

Empresas Orientadas a Dados e Análises: a tecnologia está a serviço da tomada de decisão?

David J. Soares¹

Resumo

O intenso processo de criação de dados digitais, internos e externos, tem se traduzido para as organizações em desafios e oportunidades, na medida em que, extrair informações relevantes para suportar a tomada de decisão exige mudanças em seus modelos, processos e visão de negócios. Ao tomarem decisões orientadas por dados e análises, as organizações almejam gerar valor para seus *stakeholders* e desenvolver vantagens competitivas sustentáveis. Assim, objetivou-se com esta pesquisa, analisar o nível de maturidade em dados e análises como direcionadores do processo de tomada de decisão em um grupo empresarial, sob a perspectiva de dados e tecnologia. Adaptou-se para o estudo um modelo de avaliação de maturidade em dados e análises desenvolvido pela Computer Science Corporation (CSC). Utilizou-se uma abordagem mista de pesquisa na qual, na primeira etapa, de caráter quantitativo, foi realizada uma *survey* com 175 participantes e, na segunda etapa, de caráter qualitativo, foram realizados oito grupos focais, atingindo um total de 66 participantes. Os resultados quantitativos apontam para um nível intermediário de maturidade, o nível Replicável. Contudo, os resultados qualitativos apontam para o nível básico de maturidade, denominado *Ad Hoc* ou Esporádico. Em ambos os casos, os resultados sugerem a necessidade de investimentos em tecnologias apropriadas capazes de coletar, integrar e analisar os dados e torná-los disponíveis e acessíveis para os tomadores de decisão; assim como, permitir a incorporação de soluções analíticas nas aplicações empresariais existentes.

Palavras-chave: Big Data; Maturidade em Dados e Tecnologia; Tomada de Decisão.

¹ Graduando em Gestão da Informação pela Faculdade de Gestão e Negócios – FAGEN, Universidade Federal de Uberlândia – UFU; david@ufu.br.

1 Introdução

Para especialistas, estudiosos da área e empreendedores os dados são o novo petróleo da economia global, visto o seu potencial para a geração de valor, riqueza material e inovação (WEDEL; KANNAN, 2016). Para ilustrar tal entusiasmo, estimativas apontam que o potencial de geração de riqueza por meio da análise efetiva dos dados existentes supera o valor de 300 bilhões de dólares (MCKINSEY GLOBAL INSTITUTE, 2016), logo, caracterizando-se como um fator chave para a sustentabilidade dos negócios.

A exponencialidade dos dados criados diariamente tem se imposto como um dos principais desafios enfrentados pelas organizações, na medida em que extrair informações relevantes para contribuir no processo de tomada de decisão, tanto do ponto de vista tecnológico (hardware, software e metodologias) quanto do de pessoas preparadas para extrair valor dos dados, tem exigido mudanças significativas nos modelos, processos e visão de negócios das mesmas (LOPES; SOARES, 2016). Se analisarmos apenas a quantidade de dados que são criados em apenas 60 segundos no mundo via as principais plataformas digitais, a tarefa pode parecer inviável: 46.200 posts são realizados no Instagram, 70.017 horas assistidas na Netflix, 342.000 apps são baixados no Google Play e Apple App Store, 452.000 tweets são enviados pelo Twitter, 3.5 milhões de pesquisas são realizadas no Google, 4.1 milhões de vídeos são vistos no YouTube, 16 milhões de mensagens de texto (SMS) e 156 milhões de emails são enviados (POPOMARONIS, 2017).

Ao somarmos os dados internos às organizações e os dados externos como mídias sociais, relatórios diversos, artigos acadêmicos, vídeos, páginas da Internet, hábitos de navegação, sinais de sensores, dados de localização espacial (GPS), smartphones, informação genômica, entre outros, temos o que se convencionou chamar de *Big Data*. Ou seja, uma gigantesca quantidade de dados que se gerenciados de forma apropriada é capaz de fornecer às empresas vantagens competitivas sustentáveis e também geração de valor para os *stakeholders* (JANSSEN; VOORT; WAHYUDI, 2016).

Organizações que almejam extrair a plenitude do potencial advindo dos dados devem, na visão Gupta e George (2016), se tornarem *data-driven companies*, ou seja, organizações que tomam decisões orientadas por dados e análises. Nesse sentido,

empresas tem adotado e implantado soluções de *Business Intelligence (BI)*, *Big Data*, *Analytics* em suas atividades de planejamento, execução e controle. Entretanto, as organizações, muitas vezes, têm focado suas atenções apenas no aspecto tecnológico e ferramental, o que as leva a não ter uma visão holística, necessária e *sine qua non* do ecossistema de dados, o que consequentemente diminui a efetivação da estratégia global de realizar decisões assertivas orientadas por dados e análises (BOYD; CRAWFORD, 2012).

Nesse sentido, as organizações, antes mesmo de investirem fortemente em tecnologias ou estruturas, devem avaliar qual o seu nível de prontidão para a orientação a dados e o seu nível de maturidade na temática.

Maturidade em dados e análises pode ser entendida como o nível em que uma organização coleta, integra, gerencia e potencializa todas as fontes de dados relevantes internas e externas (CSC, 2013). Isto significa criar um ecossistema que entrega resultados relevantes, possibilitando transformações de impacto nos negócios, ou seja, maturidade em dados e análises não é somente ter alguma tecnologia, ou dados, ou habilidades, mas sim, é um processo que envolve criar um ecossistema que inclui tecnologia, gestão de dados, análises, governança, habilidades e outros componentes organizacionais (HALPER; KRISHNAN, 2013).

A análise da maturidade em dados e análise e a utilização de um modelo para avaliar esta maturidade são úteis por vários motivos e, dentre eles, destaca-se: o auxílio em criar uma estrutura em torno de dados e análises evidenciando por onde começar; auxilia na identificação e definição dos objetivos organizacionais em relação a dados e análises e na criação de um processo para comunicar a visão para toda a organização; um modelo de maturidade pode prover um processo metodológico para medir e monitorar o estágio de maturidade e o esforço necessário para complementar o estágio atual, bem como os passos necessários para alcançar um próximo estágio (HALPER; KRISHNAN, 2013).

Assim, objetivou-se, com esta pesquisa, analisar o nível de maturidade em dados e análises como direcionadores do processo de tomada de decisão em um grupo empresarial, sob a perspectiva de dados e tecnologia.

A organização estudada é um grupo empresarial com atuação nos setores de agronegócio, serviços, telecomunicações, tecnologia da informação (TI), entre outros;

um grupo empresarial que depende fortemente de sua capacidade analítica em dados e tecnologias de suporte.

2 Revisão da Literatura

2.1 Dados e Tecnologia

Quarenta e quatro zettabytes (44 ZB). Ao contrário do que o senso comum possa indicar, 44 ZB não configura em si uma informação. Como apontado por Zins (2007), um dado é caracterizado por signos numéricos ou alfabéticos que sem a devida contextualização não possuem significado inteligível. Nesse sentido, 44 ZB é um dado. Além disso, o autor define como informação um determinado conjunto de dados agrupados e/ou interpretados capazes de produzir e transmitir conhecimento. Assim, ao contextualizamos o dado informado inicialmente, temos uma informação relevante para o objetivo desse artigo.

