD, coletando percepções de diferentes pontos de vistas e focos de atuação. Também serão entrevistados os envolvidos por parte da organização contratante da consultoria, a fim de captar as percepções e os benefícios que eles alcançaram com a implantação dos processos de GD.

7. AGRADECIMENTOS

Nossos agradecimentos à ACM (em especial à ACM SIGCHI) por nos permitir modificar o modelo que eles desenvolveram.

8. REFERÊNCIAS

- [1] Begg, C. e Caira, T. 2012. Exploring the SME Quandary: Data Governance in Practise in the Small to Medium-Sized Enterprise Sector. *Electronic Journal Information Systems Evaluation*, Paisley. 15, 1, 3-13.
- [2] Bruening, P. J. e Waterman, K. K. 2010. Data tagging for new information governance models. *Security & Privacy*, IEEE. 8, 5(Set/Out. 2010), 64-68.
- [3] Chapple, M. 2013. Speaking the same language: Building a data governance program for institutional impact. *Educause Review*. 48, 6(Nov/Dez. 2013), 14-27.
- [4] Cheong, L. K. e Chang, V. 2007. The need for data governance: a case study. 2007, In: *18th Australasian Conference on Information System*, Toowoomba.
- [5] Coleman, D. W., Hughes, A. A. e Perry, W. 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, 267-271.
- [6] DAMA BRASIL. 2014. Disponível em <<http://www.dama.org.br/>>. Acesso em: 10 nov.
- [7] 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.
- [8] Felici, M., Koulouris, T. e Pearson, S. 2013. *Accountability for Data Governance in Cloud Ecosystems*. IEEE International Conference on Cloud Computing Technology and Science, p. 327-332.
- [9] Freitas, P. A. et al. 2013. Aspects of data quality that cause impact on business intelligence systems. *Computational Science and Engineering (CSE), IEEE 16th International Conference*. 76-80.
- [10] Haider, W. e Haider, A. 2013. Governance structures for engineering and infrastructure asset management. *Technology Management in the IT-Driven Services (PICMET), IEEE Proceedings of PICMET'13*. 1229-1238.
- [11] Khatri, V. e Brown, C. V. 2010. Designing data governance. *Communications of the ACM*. 53, 1(Jan. 2010), 148-152.
- [12] Mosley, M. et al. 2009. *The DAMA Guide to the Data Management Body of Knowledge*, 1ª edição, USA, Technics Publications, LLC, 406 p.
- [13] Olama, M. M. et al. 2014. A qualitative readiness-requirements assessment model for enterprise big-data infrastructure investment. In: *SPIE Sensing Technology Applications*.
- [14] Otto, B. 2011a. *A morphology of the organisation of data governance*. University of St. Gallen, Institute of Information Management, Müller-Friedberg.
- [15] Otto, B. 2011b. Organizing data governance: findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems*, 29, 1(Jan. 2011), 45-66.

- [16] Rosenbaum, S. 2010. Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access. *Health services research*, 45, 5p2, 1442-1455.
- [17] Sampieri, R. H., Collado, C. F. e Lucio, M. P. B. 2006. *Metodologia de Pesquisa*, 3ª edição, São Paulo: McGraw-Hill, 583p.
- [18] Soares, S. 2012. *A Framework that Focuses on the "Data" in Big Data Governance*, Disponível em < <http://ibmdatamag.com/2012/06/a-Framework-that-focuses-on-the-data-in-big-data-governance/>>. Acesso em: 10 out. 2014.
- [19] Soares, S, Deutsch, T., Hanna, S. e Malik, P. 2012. *Big Data Governance: A Framework to Assess Maturity*, Disponível em < <http://ibmdatamag.com/2012/04/big-data-governance-a-Framework-to-assess-maturity/>>. Acesso em: 10 out. 2014.
- [20] Tallon, P. P. 2013. Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost. *Computer, IEEE*. 46, 6(Jun. 2013), 32–38.
- [21] Weber, K., Otto, B. e Osterle, H. 2009. One size does not fit all—a contingency approach to data governance. *Journal of Data and Information Quality (JDIQ)*. 1, 1(Jun. 2009).
- [22] Yin, R. K. (2009). *Case Study Research: Design and Methods*, 4th edition. California: SAGE Publications, Inc., 217p.