Catálogo de práticas para apoio a implantação do processo de gestão e governança de dados

**Uiller Del Prado, Edjalma Queiroz da Silva, Anita Fernandes**

FATEC/SENAI – Faculdade de Tecnologia SENAI Mato Grosso Av. XV de Novembro, 303 - CEP: 78020-301 - Cuiabá / MT

[uiller.prado@gmail.com](mailto:uiller.prado@gmail.com),edjalmaqs@gmail.com,anita.fernandes@univali.br

***Abstract.*** *Data has been driving digital transformation. However, to implement the digital transformation, organizations need to be able as well as being willing to access their data. Cultural adherence by its employees is also a key fact. For this purpose, companies need management and data governance, so that all the organization's components can be gathered together collaboratively - strategies, people, processes, policies and technologies. And therefore, this article is presented as a catalogue of practices to support the implementation of the data management and governance process.*

***Resumo.*** *Dados estão potencializando a transformação digital nas organizações. Contudo, para implementar a transformação digital, as organizações precisam tanto serem capazes quanto estarem dispostas a acessar seus dados. A adesão cultural por parte de seus colaboradores também é fator primordial. Por tanto, será preciso gestão e governança de dados para, de maneira colaborativa, unir todos os componentes da organização – estratégias, pessoas, processos, políticas e tecnologias. Nesses termos, o presente artigo apresenta-se como um catálogo de práticas para apoio à implantação do processo de gestão e governança de dados.*

# 1. Introdução

Dados sempre foram importantes e presentes no apoio à tomada de decisão, mesmo antes do formato digital. Na última década eles, os dados, estão sendo denominados de “novo petróleo” [Vanian 2016].

O volume de dados tem crescido em velocidade recorde. As organizações lutam para aproveitá-los ao máximo, armazenando todo tipo de dado valioso comumente gerado em suas operações de negócios. Entra em cena um termo bastante em voga denominado Big Data, que pode ser definido como área do conhecimento que estuda como tratar, analisar e obter informações a partir de conjuntos de dados grandes demais para serem analisados por sistemas tradicionais [Davenport 2014].

Quase não há dúvidas sobre o valor real e tangível que o *Big Data* representa para as organizações. Ao todo, 84% das empresas acreditam que as análises de *Big Data* estão mudando a forma como elas encaram a concorrência [Columbus 2014]. Empresas como Uber [Marr 2015] e Airbnb [Huet 2015] transformaram setores inteiros com operações mais inteligentes impulsionadas por *Big Data*.

Não deveria ser uma surpresa a conclusão de que a chave para o sucesso do *Big Data* é o gerenciamento inteligente de dados. Porém, em meio a todo o entusiasmo em relação às novas tecnologias e às novas fontes de dados, é comum que as equipes de Tecnologia da Informação (TI) ajam apressadamente, disponibilizando dados sem dispor de uma forma eficiente para promover a gestão e governança desses dados [Informática 2019]. Pior do que não ter dados é ter dados de má qualidade. Qualquer organização que espera prosperar com *Big Data*, ou qualquer arquitetura em que seus dados estejam formatados, precisa ter confiança em fornecer conjuntos de dados seguros de maneira consistente.

Gestão e Governança de Dados (GGD) é a disciplina responsável por definir, planejar, implantar e executar estratégias, procedimentos e práticas necessárias para gerenciar de forma efetiva os recursos de dados e informações das organizações, incluindo planos para sua definição, padronização, estruturação, proteção e utilização [Rêgo 2013].

Devido à rapidez da evolução das tecnologias e às inovações ligadas a elas, passa-se a existir escassez de mão-de-obra especializada, vez que nem todas as pessoas chegam ao mercado de trabalho com o conhecimento adequado para suprir a demanda [Da Costa and Martins 2016]. Não são raros comportamentos onde organizações terminam optando em formar seu próprio quadro de especialistas. Com GGD não é diferente.

Nos presentes termos, este trabalho objetivou catalogar práticas capazes de promover suporte à implantação do processo de GGD nas organizações.

# 2. Revisão de Literatura

﻿Hoje a área de GGD está intimamente ligada à área de *Big Data* e também tem contribuição no âmbito de *Analytics* o que*,* evidentemente, passa por *Business Intelligence* (BI). O quadro abaixo elucida como Davenport [2013] faz uma rápida leitura, histórica, holística, do comportamento do dado no decorrer do tempo.

| Analytics | Momento | Principais características |
| --- | --- | --- |
| 1.0 | Era do BI (~1950 a ~2005) | Pela primeira vez dados produzidos em operações de negócio (vendas, produção, etc.) foram agregados e analisados. Era marcada tanto pela *Enterprise Data Warehouse* capturando informações quanto por *softwares* de BI reportando dados. As análises além de demoradas se focavam em explicar o passo e não em predições do futuro. |
| 2.0 | Era do *Big Data* (~2005) | Era marcada pela presença da internet fomentando a geração de grandes volumes de dados (Google, eBay, iniciativas públicas, sensores, etc.). Organizações passam a demandar tecnologias mais modernas e novas competências profissionais. |
| 3.0 | Era da oferta de dados enriquecidos | Dados, provenientes de fontes internas e externas, estruturados e não estruturados, passam a ter novas finalidades além de sustentar a tomada de decisão. Agora fomentam a criação de novos produtos e serviços. |

Tabela 1 – Era dos dados e tecnologias atreladas

Na fase 3.0 da linha do tempo proposta acima, mais propriamente nos últimos 10 anos, existe um fato implícito denominado *Self Service BI* ou S*elf Service Analytics* onde, diferentemente das fases 1.0 e 2.0 quando a criação das ferramentas de análise de dados estavam sob a alçada e controle das equipes de TI, e os dados se comportando em formatos de silos (isolados em conjuntos distintos em áreas de negócios distintas), agora, na versão 3.0, a criação de relatórios, painéis de BI e tantos outros recursos de análises de dados passam a ocorrer dentro das áreas de negócio, pelos próprios usuários de negócio. É justamente neste ponto onde o exercício da GGD passa a ter significado fundamental, com vistas a qualidade dos dados, sua correta significância no contexto da organização, zelando por aspectos de segurança, e o mais importante, garantindo que os mesmos silos de antigamente não sejam novamente gerados, contudo agora apenas dentro de tecnologias mais modernas.