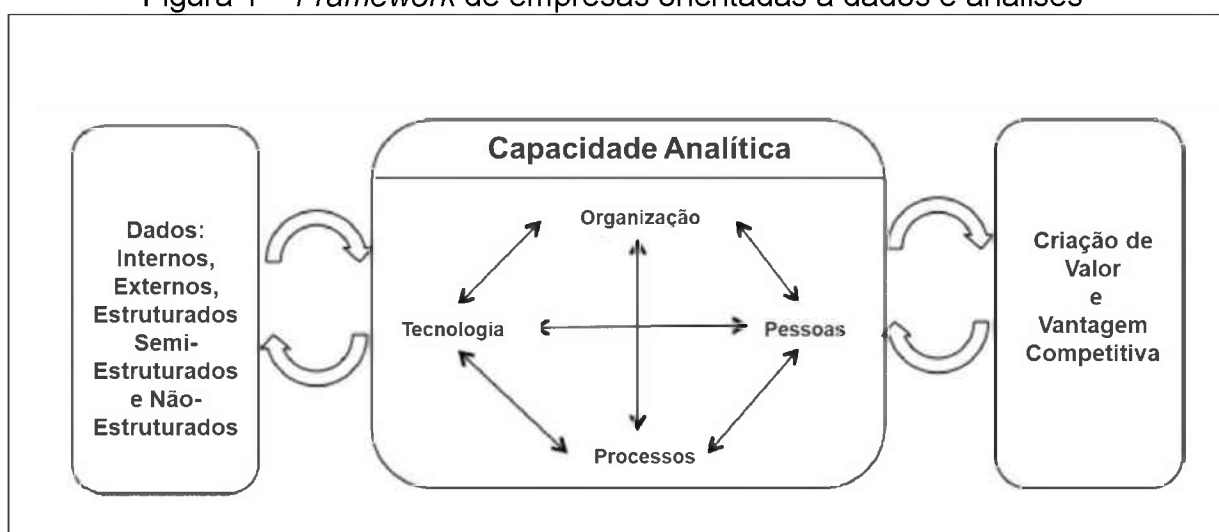
De acordo com a International Data Corporation (IDC) (IDC, 2014), em 2020, existirá no mundo um total de 44 ZB de dados digitais, criados ou copiados, o que equivale à 44 trilhões de gigabytes. Em 2013, esse número era de 4.4 ZB, ou seja, a cada dois anos a humanidade dobra a quantidade de dados criados e armazenados. Com essa quantidade de dados disponíveis interna e externamente às organizações, o desafio de transformar tais dados em informações para auxiliar no processo de tomada de decisão tem se tornado vital para a sustentabilidade dos negócios, públicos e privados. Os dados digitais podem ser caracterizados em três tipos distintos (PRASAD; ACHARYA, 2012):

- a) Estruturados: são dados que estão organizados em um modelo pré-determinado e pode ser facilmente processado por um computador. Como exemplo podemos citar os dados presentes em um banco de dados e que estão no formato de linhas e colunas;
- b) Semiestruturados: esse é o tipo de dado que não se adequa a uma modelagem de dados, porém, possui uma estrutura definida. Uma vez que, tais dados não estão em modelagem definida, como uma tabela, os mesmos não podem ser facilmente utilizados por um sistema computacional. São exemplos desse tipo de dado os e-mails, *Extensible Markup Language (XML)* e páginas HTML;

- c) Não-estruturados: são dados que não estão em conformidade com um modelo de dados previamente definido ou não estão em um formato que pode ser lido e processado por computadores de forma rápida e efetiva. Entre 80% e 90% dos dados internos às organizações são deste tipo e estão contidos em relatórios, webconferências, apresentações em PowerPoint, imagens, vídeos, corpo de e-mails, *chats*, entre outros. Neste tipo de dado é que se encontra o maior desafio para as organizações, uma vez que, para além da coleta e análise, é necessária a devida integração com os demais tipos e fontes de dados.

Na medida em que se acentua a globalização, abertura de mercados e consequente aumento concorrencial, as organizações buscam conquistar vantagens competitivas sustentáveis no intuito de se manterem relevantes e geradoras de valor para o seu segmento alvo de mercado e ao mesmo tempo, se manterem lucrativas (BHARADWAJ; VARADARAJAN; FAHY, 1993; BLACK; BOAL, 1994). Nesse sentido, Vidgen, Shaw e Grant (2017) e Akter et al. (2016) apontam que uma das formas para a criação de valor e desenvolvimento de uma vantagem competitiva sustentável é por meio da gestão e análise de dados - capacidade analítica orientada por dados e análises (Figura 1) -, ou seja, uma organização que disponibiliza dados confiáveis para o decisor no momento correto e adequado.

Figura 1 – *Framework* de empresas orientadas a dados e análises



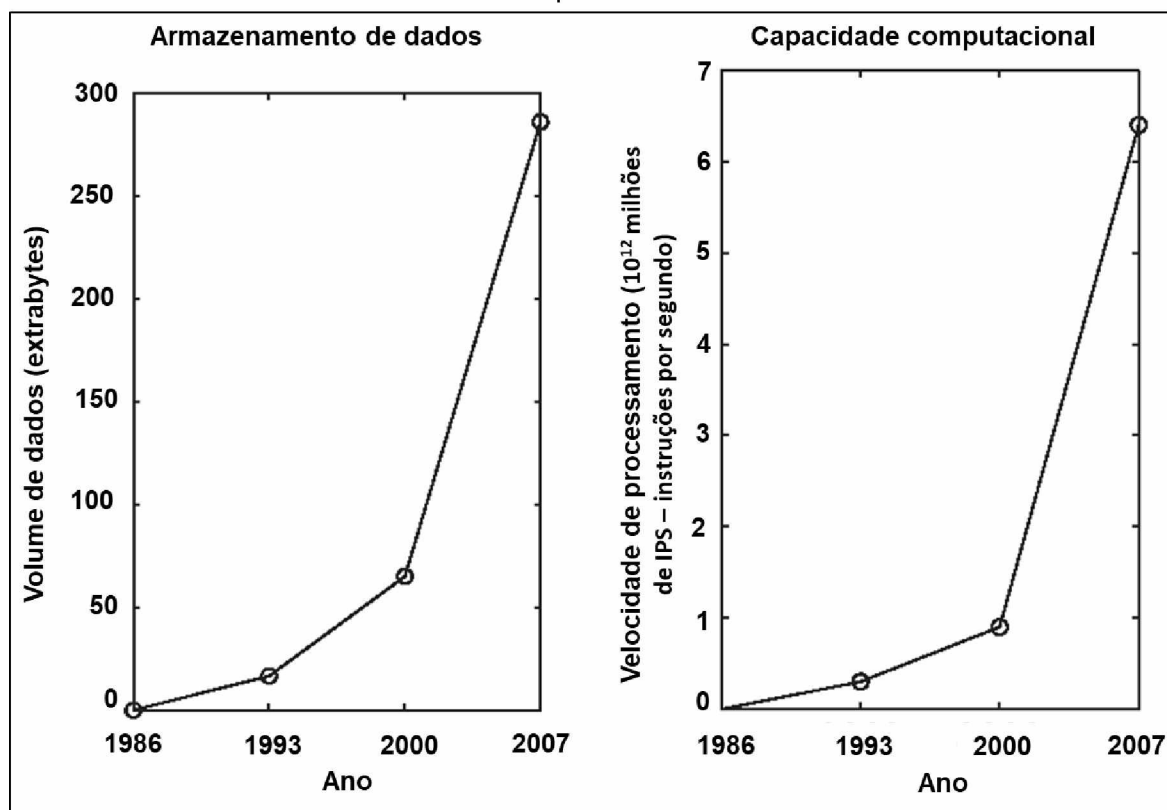
Fonte: Adaptado de Vidgen, Shaw e Grant (2017).

