

# O porquê de governança de dados em organizações de controle



**Ricardo Dantas Stumpf**

é servidor do Tribunal de Contas da União, lotado na Secretaria de Soluções de TI, onde atua na área de Gestão Técnica de Dados. É, também, bacharel em Ciências da Computação pela Universidade de Brasília.

## RESUMO

Existem diferentes níveis de governança dentro de uma organização, seja lidando com estratégia, com TI, com dados em geral ou dados abertos. Dados são ativos gerenciáveis, seja para o funcionamento saudável da instituição, seja para análises de dados sofisticadas. Este artigo trata de definir a utilidade da governança dos dados nesse contexto, assim como elementos básicos para sua implantação, com um viés para organizações de controle.

**Palavras-chave:** Governança de dados. Gestão de dados. Ciência de dados.

## 1. INTRODUÇÃO

O que é importante em uma organização? Pessoas, orçamento, equipamentos e imóveis, dados... É difícil sequer pensar no funcionamento de organizações caso falte algum desses elementos, o que prova a crucial importância de todos eles. Vários desses são **ativos** essenciais à maioria das entidades e, como tais, devem ser bem **geridos** e **governados**, de modo a maximizar o seu valor. Em síntese, gerir é garantir que fazemos certo as coisas, enquanto governar é garantir que fazemos as coisas certas (WODZINSKI et al, 2015).

E agora que o TCU está investindo fortemente em análise de dados, cuidar dos dados tornou-se ainda



mais crítico, como pode-se inferir do trecho do relatório trimestral abaixo:

[C]hegou-se ao consenso dos três principais desafios que o tribunal deverá superar para obter bons resultados com a análise de dados: o desafio técnico, especialmente vinculado à qualidade dos dados; o regulatório, atinente às restrições legais e normativas; e o cultural, concernente aos aspectos comportamentais das pessoas envolvidas (RELATÓRIO TRIMESTRAL TCU 4º TRIMESTRE/2014, p. 101).

A 2ª edição do guia de Governança do TCU (2015) traz um resumo da interação entre a gestão e a governança:

“São funções da governança:

- a. definir o direcionamento estratégico;
- b. supervisionar a gestão;
- c. envolver as partes interessadas;

**Figura 1:**

TCU - Referencial básico de Governança aplicável a órgãos e entidades da Administração Pública, 2ª edição, p. 32



- d. gerenciar riscos estratégicos;
- e. gerenciar conflitos internos;
- f. auditar e avaliar o sistema de gestão e controle; e
- g. promover a *accountability* (prestação de contas e responsabilidade) e a transparência” (TCU, REFERENCIAL BÁSICO DE GOVERNANÇA APLICÁVEL A ÓRGÃOS E ENTIDADES DA ADMINISTRAÇÃO PÚBLICA, 2ª edição, p. 31).

E essa lógica vale tanto para o nível estratégico, como para os diversos níveis e facetas mais focadas da instituição. Em torno da governança e gestão corporativas, gravitam, dentre outras, a governança e gestão de tecnologia da informação e, em torno desta, a governança e gestão de dados produzidos e custodiados. Mesmo a governança e gestão de dados também se subdivide nas organizações mais complexas.

## 2. DADO E INFORMAÇÃO

É comum usarmos dado e informação como sinônimos, assim como conhecimento e competência. A rigor, porém, representam conceitos diferentes, segun-

do o professor Valdemar Setzer (2015), da Universidade de São Paulo.

Dado é “uma sequência de símbolos quantificados ou quantificáveis”, enquanto informação é “uma abstração na mente de alguém”.

Como exemplo simples, a sequência de símbolos **880**. É um **dado**, digital ou não. Se alguém reconhece 880 numa lista de salários, então temos a **informação** de um salário numérico com 880 unidades.