Os objetivos da GGD [Tallon et. al. 2014 and Webber et. al. 2009] consistem em: definir quais dados são essenciais para a organização; definir como a organização pode fazer uso dos dados; definir quem será responsável por fazer uso dos dados; estruturar e gerenciar dados como ativos; maximizar o valor do dado para a organização, aumentando sua confiabilidade; proteger a informação de modo que seu valor para a organização não seja diminuído pelo uso da tecnologia, pelos erros humanos ou pela perda de acesso no decorrer do tempo; promover o equilíbrio entre os riscos e os benefícios da retenção dos dados e; promover a segurança da organização.

A GGD opera sobre princípios de integridade, transparência e responsabilidade com relação aos dados: diminuindo conflitos operacionais; encorajando adoção de novas práticas comuns; reduzindo custos; aumentando a efetividade através da coordenação de esforços.

Rêgo [2013] argumenta que atualmente existem diversas fontes com frameworks e modelos de maturidade disponíveis no mercado. Destaca que entre as principais estão o ﻿guia DAMA-DMBOK®, os *frameworks* do *Data Governance Institute* da IBM®, além do modelo de maturidade proposto por Tony Fisher da DataFlux®.

# 2.1. DAMA-DMBOK

*Data Management Association* (DAMA) é uma organização internacional sem fins lucrativos, que dita os princípios, melhorias e melhores práticas quanto a GGD. Seu maior produto é o DMBOK®, que é uma coleção de processos, princípios e conhecimento sobre o gerenciamento de dados, desenvolvido com a participação de mais de 120 profissionais do mundo inteiro [Rêgo 2013].

Segundo Mosley [et. al. 2009], existem dez domínios para o gerenciamento de dados:

* Governança de Dados: está no centro e concerne ao planejamento, controle e supervisão no gerenciamento e uso dos dados. É o exercício da autoridade e controle sobre a gestão de ativos de dados.
* Gestão da Arquitetura de Dados: consiste na definição da estrutura ou diagrama de arquitetura organizacional dos dados.
* Desenvolvimento de Dados: consiste em analisar, desenhar, implementar e testar soluções que atendam às necessidades de dados na organização.
* Gestão de Operações de Dados: consiste em planejar, controlar e suportar os dados durante todo seu ciclo de vida, desde a aquisição ou criação até o arquivamento ou eliminação definitiva.
* Gestão de Segurança dos Dados: consiste em planejar, desenvolver, e executar políticas e procedimentos de segurança que promovam a privacidade, confidencialidade e acesso apropriado aos dados.
* Gestão de Dados Mestres e Referência: consiste em planejar, implementar e controlar as versões de dados originais, qualificados e replicados em ambientes distribuídos.
* Gestão de *Data Warehousing* e *Business Intelligence*: consiste em planejar, implementar e controlar processos que promovam suporte às tomadas de decisões, sob várias dimensões de análise.
* Gestão de Conteúdo de Documentos: consiste em planejar, implementar, e controlar atividades para armazenar, proteger e acessar dados estruturados e não estruturados.
* Gestão de Metadados: consiste em planejar, implementar e controlar atividades para garantir integração, controle e entrega de metadados (propriedades dos dados) sobre a arquitetura de dados e informações.
* Gestão da Qualidade dos Dados: consiste em planejar, implementar e controlar a melhoria da qualidade de dados.

# 2.2. DGI Framework

O *framework* de maturidade em GGD elaborado pelo Conselho de Governança de Dados da IBM® permite classificar, organizar e comunicar as atividades envolvidas na tomada de decisão relativa aos dados. Possui onze domínios e para cada uma delas existe um bloco de questões que possibilitam a análise da maturidade de GGD da organização, baseada na categoria que está sendo analisada [Soares 2012]. As principais questões por domínios podem ser observadas a seguir.

* Conscientização Organizacional: reconhecimento da responsabilidade para governar dados em diferentes níveis de gestão.
* Administração e Gestão: disciplina de controle de qualidade designada para assegurar os cuidados de detenção de dados, valorização de ativos de dados, minimização de riscos e fomento do controle organizacional.
* Políticas: estabelecimento de fluxos de trabalho, padrões e políticas pautadas no comportamento organizacional desejado.
* Criação de Valor: processo no qual os dados são qualificados e quantificados permitindo que o negócio maximize o valor criado por eles.
* Gestão de risco de dados: metodologia através da qual os riscos dos dados são identificados, qualificados, quantificados, evitados, aceitos e minimizados.
* Segurança, Privacidade e Conformidade: descreve as políticas, práticas e controles usados pela organização para minimizar os riscos e proteger os dados.
* Arquitetura de dados: desenho da arquitetura de sistemas de dados estruturados e não estruturados que permita a disponibilidade e distribuição aos utilizadores de forma apropriada.
* Gestão da qualidade de dados: métodos para medir, melhorar e certificar a qualidade e integridade da produção, teste, uso e arquivamento dos dados.
* Gestão de Metadados: métodos e ferramentas usados para criar definições semânticas comuns para os termos do negócio, modelos, tipos e repositórios de dados.
* Gestão do ciclo de vida da informação: abordagem sistêmica para a coleta da informação, uso, retenção e eliminação.
* Auditoria e Relatórios: processos organizacionais para monitorizar e medir o valor dos dados, riscos e eficácia da Gestão e Governança.