Empresas e instituições que utilizam dados e análises para suportar suas decisões são denominadas *data-driven companies*. Tais organizações de acordo com

McAfee e Brynjolfsson (2012), são, em média, 5% mais produtivas e 6% mais lucrativas, quando comparadas aos seus competidores que não fazem o uso efetivo dos dados no direcionamento das decisões dos negócios. Outro ponto relevante a se considerar é que empresas orientadas por dados e análises são mais ágeis em detectar e atuar ativamente sobre variáveis externas - oportunidades e ameaças de mercado-, que possam impactar a sustentabilidade da organização e também, inovar em seus processos e modelos de negócios (CÔRTE-REAL; OLIVEIRA; RUIVO, 2016; SORESCU, 2017).

A efetividade do processo de coleta, armazenamento e processamento dos dados está diretamente relacionado, dentre outras variáveis relevantes, à capacidade tecnológica existente – hardware, software e metodologias. Nesse sentido, podemos traçar um paralelo da evolução tecnológica entre a capacidade de armazenamento de dados e a capacidade computacional de processamento ao longo das décadas de 1980, 1990 e 2000 (Figura 2). De uma forma geral, a capacidade de armazenamento dos dados ao longo dos anos tem sido superior à capacidade de processamento computacional, o que indica a inabilidade das organizações, até então, de transformar dados em informações úteis para suportar a tomada de decisão (CHEN; ZHANG, 2014).

Figura 2 – Evolução da quantidade de dados armazenados versus a capacidade computacional



Fonte: Chen e Zhang (2014, p. 315, tradução nossa).

A partir de avanços nas tecnologias da informação (TI), notadamente os ocorridos nos últimos anos, as organizações podem, agora, coletar e analisar grandes quantidade de dados, sejam eles internos e/ou externos a mesma (COA; DUAN; LI, 2015). Dentre tais avanços podemos destacar o desenvolvimento de softwares como Microsoft Excel, Stata, IBM SPSS, SAS, S-Plus, Apache Kafka, Apache Drill, Apache Hadoop, Apache Mahout, Cassandra, Pentaho, Jaspersoft, Skytree Server, Tableau, Qlikview, Microsoft Power BI, Storm, SQLstream, Splunk, S4, SAP Hana, Talend Open Studio, Karmasphere, entre outros; ferramentas de programação e gestão de bancos de dados MySQL, NoSQL, R, MapReduce, Perl, Pig, Dremel, Hive, Spark, Java, Python, Dryad, entre outros; técnicas para análise de dados (Mineração de dados (*Data mining*), Redes Neurais (*Neural networks*), Aprendizado de máquina (*Machine Learning*), Processamento de sinais (*Signal Processing*) e métodos e ferramentas de visualização dotadas de capacidade analítica (WEDEL; KANNAN, 2016; CHEN; ZHANG, 2014).

2.2 Business Intelligence, Big Data e Analytics

Business Intelligence (BI) é um termo usado para descrever uma ampla categoria de aplicações, tecnologias, arquiteturas, e processos para a coleta, armazenamento, acesso e análise de dados para fornecer aos usuários de negócios informações competitivas e oportunas, permitindo melhores insights para a tomada de decisões operacionais e estratégicas (NEGASH, 2004; WATSON, 2009). Os termos *Analytics* e *Big Data* tornaram-se sinônimo de *BI* em alguns círculos, e, para outros, esses termos incorporam o *BI* tradicional, mas adiciona elementos, tais como a análise preditiva, mineração de dados, abordagens científicas e ferramentas de operações de pesquisa e gestão (GUPTA; GOUL; DINTER, 2015).

Big Data geralmente se refere a dados que excedem o armazenamento, processamento e capacidade de computação de bases de dados típicos e técnicas de análise convencionais. Como recurso, *Big Data* requer ferramentas e métodos que podem ser aplicados para analisar e extrair padrões de dados em grande escala. *Big Data* tem sido impulsionado pelo aumento da capacidade de armazenamento de dados, maior poder de processamento computacional e disponibilidade de maiores volumes de dados, provendo às organizações mais dados do que elas têm recursos computacionais e tecnologias para processar. Além dos óbvios grandes volumes de dados, *Big Data* também está associado a outras complexidades específicas, muitas vezes referidas como os quatro Vs: Volume, Variedade, Velocidade e Veracidade (DUMBILL, 2012; ZIKOPOULOS et al., 2012; GARSHOL, 2013; GROBELNIK, 2013).

Para os propósitos deste artigo, usou-se o termo *Business Intelligence e Analytics (BI&A)* para representar a mais ampla interpretação do campo. *Business Intelligence e Analytics (BI&A)* têm emergido como uma importante área de estudo tanto para praticantes como para pesquisadores e refletem a magnitude e o impacto dos problemas relacionados a dados a serem resolvido nas organizações (CHEN; CHIANG; STOREY, 2012).

As oportunidades associadas a dados e análises em diferentes organizações têm ajudado a gerar interesse significativo por *BI&A*, que são frequentemente referidos como técnicas, tecnologias, sistemas, práticas, metodologias e aplicações para analisar dados críticos de negócios para ajudar uma empresa a melhor compreender o seu negócio e o mercado, tomando decisões de negócios de forma

oportuna assertiva. Além do processamento de dados e tecnologias analíticas, *BI&A* inclui práticas e metodologias centradas no negócio que podem oferecer oportunidades com boa relação custo-benefício e gerar impactos em negócios como comércio eletrônico, inteligência de mercado, governo eletrônico, saúde, emprego, produtividade da economia, crime, desastres naturais, gestão de recursos e segurança (CHEN; CHIANG; STOREY, 2012; TINATI et al., 2014).