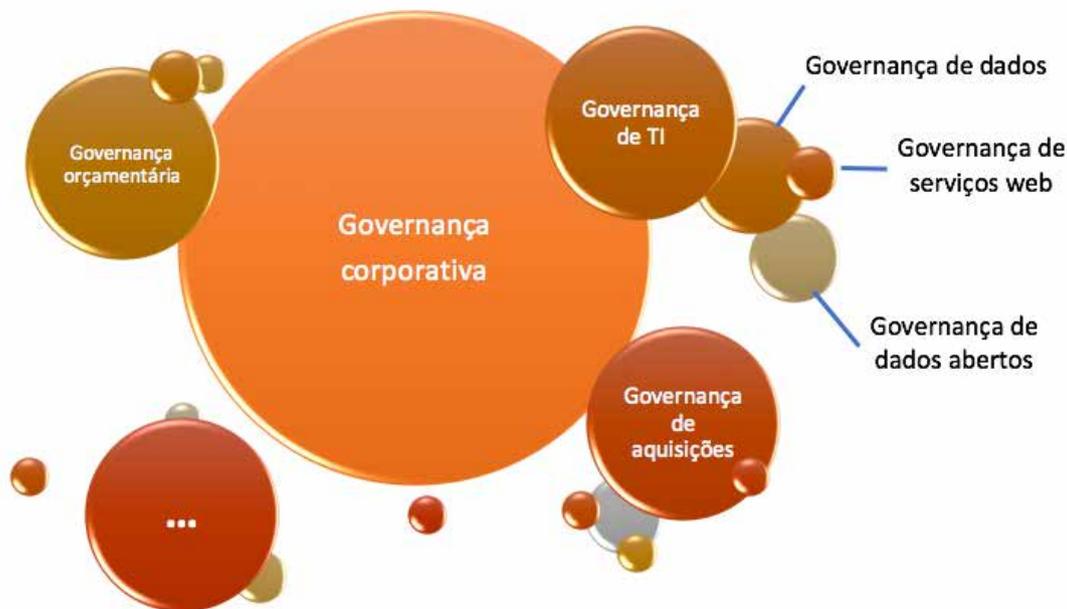
Já **conhecimento** exige uma abstração e experimentação pessoal (SETZER, 2015). Um funcionário de uma empresa sabe, da experiência, que se trata de R\$880,00, valor do salário mínimo em outubro de 2016 atribuído a um colaborador recém-contratado. O número 880 passou de dado a informação e, então, a conhecimento. Finalmente, um auditor tem **competência** para interagir e atuar sobre esse conhecimento durante uma fiscalização.

Esse processo acontece diariamente em ambientes profissionais: “processo” é um documento, fluxo de trabalho repetível, instrumento jurídico, código de computador em execução ou saliência de ossos? Depende.

É aí que a **Governança de dados** entra, ao utilizar o conhecimento e competência das pessoas para definir políticas, responsabilidades, glossários, metadados, fluxos de trabalho dos dados em movimento,

**Figura 2:**

Diferentes níveis de governança “gravitando” simultaneamente



**Figura 3:**

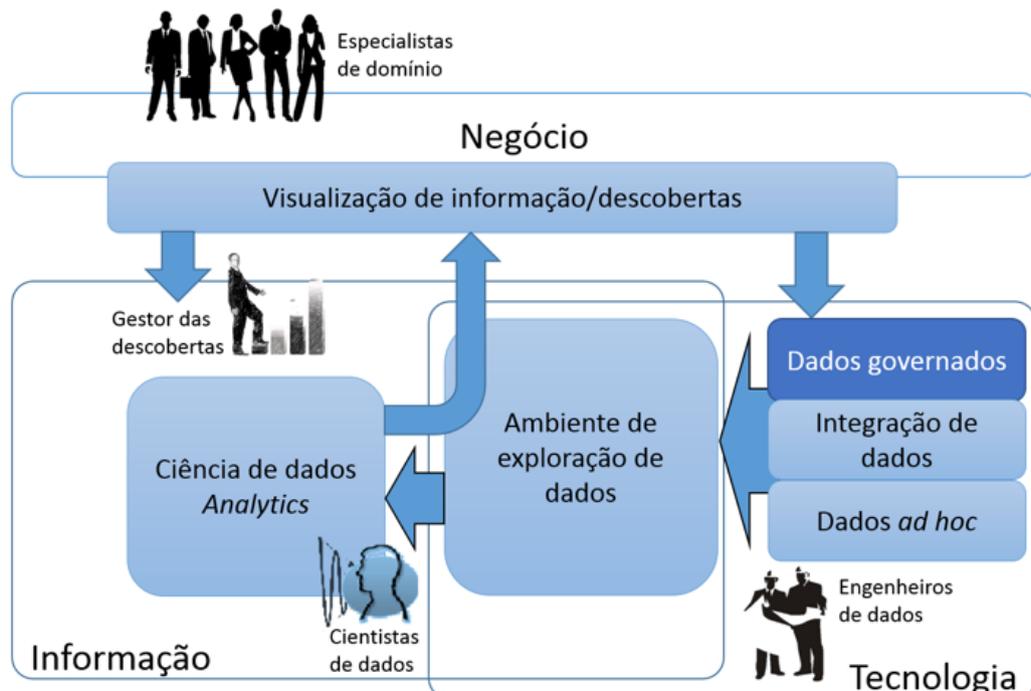
Exemplo de como símbolos relacionam-se com diversos conceitos