# 2.3. Data Governance Maturity Model

O *framework Data Governance Maturity Model* consiste em um modelo para análise da maturidade da GGD nas organizações, auxiliando-as a entender o nível de gerenciamento de dados atual e proporcionar um caminho para o crescimento 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 seria a melhor fase para o negócio, como e quando elas devem avançar para a próxima fase [DadataFlux® 2007].

Para cada fase do *Data Governance Maturity Model*, os seguintes domínios devem ser definidos e estruturados:

* Pessoas: envolvidos e suas contribuições.
* Políticas: atividades e regras de negócios para atingir a GGD.
* Tecnologias: investimentos e tecnologia necessários.
* Riscos e Recompensas: riscos que a organização encontra na fase atual e ganhos obtidos com o avanço para a próxima fase.
* Avançando para a próxima fase: ações necessárias para o avanço à próxima fase.

# 2.4. Catálogo de práticas para apoio a implantação do processo de Gestão e Governança de Dados

Abaixo cataloga-se cinco práticas. Para cada uma delas sugere-se um conjunto de definições que ao serem especificadas passam a oportunizar a documentação e regulação do arcabouço de GGD [Cochrane 2009].

1. Comece com um pequeno grupo de trabalho que represente múltiplas perspectivas de negócio. Menos é mais neste primeiro momento. Como na maioria das iniciativas de relevância, muitas vezes determina-se estratégias e objetivos durante a construção dos casos de negócios para, somente após, obter patrocínio. Deste modo é oportuno definir um pequeno grupo de trabalho, que represente múltiplas áreas da organização, para descrever os principais custos-benefícios do programa de GGD como um todo.
2. Desenvolva um *framework* operacional, composto por definições acerca de estratégias, pessoas, processos, políticas e tecnologias, definindo claramente como esses contextos se relacionam:
   * Estratégias. É neste tópico que a elaboração das estratégias acontecem. São oportunas definições claras acerca da missão, objetivos, escopo da primeira iteração ou projeto piloto.
   * Pessoas. Uma eficaz execução dos planos de governança estabelecidos pelas Estratégias depende de um conjunto certo de pessoas. É neste tópico que se idealizam o Conselho de Governança de Dados, o Comitê de Gestão de Dados (Gestor, Diretor, Curador), define-se os *Stakeholders* de alto nível (Donos ou Patrocinadores), esboça-se o Escritório de Gestão e Governança de Dados.
   * Processos. É neste tópico que os principais processos de trabalho são estabelecidos. Define-se os protocolos de comunicação, papéis e responsabilidades, direitos de decisão e controle juntamente com definições acerca da hierarquia da prestação de contas ou a quem se deve reportar.
   * Políticas. Regulamentos de conformidade, controles de qualidade e padronização de dados precisam ser definidos para que deste modo possa ser possível impulsionar a qualidade geral dos dados. Para tanto, métricas de desempenho devem ser elaboradas no sentido de medirem a progressão geral do programa de GGD.
   * Tecnologias. Neste tópico elaboraram-se definições acerca do papel que a tecnologia desempenhará no programa de GGD. Soluções de *Big Data, BI* e *Analytics*, Gestão de Dados Mestres (MDM) entre outras estratégias de implementação de tecnologia com vistas a melhor subsidiar as operações de negócio da organização são vislumbradas neste tópico. Esboça-se portanto quais processos a tecnologias está-se possibilitando, quem usará as ferramentas e como eles as usarão, e como as ferramentas podem ser corretamente utilizadas.
3. Defina uma iniciativa piloto, pequena, simples, de fácil implementação, que agregue rapidamente valor. Embora deva se ter uma visão de longo prazo, com vistas a organização como um todo, nesta primeira interação, é fator determinante de sucesso a definição de área ou escopo reduzido. É oportuno ser uma área da organização onde os colaboradores sejam adeptos a novas iniciativas e que a mudança de cultura não seja vista como empecilho. Onde seja possível agregar valor desde o início do programa. Devido à ênfase nas pessoas e nos processos, se a primeira iteração for muito grande, é comum que participantes percam o interesse no decorrer da implementação.
4. Defina e monitore métricas de desempenho estabelecendo o ritmo de evolução do programa. Assim que a iniciativa piloto estiver em andamento é importante monitorar o sucesso do programa tendo como base as métricas de desempenho especificadas. Isso mostrará ao conselho de governança e às principais partes interessadas onde o programa está funcionando perfeitamente e onde estão as áreas deficitárias oportunas de melhoria. Para estas últimas é importante estabelecer um ritmo de como o programa funcionará a longo prazo.
5. Refina e cresça. Governança de dados deve ser vista como uma competência central, como um programa constante, e não como um projeto com datas de início e término pré-definidas. O planejamento cuidadoso associado às iniciativas menores contribuem para o alcance de objetivos a curto, médio e longo prazo. Constantemente refinar e evoluir as estratégias do programa, eliminando impedimentos e processos interrompidos, conduzirá ao sucesso do programa traduzindo-se em valor para a organização na forma de dados aprimorados e qualidade da informação.