Tecnologias de *BI&A* e *Big Data* estão revolucionando drasticamente os negócios e a sociedade. O potencial ilimitado de uma economia impulsionada e orientada a dados é reconhecido por estudiosos e organizações diversas, e há cada vez mais entusiasmo para a noção de *BI&A* e *Big Data* (SHIN; CHOI, 2015). Embora estas tecnologias tenham o potencial de oferecer grandes vantagens competitivas, governos e empresas estão lutando para estabelecer governança e privacidade eficazes em relação às iniciativas de *BI&A* e *Big Data*. Enquanto o potencial é real, as práticas ainda são tímidas (EYNON, 2013).

Preocupações relacionadas a *BI&A* e *Big Data*, como invasão de privacidade, fragilidade na segurança, interoperabilidade limitada e pessoas raramente são examinadas em relação a outras preocupações relacionadas à tecnologia. Tais questões, incluindo os impactos sociais, culturais e éticos de como pode-se desenvolver e gerenciar a evolução de *BI&A* e *Big Data* será fundamental para seu sucesso (SHIN; CHOI, 2015).

BI&A e *Big Data* são componentes de um ecossistema social e, o desenvolvimento de sistemas deve focar na integração de dados com o meio social e cultural (SHIN; CHOI, 2015). Argumento esse que, consistente com a teoria sócio técnica, o projeto, análise e adoção de tecnologias devem ser baseados na compreensão contextual, ou seja, uma avaliação baseada em contexto que determina a adequação de uma tecnologia dentro de um contexto específico (RAMIREZ, 2007). A teoria está alinhada com o conceito de um ecossistema de *BI&A* e *Big Data*, onde estes representam ecossistemas de múltiplas camadas, conectado por redes de alta capacidade, serviços, aplicativos e usuários (SHIN; CHOI, 2015).

3 Metodologia

Objetivou-se com esta pesquisa, analisar o nível de maturidade em dados e análises como direcionadores do processo de tomada de decisão sob a perspectiva de dados e tecnologia.

Para alcançar o objetivo proposto, utilizou-se uma abordagem que pode ser caracterizada como mista, tendo etapas qualitativa e quantitativa (RICHARDSON, 1999). Na etapa quantitativa foi realizada uma *survey* para avaliar a percepção das pessoas do grupo empresarial sobre o nível de maturidade em dados e análises de suas respectivas empresas, ou seja, o nível de maturidade das empresas na utilização de dados e análises no processo de tomada de decisão considerando apenas a dimensão dados e tecnologia. Na etapa qualitativa foram realizados oito grupos focais com o propósito de aprofundar e elucidar as questões abordadas na etapa quantitativa.

O público alvo desta pesquisa era formado pelos executivos de todos os níveis e analistas de todas as empresas e todas as áreas operacionais.

Foram convidados a participar da *survey* aproximadamente 500 indivíduos. Destes, 310 acessaram a plataforma e iniciaram o processo de preenchimento do formulário e, ao final, 175 indivíduos preencheram completamente o formulário, sendo esta a amostra final, caracterizada como voluntária e não aleatória.

Para avaliar o nível de maturidade em dados e análises utilizou-se o instrumento adaptado de “CSC Big Data Maturity tool”, desenvolvido pela Computer Science Corporation (CSC) (CSC, 2013). O modelo original avalia as dimensões: Geral, Intenção, Dados, Tecnologia, Processo e Pessoas. Neste estudo, o recorte está circunscrito às dimensões Dados e Tecnologia. A dimensão Dados define a maturidade em atributos como qualidade, relevância, disponibilidade, confiabilidade, governança, segurança e acessibilidade dos dados. Já a dimensão Tecnologia utiliza atributos como a adequação, aplicabilidade, integração, suporte para padrões e desempenho da tecnologia e arquitetura de TI.

Os pontos avaliados dentro da dimensão Dados cobriram modificações no modo como as análises de dados são realizadas na organização; a disponibilidade dos dados para a tomada de decisão e; os tipos de dados analisados. Já em relação

à dimensão Tecnologia, a avaliação abordou a intensidade que a tecnologia existe na organização atende às necessidades para dados e análises; emprego de tecnologia especializada adequada às necessidades de dados e análises; satisfação dos usuários em relação as funcionalidades das tecnologias direcionadas a dados e análises; a velocidade de resposta da equipe de TI aos usuários finais e; incorporação de possibilidades de análises, *BI* ou funcionalidades de apoio à decisão nas aplicações empresariais.

A maturidade foi avaliada em cinco níveis, indo desde o nível mais insipiente até o nível máximo de maturidade. No primeiro nível, denominado *Ad Hoc* ou Esporádico, os dados são provenientes de repositórios internos de fácil acesso e apresentam uma visão incompleta do assunto e requerem esforço manual substancial para se transformar em um formato utilizável para análise ou consumo pelo usuário final. A tecnologia é a existente ou alguma tecnologia aberta, de baixo custo. No segundo nível, denominado Oportunista, os projetos são baseados em dados de várias fontes e podem incluir combinações de dados estruturados, semiestruturados e não estruturados. Os dados estão disponíveis a um número limitado de usuários. A tecnologia é adquirida e implantada para um projeto específico, mas não estão integradas. No terceiro nível, denominado Replicável, existem processos para coleta, monitoramento e integração de dados, mas não existem práticas seguras para governança e segurança dos dados. Várias tecnologias estão disponíveis, mas a adoção é seletiva. No quarto nível, denominado Gerenciado, cresce o apoio à medida que dados e análises são vistos como oportunidades para o desenvolvimento de novos produtos e serviços. Ferramentas de mensuração de desempenho orientam novos investimentos. Várias são ferramentas adotadas e utilizadas. No último nível, denominado Otimizado, as informações são compreensíveis e confiáveis, baseada em dados internos e externos, estruturados, semiestruturados e não estruturados. Apresenta vasta gama de tecnologias e aptidões. Hardwares e softwares otimizados, com um alto nível de automação.

O instrumento de coleta de dados foi traduzido e adaptado à realidade e necessidades do grupo empresarial analisado. Entre os principais ajustes destaca-se a adaptação de termos técnicos, inclusão do item de resposta “Não Sei” em vários itens, quando cabíveis e adaptação de conteúdo de modo a refletir a realidade brasileira. Após o ajuste final do instrumento, ele foi disponibilizado na plataforma

eletrônica SurveyMonkey e a solicitação para participação e o endereço eletrônico do instrumento foram enviados por e-mail para mais de 500 indivíduos das diferentes empresas do grupo. Os dados foram coletados entre os dias 29/12/2014 e 28/01/2015.

Após a coleta, processamento e análise preliminar dos dados quantitativos, construiu-se o roteiro de pesquisa para condução dos grupos focais.

Participaram dos grupos focais 66 indivíduos, divididos em oito grupos. A grande maioria dos participantes dos grupos focais acessou a survey e iniciou o preenchimento. Os grupos focais foram realizados entre os dias 29/01/2015 e 06/02/2015.