monitoramento e linhagem de dados que permitam à organização melhorar a qualidade, descoberta e entendimento desses, de modo a simplificar a extração de informações e conhecimentos, resultando em melhores análises e decisões de negócio.

**Figura 4:**

Modelo de ciclo de dados, adaptado (HENDERYCKX, 2016)



### 3. NOVO PETRÓLEO É DIGITAL

Dados estão sendo chamados de “novo petróleo” (VANIAN, 2016), mas sempre foram importantes e presentes, mesmo antes do formato digital. Contudo, nosso foco está nos dados digitais, que estão crescendo em tal ritmo e variedade que têm sido associados a “tsunamis”, “avalanches” e “tempestades” de dados, ou *Big Data*. Os dados internos, normalmente os mais alinhados com os processos de trabalho internos da corporação, são chamados de *Small Data* (HENDERYCKX, 2016).

Uma tendência nas empresas é investir cada vez mais em análise e visualização de dados, de modo a vislumbrar mais oportunidades de negócio e a aumentar a eficiência operacional. Da mesma forma, órgãos públicos, em especial os de controle, estão coletando grandes quantidades de dados de seus parceiros e auditados, além do *feedback* da sociedade nas redes sociais.

Jan Henderyckx, presidente da DAMA Benelux, apresentou um modelo genérico para o ciclo dos dados (figura 4):

Nesse modelo, observe que os dados que vêm das áreas de negócio precisam ser bem governados. Isso não significa que os dados externos custodiados não precisam de governança, mas as preocupações primárias podem ser outras. Por exemplo, o TCU possui um ambiente utilizado pelo Controle Externo como

laboratório de exploração e cruzamento de dados externos e internos.

A área responsável por esse laboratório adota, quando necessário, o “Princípio da Oportunidade”, que justifica a inclusão *ad hoc* de novas bases, mesmo antes de haver tratamentos de qualidade e documentação mais detalhados. O motivo é não deixar, no contexto de alguma auditoria, a janela da oportunidade de análise passar. A partir daí, porém, tal base torna-se elegível a um processo de tratamento mais completo.

Além disso, existe a questão **contextual**. Bases externas de baixa qualidade podem ser úteis para se descobrir achados de auditoria em sistemas de informação, enquanto bases derivadas de boa qualidade, com documentação completa de metadados, facilita a análise relativa ao uso efetivo dos repasses de recursos financeiros, por exemplo.

Quando o modelo apresentado tem baixa qualidade e organização, surgem problemas em três momentos (DYKES, 2016):

- a. as pessoas não confiam nos dados *a priori*;
- b. há tantas fontes que não se sabe o que usar; e
- c. há paralisia de análise, quando é difícil saber quando a análise está boa o suficiente.

Isso dito, podemos nos voltar para gestão e governança dos dados.

#### 4. GESTÃO DE DADOS

Organizações costumeiramente já gerem seus dados, em especial os estruturados e semiestruturados. Quanto à forma, é comum que a gestão siga boas práticas em seus nichos, normalmente pela força da necessidade e pelo conhecimento compartilhado pelos pares internos e parceiros, como os da comunidade TI-Controle, com cerca de 20 órgãos de controle que discutem melhores práticas em TI. A sistematização dos nichos de atuação aparece em *frameworks* como o preconizado no DMBoK (MOSLEY *et al.*, 2012), da DAMA International.