Figura 1 – Exemplo de framework operacional proposto por Cochrane [2009]

Para tanto, um framework de modelos de documentos foi criado. De modo a subsidiar cada um dos 5 passos supracitados.

Prevendo tanto alterações pertinentes a evolução natural dos modelos de documentos quanto a cópia e/ou ramificação por outros profissionais, o *framework* foi disponibilizado em formato distribuído, de código aberto e gratuito de (GIT) e pode ser acessado através do endereço eletrônico <http://bit.ly/gestao-governanca-dados>.

A estrutura deste framework é composta de dois níveis que visam isolar contextos. O primeiro aborda sobre as estratégias de GGD no seu conceito mais amplo, versa acerca dos domínios para o gerenciamento de dados que devem ser oportunizados na organização. O segundo tem por objetivo sustentar a execução de projetos definidos pela estratégia de governança e, por tanto, baseia-se no guia PMBOK®, que é uma publicação do *Project Management Institute* (PMI) e que traz um conjunto de conhecimentos e boas práticas em gerenciamento de projetos. São os documentos que compõem o *framework* de práticas:

| Contexto 1 - Modelo de Documentação de Plano de Gestão e Governaça de dados | |
| --- | --- |
|  | 1 - Estratégia - Definição do grupo de promotores da implantação do processo de gestão e governança de dados.doc |
|  | 2 - Estratégia - Plano de implantação do processo de gestão e governança.doc |
|  | 3 - Pessoas - Designação do conselho de governança de dados.doc |
|  | 4 - Pessoas - Designação do comitê gestor de dados.doc |
|  | 5 - Processos - Definições dos processos de trabalho.doc |
|  | 6 - Políticas - Definições das políticas de dados.doc |
|  | 7 - Tecnologias – Definições acerca da implantação e uso da tecnologias.doc |
| Contexto 2 - Modelo de Documentação de Plano de Projeto | |
|  | 1 - Plano de Gerenciamento de Projeto.doc |
|  | 2 - Termo de Abertura do Projeto.doc |
|  | 3 - Equipe de Planejamento do Projeto.doc |
|  | 4 - Plano de Gerenciamento de Escopo e de Mudanças.doc |
|  | 5 - Declaração de Escopo do Projeto.doc |
|  | 6 - Estrutura Analítica do Projeto.doc |
|  | 7 - Plano de Gerenciamento de Tempo.doc |
|  | 8 - Plano de Gerenciamento de Custos.doc |
|  | 9 - Plano de Gerenciamento da Qualidade.doc |
|  | 10 - Plano de Recursos Humanos.doc |
|  | 11 - Plano de Gerenciamento das Comunicações.doc |
|  | 12 - Plano de Gerenciamento de Riscos.doc |
|  | 13 - Plano de Aquisições.doc |
|  | Compilado de modelos de documentos para Gestão de Projetos.xls |

Tabela 2 – Lista de modelos de documentos do catálogo de práticas

# 3. Método de Pesquisa

O método utilizado neste trabalho é o hipotético-dedutivo, que parte da descrição de um problema, passando para a definição das hipóteses que serão testadas e comprovadas sua veracidade ou não [Prodanov and Freitas 2013]. A utilização das hipóteses se faz necessária pois ajuda na coleta e análise de dados, sendo também consideradas como “tentativa de resposta ao problema de pesquisa” [Barros and Lehfeld 2007].

Do ponto de vista da natureza deste trabalho é considerado como pesquisa aplicada, pois tem como objetivo a produção de conhecimento para a aplicação prática, solucionando problemas específicos [Silva and Meneses 2005, Prodanov and Freitas 2013]. A utilização deste método se justifica, uma vez que, como produto final, foi criado um catálogo de práticas para apoio à implantação do processo de GGD, sendo este validado por especialistas.

A abordagem do problema é classificada como pesquisa qualitativa, pois não utiliza dados estatísticos para sua avaliação, mas sim, analisa dados extraídos do ambiente natural de estudo do pesquisador [Perovano 2016]. Os dados serão coletados por meio de pesquisas bibliográficas, de onde serão avaliadas as informações pertinentes para a construção do catálogo de práticas. Desta forma, sob o ponto de vista dos objetivos, este trabalho se classifica como pesquisa exploratória, por efetuar pesquisas bibliográficas e construir hipóteses.

Como estratégia foram utilizadas duas abordagens distintas. A primeira consistiu na realização de pesquisa aberta em organizações a título de avaliar o grau de maturidade de GGD em organizações que provêem plataformas de BI e *Analytics* nas áreas de negócio. A segunda avaliou com especialistas de TI a relevância do catálogo de práticas.

A primeira fase da pesquisa foi constituída por 10 perguntas objetivas divididas em três contextos diferentes sendo duas perguntas de caráter profissional daquele que responde a pesquisa; duas perguntas acerca da organização onde, na data da pesquisa, o profissional estaria atuando; e, por fim, seis questões específicas objeto desta pesquisa. A segunda fase foi constituída de apenas uma pergunta acerca do grau de relevância do catálogo.

# 3.1 Problemas de Pesquisa

As organizações têm implantado projetos de BI *e Analytics* indistintamente a implantação de GGD? A desassociação desses dois universos teria relação com baixos níveis de maturidade em dados? Um catálogo de prática apoiaria a implantação do processo de GGD nas organizações?

# 3.2. Solução Proposta

No contexto da solução proposta, foi formulado um conjunto de hipóteses. As hipóteses nulas (H0) são aquelas que a pesquisa pretende refutar para que as hipóteses alternativas (HA) possam ser consideradas como aceitas.