Os dados quantitativos recolhidos foram submetidos a diversos tratamentos estatísticos, valendo-se de procedimentos contidos no software estatístico R e os dados qualitativos foram submetidos à análise de conteúdo.

4 Análise dos Resultados

O uso de dados e análises em processos de tomada de decisão representam muitas oportunidades para alavancar ou mesmo redesenhar processos de negócios e até oportunidades para introdução de novos serviços. Contudo, muitos conceitos ainda não estão claros e o processo de tomada de decisão baseados em dados e análises não é simplesmente um caso, uma aplicação, uma única tecnologia ou uma única habilidade. Dados e análises impactam todos os processos de negócios.

Assim, os resultados ora apresentados sobre o nível de maturidade em dados e análises, têm o potencial de subsidiar planos de ação para a organização com vistas à uniformização e alavancagem dos conceitos e uso de tecnologias relacionadas ao assunto.

Participaram da pesquisa quantitativa 10 presidentes/vice-presidentes e superintendentes, 25 diretores, 56 coordenadores, 26 consultores e especialistas e 58 supervisores e analistas.

Quando se observa a participação por áreas operacionais, nota-se que as áreas de Marketing, Operações e Comercial foram as áreas com maior participação, sendo representadas por 42, 34 e 31 indivíduos, respectivamente. Neste aspecto,

destaca-se a moderada participação dos indivíduos da área de Tecnologia da Informação, sendo que apenas 11 indivíduos desta área participaram do processo.

Ao adaptar o instrumento para avaliação de maturidade em dados e análises, optou-se por incluir a opção de resposta “Não Sei” em vários itens. Para a avaliação da dimensão Dados, foram incluídos 17 itens com a opção de resposta “Não Sei”. Dos 175 respondentes, 91 (52%) disseram não saber sobre pelo menos um dos itens avaliados. Em relação à dimensão Tecnologia, foram incluídos 12 itens com a opção de resposta “Não Sei”, sendo que 71 (41%) dos respondentes afirmaram não saber sobre pelo menos um dos itens avaliados.

Ao analisar o nível de maturidade em dados e análises, as respostas “Não Sei” não foram utilizadas no cálculo dos indicadores. Contudo, esta informação, por si só, já é sinalizadora sobre uso e disseminação limitados de dados e análises no processo de tomada de decisão na organização em estudo, sinalizando um estágio inicial de maturidade em dados e análises.

Em contrapartida, os participantes dos grupos focais apresentaram alto grau de interesse e expectativa pelo tema tratado, sinalizando um substrato para direcionamento da empresa para uma empresa baseada em dados. Alguns dos participantes mencionaram que haviam estudado sobre o assunto e se preparado, tanto para participarem da *survey* quanto dos grupos focais.

A dimensão Dados define a maturidade em atributos como qualidade, relevância, disponibilidade, confiabilidade, governança, segurança e acessibilidade dos dados. Nessa dimensão, a nota média consolidada foi 2,95, como se observa na figura 3.

Figura 3 - Maturidade consolidada em dados e análises da dimensão Dados.

Dados: 2,95

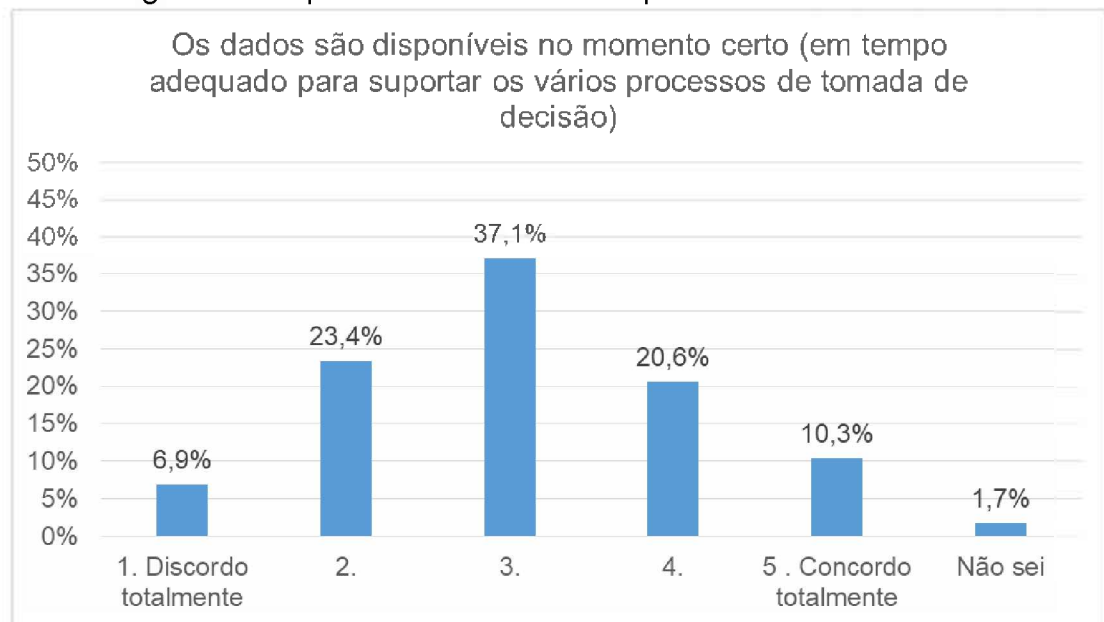


Fonte: Os autores.

A nota média atribuída pelos participantes em relação à maturidade em dados, sugere uma transição entre o nível Oportunista, onde os projetos são baseados em dados de várias fontes e podem incluir combinações de dados estruturados, semiestruturados e não estruturados, mas estes dados estão disponíveis a um número limitado de usuários, para o nível Replicável, em que existem processos para coleta, monitoramento e integração de dados, mas não existem práticas seguras para governança e segurança dos dados.

Ao analisar a variável disponibilidade dos dados para a tomada de decisão, destaca-se a percepção de que os dados não estão disponíveis no momento certo para suportar a tomada de decisão (Figura 4). Mais de 65% dos pesquisados tem esta percepção. Este resultado é contrário ao apontado por Vidgen, Shaw e Grant (2017) e Akter et al. (2016), quando afirmam que para a criação de valor e desenvolvimento de uma vantagem competitiva sustentável é necessário que os dados estejam disponíveis para o tomador(a) de decisão no momento adequado.