Na nova versão desse *framework*, até técnicas mais modernas, como integração por serviços web e *publish-subscribe* são contemplados como formas de dados circularem entre sistemas ou módulos de sistemas, localmente ou em nuvem (BRADLEY, 2016).

Apesar disso, o tema Governança de dados (que é uma espécie de “cola” a todos os processos ligados ao as-

sumo) só agora está ganhando energia, visto que as empresas e organizações que dependem de dados querem extrair o máximo valor desse ativo, cada vez mais abundante e mais complexo. Cabe lembrar que já podem existir áreas que governam bem seus dados. Nesses casos, segundo o político inglês Lucius Cary, “quando não é necessário mudar, é necessário não mudar” (In: SEINER, 2014).

A matéria-prima para o trabalho de qualquer órgão de controle é informação. E seus produtos finais (determinações, recomendações) são fontes primárias de informação a todos os seus auditados.

#### 5. GOVERNANÇA DE DADOS

Finalmente chegamos a uma definição formal, segundo John Ladley (2012):

Governança de dados é a organização e implementação de políticas, procedimentos, estrutura, papéis e responsabilidades que delineiam e reforçam regras de comprometimento, direitos decisórios e prestação de contas para garantir o gerenciamento apropriado dos ativos de dados.

Da definição, discutiremos brevemente a necessidade de princípios, políticas, curadores de dados e clas-

Figura 5:

Visão sintética do *framework* para dados, segundo DMBoK



sificação dos dados, glossários e metadados, lideranças e abordagens de implantação.

O importante é saber dosar. Governança de dados muito burocrática é um convite à desobediência; já o excesso de flexibilidade pode levar à desgovernança, a uma gestão de dados menos eficiente. Por isso, deve-se começar pela definição dos **princípios**.

### 5.1 PRINCÍPIOS

Princípios bem específicos, lógicos e viáveis facilitam todos os passos seguintes. Por exemplo, se o princípio “dados são ativos corporativos” for adotado pela organização, devemos tratar a entidade “dados” como tratamos outros ativos: melhoria contínua, definição de responsabilidades etc (LADLEY, 2012).

Cada organização tem seus princípios. A tabela 1 abaixo é exemplificativa e não-exaustiva.

### 5.2 POLÍTICAS

Como diz John Ladley (2012), a política é o que dá “dentes aos princípios”. É o documento formal que faz a organização adotar os princípios discutidos. É comum também a política definir responsabilidades pelos dados (curadoria), assim como a estrutura organizacional que irá conduzir e monitorar os esforços de Governança de Dados.

Como um bom exemplo, deixamos a sugestão de leitura da Política de Governança da Informação do Banco Central do Brasil, publicada pela Portaria nº47, de 20 de fevereiro de 2013, atualizada em 2016.

### 5.3 CURADORIA E CLASSIFICAÇÃO

Os curadores (*data stewards*, em inglês) são as pessoas ou grupos de pessoas que têm responsabilidades de **cuidar dos dados sob sua alçada de negócio**. Essa é uma mudança fundamental, pois compartilha com a TI a missão de cuidar dos dados corporativos.

Suponha que o princípio da qualidade tenha sido adotado. Isso significa que as áreas de negócio têm que se preocupar com a qualidade do dado que geram para si e para as outras, e devem reportar problemas de qualidade em relatórios e outras fontes que venham de outras áreas de negócio. O Escritório de **Governança de Dados**, ou equivalente, é o setor que auxiliará os curadores nessa tarefa.

A partir da Lei nº 12.527/2011, conhecida como a Lei de Acesso à Informação, formalizou-se como diretriz para a Administração Pública a “publicidade [das informações] como preceito geral e do sigilo como exceção”. Assim, os curadores também devem ativamente classificar as informações que produzem ou custodiam, de modo a acelerar a disponibilização dos dados públicos ou de restringi-los nos termos da lei.