H01: Raramente organizações agem apressadamente disponibilizando dados, em plataformas de BI e *Analytics*, sem dispor de uma forma eficiente para promover a gestão e governança desses dados.

HA1: É comum que organizações ajam apressadamente, disponibilizando dados, em plataformas de BI e *Analytics*, sem dispor de uma forma eficiente para promover a gestão e governança desses dados.

H02: Organizações que disponibilizam dados em plataformas de BI e *Analytics* e promovem de forma eficiente a GGD não possuem maiores níveis de maturidade em dados.

HA2: Organizações que disponibilizam dados em plataformas de BI e *Analytics* e promovem de forma eficiente a GGD possuem maiores níveis de maturidade em dados.

H03: As práticas descritas no catálogo apresentado não promovem apoio a implantação do processo de GGD.

HA3: As práticas descritas no catálogo apresentado promovem apoio à implantação do processo de GGD.

# 3. Discussão dos resultados

Participaram da pesquisa 53 distintas organizações presentes majoritariamente no Brasil e algumas no exterior.

Na conjuntura da primeira fase da pesquisa, constituída por perguntas objetivas de contextos profissional e da organização, avalia-se:

No contexto do perfil profissional, que os participantes da pesquisa ocupam funções de alto nível, 40,4% se declararam técnicos ou analistas e 33,3% coordenadores ou gerentes. Destes 33,3 % se declararam com mais de 15 anos (másters) de experiência profissional, parecido de 31,6% entre 10 e 15 anos de profissão (sênior).

|  |  |
| --- | --- |
| Figura 2 – Cargo profissional dos participantes da pesquisa. | Figura 3 – Nível profissional dos participantes da pesquisa. |

Quanto às organizações de atuação desses profissionais, 47,4% enquadram-se como organização de grande porte (de 500 ou mais empregados), seguido por 21,1% de pequeno porte (de 20 a 99 empregados), e 15,8% de médio porte (entre 100 e 499 empregados). Ainda nesse contexto, 63,2% atua no segundo setor ou setor privado e 24,6% atua no primeiro setor ou setor de governo.

|  |  |
| --- | --- |
| Figura 4 – Porte da organização | Figura 5 – Setor de atuação da organização |

No tocante às perguntas específicas de pesquisa, inicialmente fundamentou-se por "*Business Intelligence & Analytics*" conjunto de tecnologias, sistemas e ferramentas de visualização de dados que ajudam as organizações na análise dos dados críticos de negócio, orientando na tomada de decisões e obtendo valor a partir de dados [CHEN and CHIANG and STOREY 2012], de modo a contemplar a utilização de ferramentas de software como: Power BI, Qlik View/Sense, Tableau, ThoughtSpot, Domo, QuickSight, ou similares. Como resposta ao questionamento objetivo no sentido de a organização prover "*Business Intelligence & Analytics*" em apoio às operações de negócio, 77,2% dos entrevistados responderam positivamente.

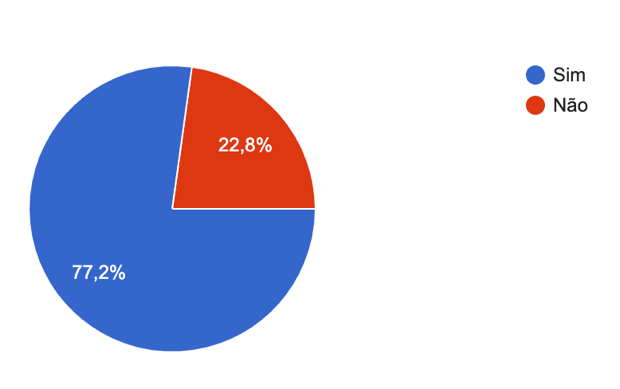


Figura 6 – Volume de organizações que disponibilizam BI e Analytics

Deste ponto em diante, para fins de sustento das hipóteses, as análises da amostra passaram a se ater somente no sentido dos 77,2% das organizações que provêem "*Business Intelligence & Analytics*" em apoio às operações de negócio. A constar passa-se a tratar de um total de 41 organizações. Desta forma, a fatia de 22,8% que não provêem as citadas plataformas passou a ser desconsiderada.

Definiu-se por “Governança de Dados” como o exercício de autoridade e controle (planejamento, monitoramento e engajamento) sobre o gerenciamento de ativos de dados. A função de “Governança de Dados” guia como serão realizadas todas as demais funções da “Gestão de Dados”. “Governança de Dados” é de alto nível, ou seja, é gestão estratégica de dados na esfera executiva, enquanto “Gestão de Dados” atua em níveis táticos e operacionais [Rêgo 2014]. Abrangem definições acerca de: Estratégias (planos de governança, missão); Políticas (regras, padrões, conformidades, métricas de desempenho, qualidade); Pessoas (papéis e responsabilidades, direitos de decisão e controle, conselhos, comitês); Tecnologias (aquisição e implantação de ferramentas, integração, armazenamento e qualificação de dados). Questionando-se, na escala de 1 a 5, onde 1 se denominaria por fracamente implementado e 5 por fortemente implementado, qual o grau de implementação dos preceitos de "Gestão e Governança de Dados" na organização. A título de resposta observou-se que a maioria das organizações se enquadrou nas proporções de grau 4 (12 organizações ou 29,27%), precedido do grau 3 (11 organizações ou 26,83%) e grau 5 (9 organizações 21,85%) predominando elevado nível de GGD declarado.