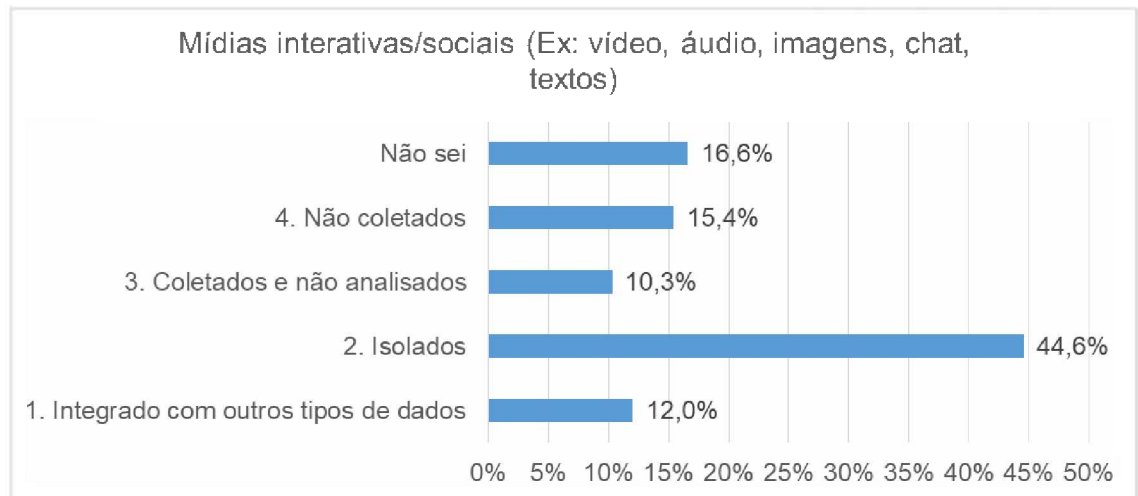
Figura 4 - Disponibilidade dos dados para a tomada de decisão



Fonte: Os autores.

Em relação à variável tipos de dados analisados (Figura 5), observava-se que os dados advindos de mídias interativas/sociais (internos e externos) (44,6%) como vídeos, áudio, imagens, chats, textos, entre outros tipos não-estruturados, são analisados de forma isolada. Também, 15,4% deste tipo de dado não é coletado e para 10,3% dos entrevistados este tipo de dado é coletado, porém não analisado. Tal cenário corrobora com os apontamentos realizados por Prasad e Acharya (2012), ao afirmarem que a coleta, análise e integração de dados não-estruturados se configura como um dos grandes desafios para as organizações.

Figura 5 – Análise de mídias interativas/sociais



Fonte: Os autores.

Na dimensão Tecnologia utiliza-se atributos como a adequação, aplicabilidade, integração, suporte para padrões e desempenho da tecnologia e arquitetura de TI para determinar o nível maturidade. Nesta dimensão, a nota média consolidada do grupo empresarial foi 2,69 (Figura 6). Esta nota média sugere uma transição entre o nível Oportunista, em que a tecnologia é adquirida e implantada para algum projeto específico, mas não estão integradas, para o nível Replicável, em que várias tecnologias estão disponíveis, mas a adoção é seletiva.

Figura 6 - Maturidade consolidada em dados e análises da dimensão Tecnologia

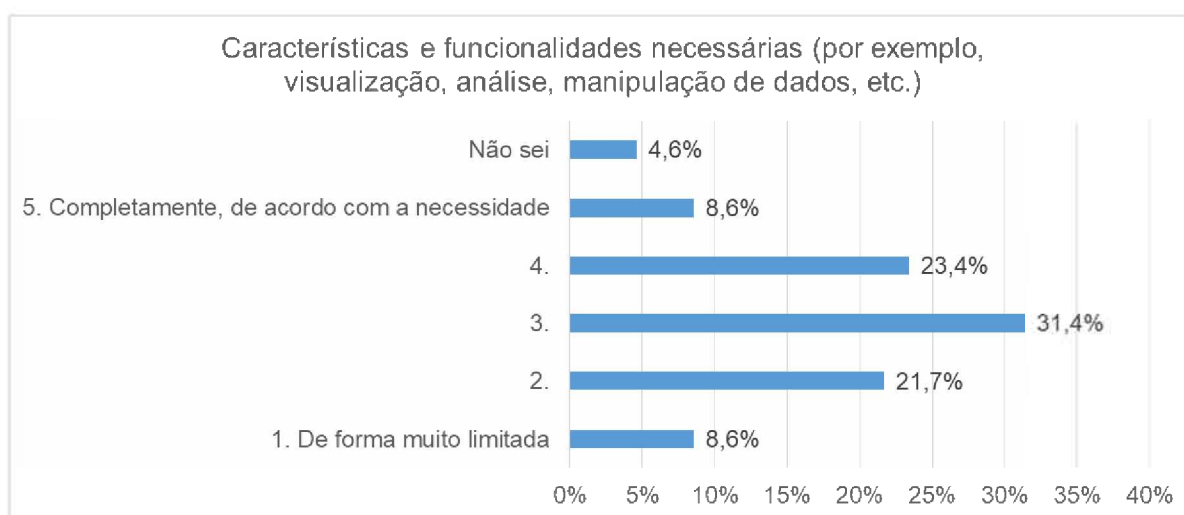
Tecnologia: 2,69



Fonte: Os autores.

Destaca-se na variável características e funcionalidades necessárias para dados e análises nas tecnologias existentes dentro do grupo empresarial a percepção, dos pesquisados, de que as mesmas não atendem as necessidades e demandas exigidas (Figura 7). Aproximadamente 60% dos pesquisados tem esta percepção. Tal resultado é contrário ao que foi apontado por Coa, Duan e Li (2015), na medida em que, o grupo empresarial não se vale dos avanços das tecnologias da informação (TI), para suprir sua necessidade tecnológica voltada a dados e análises.

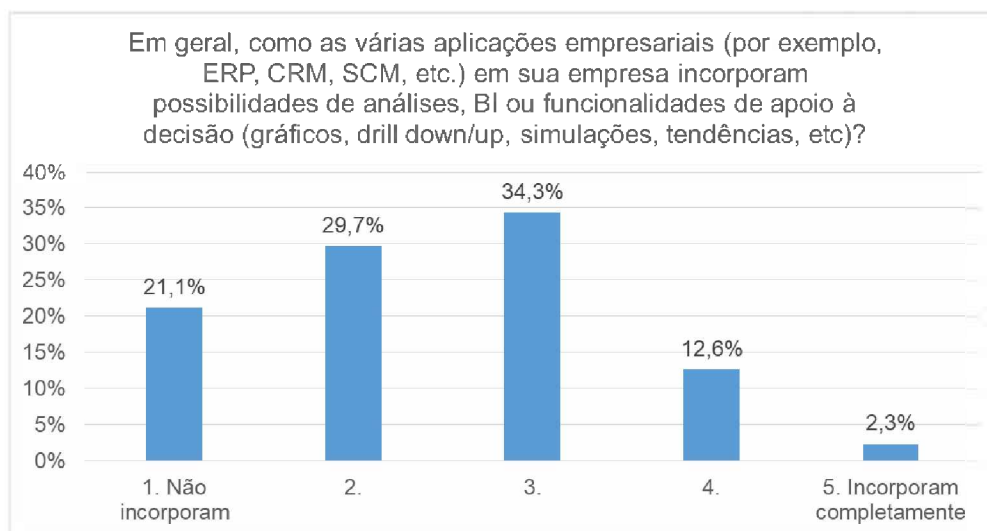
Figura 7 – Intensidade em que a tecnologia existente (características e funcionalidades) atende às necessidades para dados e análises



Fonte: Os autores.