### 5.4 GLOSSÁRIOS E METADADOS

Glossários, ou vocabulários, definem e eliminam ambiguidades em termos de negócio, como o termo “processo”, citado na seção 2. O TCU mantém o Vocabulário de Controle Externo (VCE), criado para, segundo o Ministro-Presidente Aroldo Cedraz, “padronizar o tratamento de informações especializadas e conferir maior

**Tabela 1:**

Lista de alguns princípios para implantação de Governança de Dados

| Princípios       | Descrição   |
|------------------|---|
| Regra de ouro    | Todos os dados são tratados como ativo corporativo.   |
| Federação        | Há padrões definidos para as estruturas de dados.   |
| Eficiência       | Dados relevantes devem estar disponíveis no momento certo, no lugar certo e no formato certo para usuários autorizados. |
| Qualidade        | Dados corporativos são medidos e geridos para terem qualidade.  |
| Gestão de risco  | Manter conformidade com a legislação, políticas e normativos internos relativos a dados.                                |
| Colaboração      | Dados corporativos são recursos compartilhados e publicizados.  |
| Contextualização | O contexto de uso do dado muda sua forma de armazenamento, tratamento e utilização.                                     |
| Inovação         | Novas técnicas são incentivadas, seguindo-se os demais princípios.  |

Fonte: adaptado de LADLEY (2012)

agilidade e precisão na recuperação dos conteúdos presentes nos sistemas de informação do TCU” (VCE, p.5).

A definição clássica de metadados é “dado sobre o dado”. Mas podemos ampliar esse conceito em duas vertentes: de negócio e técnicos.

Os metadados de negócio são aqueles que contextualizam os dados armazenados ou em movimento. Por exemplo, considere a tabela WZV da base de dados TTNN (hipotéticos). Como a área de negócio e os especialistas de domínio conhecem esse assunto, eles podem **enriquecer os metadados** existentes com informações específicas. Agora, os cientistas de dados saberão que a tabela WZV tem os dados cadastrais de fornecedores e que a base TTNN traz organizações auditadas que, dentre outros, compraram daqueles fornecedores.

Já os metadados técnicos são aqueles que trazem informações de TI sobre os dados, como nomes de campos e tabelas, datas de carga, versão, tamanhos, tipos etc.

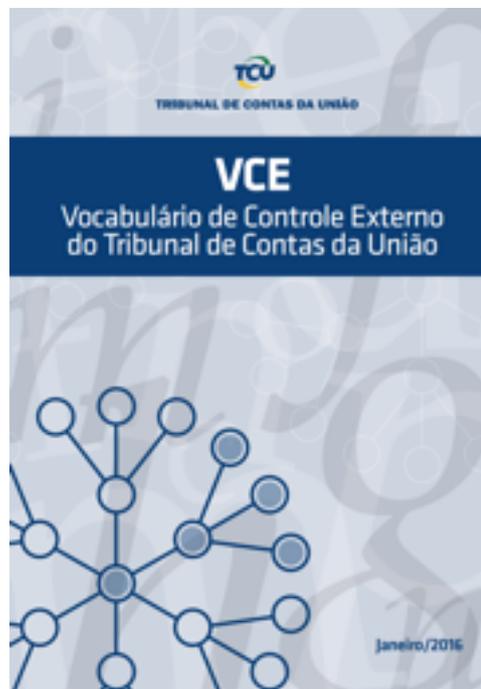
A “magia” acontece quando os metadados técnicos e de negócio são **integrados para descoberta** e uso das áreas de análise e de gestão de mudanças. No TCU, há algumas iniciativas nesse caminho. Os dados do VCE estão aos poucos sendo integrados a outras soluções de tecnologia da informação, de forma a otimizar e uniformizar o uso dos conceitos. Outra é uma ferramenta especializada em metadados que mostra, de forma integrada, informações oriundas de modelagens de dados, integrações por ETL (*Extract Transform Load*) e bases de dados corporativas (figura 7).