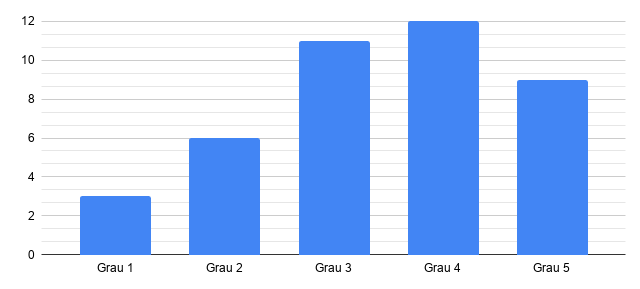


Figura 7 – Níveis de implementação de Gestão e Governança de Dados.

De modo a melhor entender os níveis da GGD questionados acima, vez que Rêgo [2013] sustenta a oportunidade de estruturas formais de apoio a GGD, conforme ilustra a figura abaixo, foram formuladas três novas perguntas.

  
﻿Figura 8 – Estruturas de apoio à Governança de Dados

Denomina-se então por “Conselho de Governança de Dados” grupo de executivos da alta administração como presidente, diretores, seniores e masters que deliberam acerca de dados no nível estratégico. Com base nesta denominação, questionou-se se a organização disporia de um “Conselho de Governança de Dados”. Observa-se que 68,3% das organizações operam sem a presença de um conselho devidamente definido.

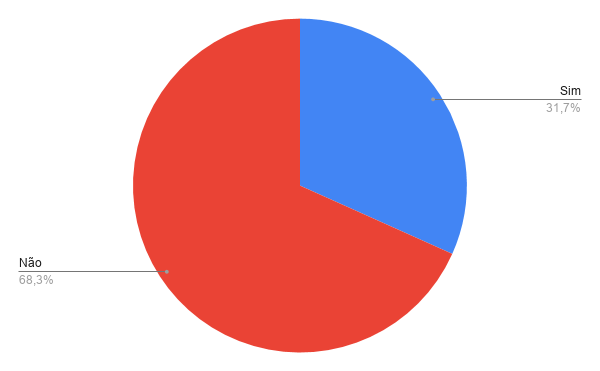


Figura 9 – Percentual de organizações que possuem Conselho de Governança de Dados

Denomina-se por “Comitês de Gestão de Dados” grupo de profissionais que atuam no nível tático da organização sendo responsáveis por tratar novas iniciativas ou ações de melhoria propostas pelo Conselho. Com base nesta denominação, questionou-se se a organização disporia de “Comitês de Gestão de Dados”. Observa-se que 51,2% das organizações operam sem a presença de comitês devidamente definidos.

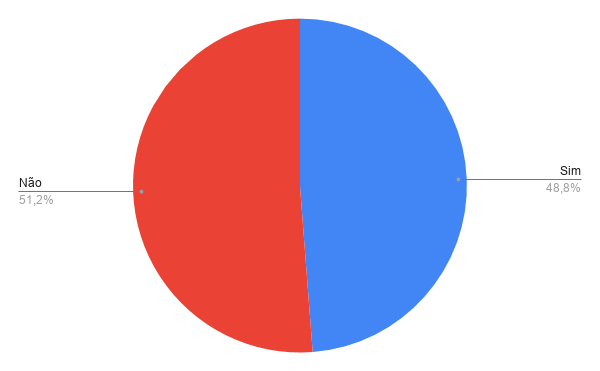


Figura 10 – Percentual de organizações que possuem Comitê de Gestão de Dados

Denomina-se por “Escritório de Gestão de Dados” estrutura formal de apoio constituída majoritariamente por profissionais da área técnica e tem por objetivo a execução de tarefas nos níveis tático e, principalmente, operacional. Com base nesta denominação, questionou-se se a organização disporia de um “Escritório de Gestão de Dados”. Observou-se comportamento inverso àqueles acima elucidados, vez que neste 58,5% das organizações passam a operar com a presença de escritório de gestão de dados devidamente definido.

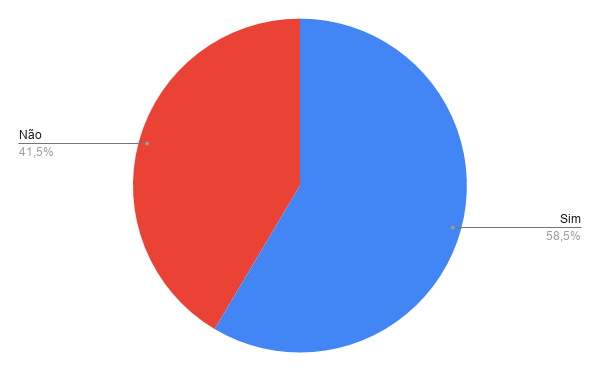


Figura 11 – Percentual de organizações que possuem Escritório de Gestão de Dados

Observa-se que 2/3 das estruturas formais de apoio a GGD denominadas acima recebem índices de existência inferiores a 50%. Evidenciando que quanto mais próximo do nível operacional maior o grau de existência da estrutura. Quanto mais próximo do nível estratégico mais rara ela é.