Ao analisar a variável relacionada a incorporação de possibilidades de análises, BI ou funcionalidades de apoio à decisão nas aplicações empresariais (Figura 8), verifica-se que, para 85% dos pesquisados, as aplicações empresariais atuais não incorporam possibilidades relacionadas a dados e análises. O resultado encontrado reafirma o que foi apontado por Eynon (2013), quando diz que embora as tecnologias de BI&A tenham o potencial de oferecer grandes vantagens competitivas, as práticas ainda são tímidas. Ou seja, neste caso, os dados existem nas diversas aplicações empresariais, contudo, as mesmas não possuem capacidade analítica orientada por dados e análises; item fundamental, como apontado por Vidgen, Shaw e Grant (2017) e Akter et al. (2016), para a geração de valor e desenvolvimento de uma vantagem competitiva sustentável.

Figura 8 – Grau de incorporação de possibilidades de análises, BI ou funcionalidades de apoio à decisão nas aplicações empresariais



Fonte: Os autores.

Embora a pesquisa quantitativa tenha apontado para um nível intermediário de maturidade, o nível Replicável, para Dados e Tecnologia, ao aprofundar as discussões nos grupos focais, percebeu-se que as empresas do grupo ainda se posicionam próximas ao menor nível da maturidade em dados e análises, nível este denominado *Ad Hoc* ou *Esporádico* em que, normalmente, os dados para tais projetos são provenientes de repositórios internos de fácil acesso e apresentam uma visão incompleta do assunto e requerem esforço manual substancial para se transformar em um formato utilizável para análise ou consumo pelo usuário final. A tecnologia é a existente ou alguma tecnologia aberta, de baixo custo.

Exceção feita a duas empresas do grupo empresarial, em que o nível de maturidade apontado pela pesquisa quantitativa indicou se confirmar durante as discussões nos grupos focais, embora algumas ressalvas tenham sido feitas, apontando para uma fragilidade de visão estratégica em tratar dados e análises, como pode ser ilustrado pela afirmação de um dos participantes de um grupo focal:

“[...] temos muita informação, mas não temos uma visão estratégica de como trata-las. Analisamos os dados hoje de forma fragmentada, falta uma visão mais ampla, holística desse processo”.

Em uma destas empresas com o maior nível de maturidade, sob o ponto de vista da equipe de TI, tem-se a percepção que o problema não seja tecnológico, mas sim, relacionado às habilidades: “hoje o Data Warehouse da organização está em boas condições sim. Tecnologia não é o ponto que “pega”. Está disponível. A questão é maturidade para analisar os dados, qualificação, pessoas que saibam fazer”. Entretanto, as áreas de negócios consideram que ainda é necessário evoluir em tecnologia para extração e tratamento de dados.

De forma geral, nas diversas empresas do grupo empresarial analisado, existem tecnologias e políticas voltadas para análise e gestão de dados. Contudo, a pesquisa indica que há problemas com os dados internos e externos às empresas do grupo como cadastros inadequados, desatualizados, dados isolados, dados incompletos, disponibilidade no momento correto para suportar a tomada de decisão, coleta, integração e análise de dados não-estruturados, entre outros. Do ponto de vista da tecnologia, observou-se desafios em relação a ferramentas inadequadas - muitas empresas utilizam apenas o Microsoft Excel-, e a incorporação de possibilidades de análises, *BI* ou funcionalidades de apoio à decisão nas aplicações empresariais, entre outros.

Assim, conforme argumenta Coa, Duan e Li (2015), para que haja uma evolução na maturidade em dados e análises, do ponto de vista de dados e tecnologia, do grupo empresarial analisado, há de se investir em tecnologias da informação (TI) apropriadas capazes de coletar, integrar e analisar grandes quantidades de dados (internos, externos, estruturados, semiestruturados e não-estruturados), assim como, permitir a incorporação de soluções analíticas nas aplicações empresariais existentes.

5 Considerações Finais

De uma forma geral, avalia-se que os respondentes, na etapa quantitativa, superestimaram o nível de maturidade em dados e análises de suas respectivas empresas. Ao confrontar os resultados quantitativos com os resultados qualitativos, em profundidade, pôde-se perceber que em várias empresas as iniciativas em dados e análises ainda são incipientes, tendo um caminho a ser trilhado e desafios a serem superados.

Ainda, no levantamento quantitativo, destaca-se o nível elevado de respondentes que atribuíram a opção de resposta “Não Sei” a vários itens avaliados.

Este fato reforça a constatação da análise em profundidade, evidenciando níveis iniciais do processo de uso de dados e análises no processo de tomada de decisão.

Além da avaliação do nível de maturidade em dados e análises como direcionadores do processo de tomada de decisão, sob a perspectiva de dados e tecnologia, considera-se que esta avaliação, por si só, já tenha contribuído de forma substancial para o processo de conscientização do grupo empresarial em Dados e Análises. Relatos apontam que vários associados buscaram informações em diversas fontes para entenderem e se familiarizarem com os assuntos e termos em pauta, além de se prepararem para participarem dos grupos focais.

Sob o ponto de vista teórico, esta pesquisa buscou direcionar esforços para os benefícios potenciais das empresas orientada por dados e análises (*data-driven companies*). Como apontado Vidgen, Shaw e Grant (2017), Akter et al. (2016) e Janssen, Voort e Wahyudi (2016), por meio da gestão e análise de dados é possível criar valor para os *stakeholders* e desenvolver vantagem competitiva sustentável; o que resultará, de acordo com McAfee e Brynjolfsson (2012), em organizações mais produtivas e lucrativas.

Como sugestão para trabalhos futuros, destaca-se a possibilidade de replicar tal estudo a um número maior de organizações, permitindo compreender o nível de maturidade em dados e análises das organizações locais, estaduais e nacionais, públicas e privadas.

Data-Driven Companies: is technology at the service of decision-making?

Abstract

The intense process of creating digital data (internal and external) has translated into challenges and opportunities to the organizations, as long as extracting relevant information to support decision making requires changes in their models, processes and business vision. By making data-driven decisions, organizations aim to generate value for their stakeholders and develop sustainable competitive advantages. Thus, the objective of this paper was to analyze the level of data maturity and analysis as drivers of the decision-making process in a business group, from a data and technology perspective. A data maturity and analysis model developed by the Computer Science Corporation (CSC) was adapted for the study. A mixed research approach was used in which, in the first stage, a quantitative study, a survey was carried out with 175 participants and, in the second stage, a qualitative study, eight focal groups were carried out, reaching a total of 66 participants. The quantitative results point to an intermediate level of maturity, the Repeatable level. However, the qualitative results point to the basic level of maturity, named Ad Hoc or Sporadic. In both cases, the results suggest the need for investments in appropriate technologies capable of collecting, integrating, analyzing data and making it available for the decision makers as well as allowing the incorporation of analytical solutions in the existing business applications.