## 5.5 ESTRUTURAS

Cada corporação tem sua cultura e necessidades, e não faria sentido haver apenas uma forma de estrutura

**Figura 6:**

Capa da edição de janeiro/2016 do VCE do TCU

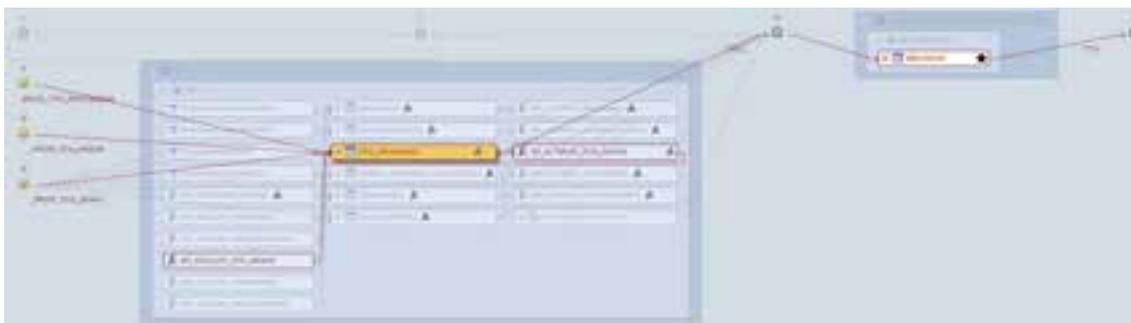


organizacional para governar dados. A figura do CDO (*Chief Data Officer*) é cada vez mais comum. Essas são algumas das formas utilizadas, segundo Bergson Rêgo (2013):

- a. CIO liderando equipe centralizada na TI, que interage com o negócio;
- b. CDO interagindo com TI e uma área de gestão de dados central do negócio;

**Figura 7:**

Exemplo editado de metadados técnicos integrados



- c. CDO interagindo com TI e várias áreas de gestão de dados do negócio; ou
- d. um CDO para cada grande linha de negócio, interagindo com TI (híbrido).

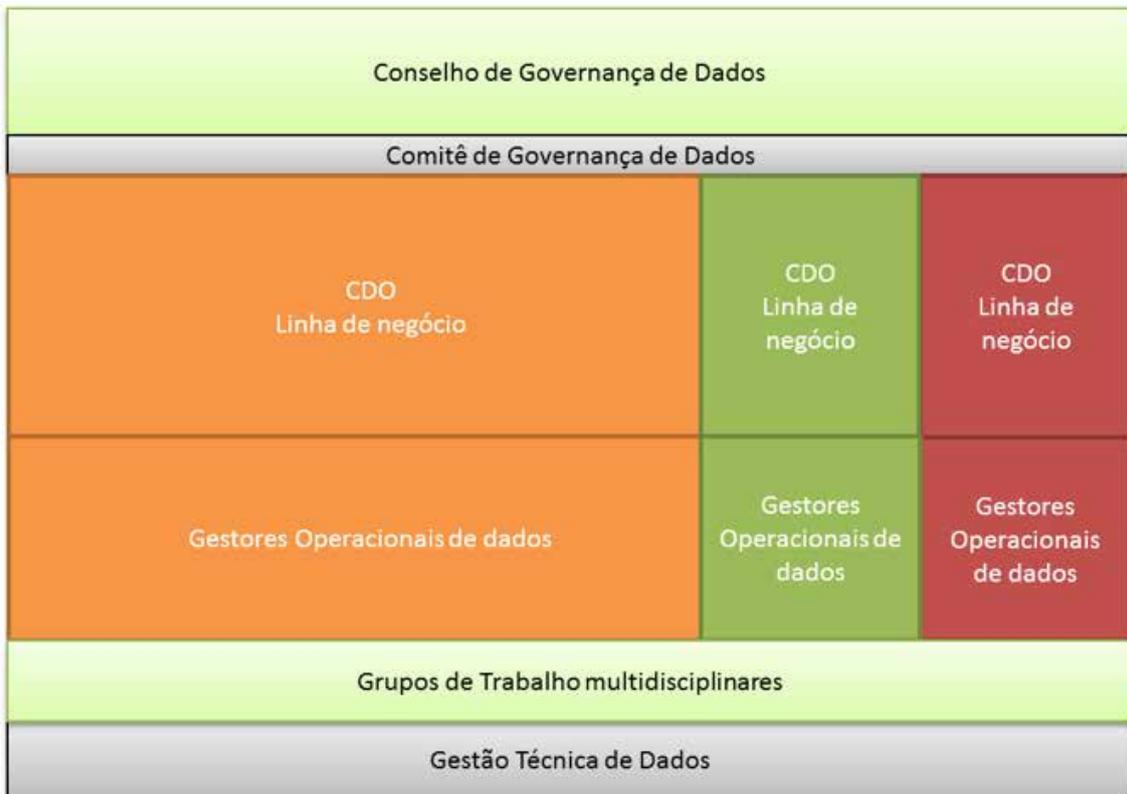
**Figura 8:**

Estrutura simples hierárquica (RÉGO, 2013)