Por fim, questionou-se acerca do nível de maturidade a partir da perspectiva de dados, para tanto utilizando-se o *Data Maturity Model* (DMM) do Instituto *Capability Maturity Model Integration* (CMMI) vez que este estabelece um conjunto de princípios que permitem às organizações, além de melhorarem suas práticas de gerenciamento de dados por todas as áreas de negócio, avaliar-se sob a ótica da maturidade a partir de uma perspectiva de dados. A constar, os níveis se denominam:

* NÍVEL 1 - INICIAL: a) Não há regras ou procedimentos sobre gerência de dados; múltiplos arquivos e bancos de dados; múltiplos formatos; redundância; mudanças “on the fly”; geralmente não há um grupo de gerência de dados; caso exista, apenas aplicam as mudanças requisitadas de acordo com as necessidades. b) Qualidade de dados depende das habilidades do pessoal de desenvolvimento. Frequentemente, projetos são cancelados ou, pior ainda, produzem dados incorretos e relatórios inválidos.
* NÍVEL 2 - GERENCIADO: a) Planejado e Gerenciável. b) Política de gerência de dados: como e quando as estruturas de dados são criadas, mudadas e gerenciadas. Em geral, a política não é institucionalizada, ficando nas mãos de um grupo ou pessoa, com função de Administrador de Banco de Dados - DBA. c) O sucesso depende das habilidades dos DBAs. Porém, não há muito esforço em documentar e capturar o significado de negócio dos dados. Pouca diferenciação entre modelos lógicos e físicos. Preocupação maior com os aspectos técnicos da gerência de dados.
* NÍVEL 3 - DEFINIDO: a) Política de gerência de dados documentada e estabelecida como um componente central do ciclo de desenvolvimento de aplicações. Há um entendimento do significado de negócio dos dados, através de uma função de Administrador de Dados - DA. Dados são tratados como recursos corporativos. b) O sucesso depende da interação das funções do DA e do DBA, do uso apropriado de ferramentas para criar modelos de dados, automatizar processos e, proativamente, monitorar e sintonizar o desempenho das bases de dados.
* NÍVEL 4 - QUANTITATIVAMENTE GERENCIADO: a) Gerenciamento de metadados permite ao grupo de gerência de dados (Administrador de Dados - DA e Administrador de Banco de Dados - DBA) catalogar e manter metadados de estruturas de dados corporativas. Sabe-se que dados existem onde. O grupo envolve-se em todos os esforços de desenvolvimento (no nível lógico sempre; no nível físico quando necessário). b) O sucesso depende do suporte da alta administração à máxima “dados são recursos corporativos”. Ferramentas avançadas são necessárias para gerenciar metadados, qualidade de dados e as bases de dados.
* NÍVEL 5 - OTIMIZADO: a) Organizações neste nível usam as práticas desenvolvidas nos níveis anteriores para continuamente melhorar o acesso a dados, qualidade de dados e desempenho dos bancos de dados. Nenhuma mudança é introduzida sem primeiro ser escrutinada pelo grupo de gerência de dados e documentada no repositório de metadados. b) Deve haver uma preocupação contínua na melhoria do processo de gerência de dados com uso de ferramentas avançadas.

Ao questionar o nível de maturidade em dados que as organizações se enquadram perante o *Data Maturity Model* (DMM) observou-se que 33,3% se enquadram no nível 1 (inicial), seguido de 33,3% enquadrados no nível 2 (gerenciado) e 15,2% no nível 3 (definido). Apenas 15,2 % se enquadram no nível mais alto (nível 5, otimizado).

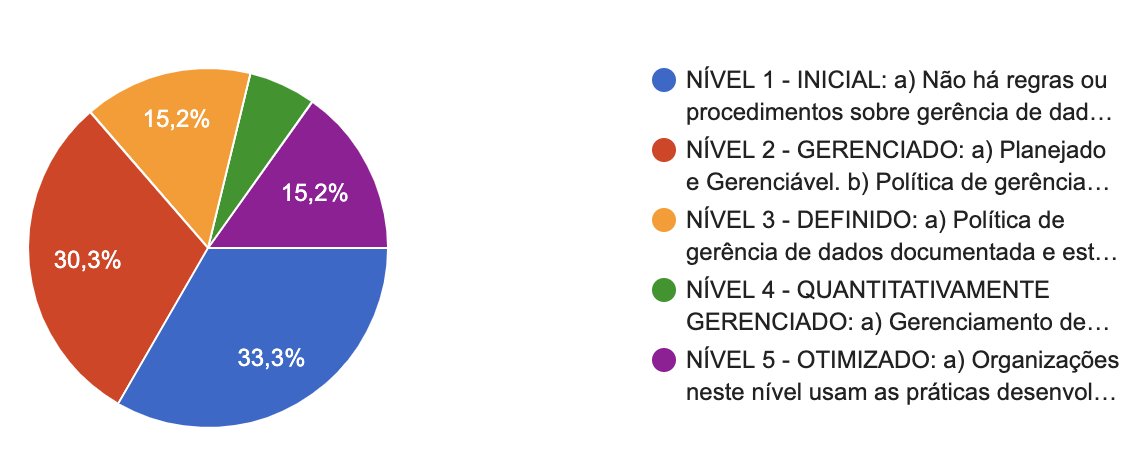


Figura 12 – Nível de maturidade em dados das organizações

Observou-se, ainda, o nível de maturidade comparada. Em outros termos, grau declarado de implementação de GGD em relação ao nível médio de maturidade em dados. Nota-se correlação proporcional entre os dois universos.