Keywords: Big Data; Data and Technology Maturity; Decision-Making.

6 Referências

- AKTER, S. et al. How to Improve Firm Performance Using Big Data Analytics Capability and Business Strategy Alignment?. **International Journal of Production Economics**, v. 182, p. 113-131, 2016.
- BHARADWAJ, S. G.; VARADARAJAN, P. R.; FAHY, J. Sustainable Competitive Advantage in Service Industries: a conceptual model and research propositions. **Journal of Marketing**, v. 54, n. 4, p. 83-99, 1993. ABI/INFORM Global.
- BLACK, J. A.; BOAL, K. B. Strategic Resources: traits, configurations and paths to sustainable competitive advantage. **Strategic Management Journal**, v. 15, p. 131-148, 1994. Issue Supplement S2.
- BOYD, D.; CRAWFORD, K. Critical Questions for Big Data: provocations for a cultural, technological, and scholarly phenomenon. **Information, Communication & Society**, v. 15, n. 5, p. 662-679, jun. 2012.

CHEN, C. L. P.; ZHANG, C-Y. Data-Intensive Applications, Challenges, Techniques and Technologies: a survey on Big Data. **Information Sciences**, v. 275, p. 314-347, 2014.

CHEN, H.; CHIANG, R. HL.; STOREY, V. C. Business Intelligence and Analytics: from big data to big impact. **MIS quarterly**, v. 36, n. 4, p. 1165-1188, 2012.

COA, G.; DUAN, Y.; LI, G. Linking Business Analytics to Decision Making Effectiveness: a path model analysis. **IEEE Transactions on Engineering Management**, v. 62, n. 3, p. 384-395, 2015.

CÔRTE-REAL, N.; OLIVEIRA, T.; RUIVO, P. Assessing Business Value of Big Data Analytics in European Firms. **Journal of Business Research**, v. 70, p. 379-390, 2016.

CSC - Computer Sciences Corporation. **Big Data Maturity Model: Planning Your Strategy**. 2013. Disponível em: <http://www.csc.com/big_data/insights/111325-big_data_maturity_model_planning_your_strategy>. Acesso em: 08 fev. 2016.

DUMBILL, E. What Is Big Data? An Introduction to the Big Data Landscape. **Strata 2012: Making Data Work**, 2012.

EYNON, R. The Rise of Big Data: what does it mean for education, technology, and media research?. **Learning, Media and Technology**, v. 38, n. 3, p. 237-240, 2013.

GARSHOL, L. M. **Introduction to Big Data/Machine Learning**, 2013. Disponível em: <<http://www.slideshare.net/larsga/introduction-to-big-datamachine-learning>>. Acesso em: 01 fev. 2016.

GROBELNIK, M. Big Data Tutorial. **European Data Forum**, 2013. Disponível em: <<http://pt.slideshare.net/EUDataForum/edf2013-big-datatutorialmarkogrobelnik?related=1>>. Acesso em: 01 fev. 2016.

GUPTA, B.; GOUL, M.; DINTER, B. Business Intelligence and Big Data in Higher Education: status of a multi-year model curriculum development effort for business school undergraduates, MS graduates, and MBAs. **Communications of the Association for Information Systems**, v. 36, 2015.

HALPER, F.; KRISHNAN, K. TDWI. Big Data Maturity Model Guide. TDWI, 2013.
KITCHIN, R. **The data revolution: big data, open data, data infrastructures and their consequences**. Sage Publications, London, 2014.

INTERNATIONAL DATA CORPORATION (IDC). **The Digital Universe of Opportunities: rich data & the increasing value of the internet of things**, 2014. 17 p. Infobrief. Disponível em: <<https://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>>. Acesso em: 5 nov. 2017.

JANSSEN, M.; VOORT, H, V der.; WAHYUDI, A. Factors influencing big data decision-making quality. **Journal of Business Research**, v. 70, p. 338-345, 2016.

LOPES, J. E. F.; SOARES, D. J. Empresas Orientadas a Dados e Análises: as pessoas estão preparadas?. In: Encontro de Gestão e Negócios - EGEN, 2016, Uberlândia - MG. **Anais**...Uberlândia, MG.

MCAFEE, A.; BRYNJOLFSSON, E. Big Data: the management revolution. **Harvard Business Review**, out. 2012. Decision Making. Disponível em:

<<https://hbr.org/2012/10/big-data-the-management-revolution>>. Acesso em: 10 out. 2017.

MCKINSEY GLOBAL INSTITUTE. **The Age of Analytics: competing in a data-driven world**, 2016. 136 p. Report. Disponível em: <<https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world>>. Acesso em: 5 nov. 2017.

NEGASH, S. Business Intelligence. **Communications of the Association for Information Systems**, v. 13, 2004.

POPOMARONIS, T. Here's an Incredible Breakdown of What Happens on the Internet (Every 60 Seconds). **Inc.**, 4 maio 2017. Innovate. Disponível em: <<https://www.inc.com/tom-popomaronis/every-60-seconds-an-incredible-breakdown-of-what-happens-on-the-internet.html>>. Acesso em: 18 set. 2017.

PRASAD, R. N.; ACHARYA, S. **Fundamentals of Business Analytics**. New Delhi: Wiley India, 2012.

RAMIREZ, R. Appreciating the Contribution of Broadband ICT with Rural and Remote Communities: stepping stones toward an alternative paradigm. **The Information Society**, v. 23, p. 85-94, 2007.

RICHARDSON, R. J. **Pesquisa Social: métodos e técnicas**. São Paulo: Atlas, 1999.

SHIN, D. H.; CHOI, M. J. Ecological views of big data: perspectives and issues. **Telematics and Informatics**, v. 32, n. 2, p. 311-320, 2015.

SORESCU, A. Data-Driven Business Model Innovation. **Journal of Product Innovation Management**, v. 34, p. 691-696, 2017

TINATI, R. et al. Big data: methodological challenges and approaches for sociological analysis. **Sociology**, v. 48, n. 1, p. 23-39, 2014.

VIDGEN, R.; SHAW, S.; GRANT, D. B. Management challenges in creating value from business analytics. **European Journal of Operational Research**, v. 261, p. 626-639, 2017.

WATSON, H. J. Tutorial: Business Intelligence - past, present, and future. **Communications of the Association for Information Systems**, v. 25, 2009.

WEDEL, M.; KANNAN, P, K. Marketing Analytics for Data-Rich Environments. **Journal of Marketing**, v. 80, p. 97-12, 2016. AMA/MSI Special Issue.

ZIKOPOULOS, P. C. et al. **Understanding Big Data: analytics for enterprise class Hadoop and streaming data**. McGraw-Hills, 2012.

ZINS, C. Conceptual Approaches for Defining Data, Information, and Knowledge. **Journal of the American Society for Information Science and Technology**, v. 58, p. 479-493, 2007.