**Figura 9:**

Estrutura híbrida, adaptado (IPPOLITI, 2016)



A estrutura híbrida é particularmente interessante, e utilizada na prática em um grande banco multinacional (IPPOLITI, 2016). Dada a grande variedade interna de dados disponíveis, eles resolveram dividir a gestão desses em grandes linhas de negócio (figura 9).

### 5.6 ABORDAGENS DE IMPLANTAÇÃO

Há três formas estruturadas mais comuns de se implantar Governança de Dados (SEINER, 2016):

- a. Comando e controle;
- b. Tradicional; e
- c. Não-invasivo.

Na abordagem comando e controle, o comitê de mais alto nível define as regras, determina os curadores, compra ferramentas e cobra resultados.

A tradicional identifica os curadores e os orienta com processos mais genéricos e ferramentas existentes, e mede resultados analisando os dados envolvidos.

Por fim, a não-invasiva formaliza a curadoria tácita, foca em aplicar os processos existentes, estimula construção de ferramentas e uso das atuais e mede resultado pelo aumento percebido em eficiência e efetividade de capacidades de análise.

Como já frisamos, nenhuma é necessariamente melhor que a outra. Depende da cultura e necessidades de cada organização.

#### 4. CONCLUSÃO

Apesar da importância atual do tema Governança de Dados, esta será ainda maior no futuro, no qual haverá muita automatização, mesmo de atividades intelectuais, desde que repetíveis. E os dados, sejam oriundos de sistemas informatizados ou de sensores em objetos do dia a dia, serão um instrumento básico para essa automatização.

Isso significa que as organizações deverão capturar dados, armazená-los de forma eficiente e automatizar análises com algoritmos cognitivos (HENDERYKCX, 2016). Segundo Ian Rowlands (2016), metadados de negócio terão sua semântica, e não apenas sua sintaxe, inferida por algoritmos e *crowdsourcing* (grandes grupos pensando e colaborando juntos).

Também é desafio crescente identificar e reduzir o *dark data*, que representa perda de oportunidades, desperdício de recursos e risco (ROWLANDS, 2016). Rowlands reitera trecho de declaração do Au-

ditor-Geral do Canadá, em que esse é um problema a ser combatido:

Um dos temas que unem muitas de nossas auditorias é que os dados coletados de muitas organizações governamentais ou são inutilizáveis, ou não são usáveis ou não são utilizados (Auditor-Geral do Canadá em ROWLANDS, 2016).

Para que tal redução seja possível, impera a necessidade de governar tais dados – na dose certa para cada caso e contexto, sempre com o **olhar na missão e visão da instituição**.

#### REFERÊNCIAS

BRADLEY, C. The new DMBOK 2 discipline of Data Integration. Disponível em: <<https://www.brighttalk.com/webcast/12405/186095/the-new-dmbok-2-discipline-of-data-integration>>. Acesso em: 12 abr. 2016.

BRASIL, BANCO CENTRAL DO BRASIL – BCB. Portaria nº 47, de 20 de fevereiro de 2013. Diário Oficial da União – DOU, 21 de fevereiro de 2013. Seção 1, p. 24.

BRASIL, Comunidade de Tecnologia da Informação Aplicada ao Controle – TI Controle. Sobre a Comunidade. Disponível em: <<http://www.ticontrôle.gov.br/ticontrôle/sobre-a-comunidade/sobre.htm>>. Acesso em: 20 out. 2016.



BRASIL. LEI nº 12.527, de 18 de novembro de 2011. Lei de Acesso à Informação. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/l12527.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm)>. Acesso em: 20 out. 2016. 2011.

BRASIL, TCU – TRIBUNAL DE CONTAS DA UNIÃO. Referencial básico de governança aplicável a órgãos e entidades da administração pública. Versão 2 – Brasília: Tribunal de Contas da União. Secretaria de Planejamento, Governança e Gestão. 80p. Disponível em: <<http://www.tcu.gov.br/governanca>>. Acesso em: 20 out. 2016. 2014.