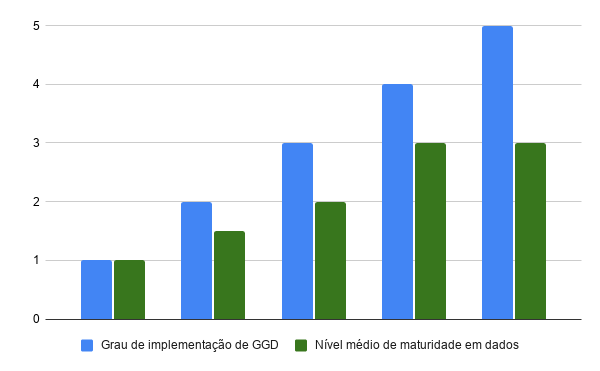


Figura 13 – Nível de maturidade comparada

# 4 Considerações finais

Apresentadas as devidas evidências, em que pese sugere-se refutar as hipnoses em H0, sugere-se que:

* Não raro, organizações têm agido apressadamente disponibilizando dados em plataformas de *BI e Analytics* sem antes dispor de uma forma eficiente para promover a GGD.
* Organizações que disponibilizam dados em plataformas de *BI e Analytics* e, associado a isto, promovem a GGD, desfrutam de maiores níveis de maturidade em dados, o que pressupõe tomadas de decisões mais assertivas.

Não obstante, elucidados os fatores positivos quanto a GGD, este trabalho ateve-se, ainda, em catalogar práticas para apoio à implantação do processo de GGD nas organizações, tendo sua validade submetida ao crivo de duas diferentes empresas dos ramos de consultoria de tecnologias associadas à *Big Data,* BI *e Analytics*, obtendo, em escala de 1 a 5, onde 1 se denominaria por baixo grau de relevância e 5 por alto grau de relevância, da empresa A nota 3 enquanto da empresa B nota 4, resultando em nota média 3,5. Em suma, obtendo um grau de relevância em torno de 70%.

Por fim, oportuniza-se a implementação de GGD de maneira anterior a projetos de *Self Service* BI e *Analytics*, vez que sólidas definições acerca de GGD tendem a fomentar os projetos de dados. De modo que estes – os projetos de dados - iniciem sabendo dos padrões e arquiteturas estabelecidos pela GGD. Seriam estes os fatores oportunos a promoverem ambientes de dados íntegros e organizados. Especialmente para fins de "*Business Intelligence & Analytics*" apoiando as tomadas de decisões em operações de negócio.

# Referências

Barros, Aidel Jesus da Silveira; LEHFELD, Neide Aparecida de Souza. A pesquisa e a iniciação científica. Barros, AJP, Lehfld, NAS, p. 81-104, 2007.

Cochrane, Mike. 5 Steps to Data Governance. 2009. Disponível em < <https://www.information-management.com/news/5-steps-to-data-governance>>. Acesso em 19 de setembro de 2019.

Columbus, Louis, 84% of enterprises see big data analytics changing their competitive landscape. 2014. Disponível em <https://www.forbes.com/sites/louiscolumbus/ 2014/10/19/84-of-enterprises-see-big-data-analytics-changing-their-industries-compe titive-landscapes-in-the-next-year/#3f4b31df17de>. Acesso em 20 de outubro de 2019.

Da Costas, M. A. M.; Martins, H. C. Análise quantitativa da formação do estoque de mão de obra qualificada de profissionais na área de TI. Revista Eletrônica de Sistemas de Informação, v. 15, n. 1, p. 1–45, 2016.

DAMA-DMBOK2. DAMA International. Data Management Body of Knowledge Data Management Association. 2017.

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 17 Agosto de 2019.

Davenport, T. H. (2013). Analytics 3.0. Harvard Business Review, v. 91, n. 12.

Davenport, T. H. (2014). How strategists use “big data” to support internal business decisions, discovery and production. Strategy and Leadership, 42, 45–50.

Huet, Ellen. How Airbnb uses big data and machine learning to guide hosts to the perfect price, 2015. Disponível em <https://www.forbes.com/sites/ellenhuet/2015/06/05/how-airbnb-uses-big-data-and-machine-learning-to-guide-hosts-to-the-perfectprice/#590b55486d49>. Acesso em 20 Outubro de 2019.

Informática, 2019. Disponível em <https://www.informatica.com/resources.asset.2e2200dae0a49c9efb20ef680abb45a6.pdf> Acesso em 15 Agosto 2019.

Marr, Bernard. The amazing ways Uber is using big data, 2015. Disponível em   
<https://www.datasciencecentral.com/profiles/blogs/the-amazing-ways-uber-is-using-big-data>. Acesso em 20 de Outubro de 2019.

Mosley, M. et al. 2009. The DAMA Guide to the Data Management Body of Knowlegde, 1ª edição, USA, Technics Publications, LLC, 406 p.

Perovano, D. G. Manual de metodologia da pesquisa científica. Curitiba: Intersaberes, 2016.

Prodanov, C. C.; FREITAS, E. C. Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico. 2. ed. Novo Hamburgo: FEEVALE, 2013.

Rêgo, Bergson Lopes. Gestão e Governança de Dados: Promovendo dados como ativo de valor nas empresas. Rio de Janeiro, Brasport, 2013.

Silva, E. L.; Menezes, E. M. Metodologia da pesquisa e elaboração de dissertação. 4. ed. Florianópolis: UFSC, 2005.

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/bigdatagovernance-a-Framework-to-assess-maturity/>. Acesso em 20 outubro de 2019.

Tallon, P. P.; Short, J. E.; Harkins, M. W.The Evolution of Information Governance at Intel. MIS Quarterly Executive, v.12, n.4. p. 189 – 198, 2013.

Vanian, J. Why Data Is The New Oil. Fortune, Brainstorm Tech. 2016. Disponível em <http://fortune.com/2016/07/11/data-oil-brainstorm-tech/>. Acesso em 17 Agosto de 2019.

Weber, K.; Otto, B.; Osterle, H. One Size Does Not Fit All—A Contingency Approach to Data Governance. ACM Journal of Data and Information Quality, v. 1, n. 1, 2009.