\_\_\_\_\_. Relatório de Atividades: 4º trimestre de 2014 – Brasília: Tribunal de Contas da União. Secretaria-Geral da Presidência, p.101. Disponível em: <<http://portal.tcu.gov.br/publicacoes-institucionais/relatorios/relatorios-de-atividades>>. Acesso em: 20 out. 2016. 2015.

\_\_\_\_\_. Vocabulário de Controle Externo do Tribunal de Contas da União – Brasília: Tribunal de Contas da União. Instituto Serzedello Corrêa. Disponível em: <<http://portal.tcu.gov.br/vocabulario-de-controle-externo/>>. Acesso em: 20 out. 2016. 2015.

CANADÁ, Office of the Auditor General of Canada – OAG. Spring Reports of the Auditor General of Canada. Auditor General’s Opening Statement. Disponível em: <[http://www.oag-bvg.gc.ca/internet/English/osm\\_20160503\\_e\\_41358.html](http://www.oag-bvg.gc.ca/internet/English/osm_20160503_e_41358.html)>. Acesso em: 20 out. 2016. 2016.

DMBOK. MOSLEY, M.; BRACKETT, M.; EARLEY, S. HENDERSON, D. DAMA Guia para o corpo de conhecimento em gerenciamento de dados. Technics Publications, versão brasileira 2012.

DYKES, B. Big Data Paralyzing Your Business? Avoid These 3 Common Traps. Forbes, Entrepreneurs. Disponível em: <<http://www.forbes.com/sites/brentdykes/2016/09/28/big-data-paralyzing-your-business-avoid-these-3-common-traps/>>. Acesso em: 20 out. 2016.

HENDERYCKX, J. Sustaining Data Driven Innovation. In: DATA MANAGEMENT CONFERENCE LATIN AMERICA, 2016, São Paulo. Anais eletrônicos... São Paulo: DAMA Brasil, 2016. 42 slides. Disponível em: <<http://www.dmc-latam.com/palestrantes/jan-henderyckx/>>. Acesso em: 19 set. 2016.

IPPOLITI, J. Implementing Data Governance Strategies. In: ENTERPRISE DATA GOVERNANCE ONLINE, 2016. Anais eletrônicos... 20 slides. Disponível em: <<http://video.dataversity.net/video/keynote-implementing-data-governance-strategies/>>. Acesso em: 12 fev. 2016.

LADLEY J. Data Governance: How to Design, Deploy and Sustain an Effective Data Governance Program. The Morgan Kaufmann Series on Business Intelligence. Morgan Kaufmann. 2012.

RÊGO, B. L. Gestão e Governança de Dados: Promovendo Dados Como Ativo de Valor Nas Empresas. Rio de Janeiro: Brasport, p. 60-74. 2013.

ROWLANDS, I. Data Management in Motion. In: DATA MANAGEMENT CONFERENCE LATIN AMERICA, 2016, São Paulo. Anais eletrônicos. São Paulo: DAMA Brasil, 2016. 25 slides. Disponível em: <<http://www.dmc-latam.com/palestrantes/ian-rowlands/>>. Acesso em: 19 set. 2016.

SEINER, R. Non-Invasive Data Governance: The Path of Least Resistance and Greatest Success. Technics Publications. 2014.

\_\_\_\_\_. Comparing Approaches to Data Governance. Disponível em: <<http://tdan.com/comparing-approaches-to-data-governance/20386>>. Acesso em: 5 out. 2016.

SETZER, V. W. Dado, informação, conhecimento e competência. Journal Data & Information 1(1), DAMA Brasil: p. 38-55. 2015.

VANIAN, J. Why Data Is The New Oil. Fortune, Brainstorm Tech. Disponível em: <<http://fortune.com/2016/07/11/data-oil-brainstorm-tech/>>. Acesso em: 20 out. 2016. Julho de 2016.

WODZINSKI, M. et al. Building an Impactful Data Governance Program One Step at a Time. Disponível em: <<https://www.brighttalk.com/webcast/10477/142503/building-an-impactful-data-governance-program-one-step-at-a-time>>. Acesso em: 26 fev. 2015.