

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
DEPARTAMENTO DE CIENCIA DA INFORMAÇÃO
CURSO DE GESTÃO DA INFORMAÇÃO

BRENNA NICOLE JURCHACKS SANTOS DO AMARAL

BIG DATA:

Um *benchmark* de grande volume de dados de serviços de internet

RECIFE
2017

BRENNA NICOLE JURCHACKS SANTOS DO AMARAL

BIG DATA:

Um *benchmark* de grande volume de dados de serviços de internet

Trabalho de Conclusão de Curso do Curso de Gestão da Informação do Departamento de Ciência da Informação do Centro de Artes e Comunicação da Universidade Federal de Pernambuco como requisito para obtenção do Grau em Bacharel em Gestão da Informação.

Professor Orientador Dr.: Celio Andrade de Santana Junior

RECIFE
2017

Catálogo na fonte
Bibliotecário Jonas Lucas Vieira, CRB4-1204

A485b Amaral, Brenna Nicole Jurchacks Santos do
 Big data: um benchmark de grande volume de dados de serviços de internet / Brenna Nicole Jurchacks Santos do Amaral. – Recife, 2017.
 54 f.: il., fig.

 Orientador: Célio Andrade de Santana Júnior.
 Trabalho de Conclusão de Curso (Graduação) – Universidade Federal de Pernambuco. Centro de Artes e Comunicação. Ciência da Informação, 2017.

 Inclui referências e glossário.

 1. *Big data*. 2. *Benchmark*. 3. Gestão da informação. 4. Dados. 5. Tomada de decisão. I. Santana Júnior, Célio Andrade de (Orientador). II. Título.

020 CDD (22.ed.)

UFPE (CAC 2017-160)

DEDICATÓRIA

*Dedico este trabalho a todos que
ajudaram diretamente ou indiretamente na
minha formação acadêmica.*

AGRADECIMENTOS

*Agradeço a todos que contribuíram no
decorrer desta jornada, em especialmente:*

A Deus, a quem devo minha vida.

*A minha família que sempre me apoiou
nos estudos e nas escolhas tomadas.*

*Ao orientador Prof. Dr. Celio Andrade de Santana Júnior
que teve papel fundamental na elaboração
deste trabalho e em minha formação acadêmica.*

*Aos meus amigos e colegas pelo companheirismo e
disponibilidade para me auxiliar em vários momentos.*

*“Se você quer ser bem sucedido,
precisa ter dedicação total, buscar seu
último limite e dar o melhor de si mesmo.”*

Ayrton Senna

RESUMO

As empresas convivem com grandes volumes de informações geradas diariamente, que são dados muito importantes para cada uma, porém apenas ter os dados não basta mais, o importante é saber como usá-los, assim é a partir deste visão que o

conceito de *big data* é aplicado. O *big data* é caracterizado como sendo conjuntos de dados extremamente amplos e que, por este motivo, necessitam de ferramentas especialmente preparadas para lidar com grandes volumes, de forma que toda e qualquer informação nestes meios possa ser encontrada, analisada e aproveitada em tempo hábil. Assim, podemos classificar como uma análise de grande quantidade de informação transformada e simplificada em um volume menor, para um melhor entendimento na tomada de decisão. A Informação é poder, logo, se uma empresa/gestor souber como utilizar os dados que tem em mãos, poderá entender como melhorar um produto, como criar uma estratégia de marketing mais eficiente, como cortar gastos, produzir mais em menos tempo, evitar o desperdício de recursos, superar um concorrente, disponibilizar serviços para a um cliente especial de maneira satisfatória e assim por diante. As tecnologias atuais nos permitiram e permitem aumentar exponencialmente a quantidade de informações no mundo, e agora, empresas, governos e outras instituições precisam saber lidar com esta "explosão" de dados. O *Big Data* se propõe a ajudar nesta tarefa, uma vez que as ferramentas computacionais usadas até então para gestão de dados, por si só, já não podem fazê-lo satisfatoriamente.

Palavras-chave: *Big data. Benchmark. Gestão da Informação. Dados. Tomada de decisão.*

ABSTRACT

The companies coexist with large volumes of information generated daily, which are very important data for each one, however only having the data is not enough anymore, the important thing is to know how to use them, so it is from this view that the concept of big data is applied. Big data is characterized as being extremely large

data and therefore requires specially prepared tools to handle large volumes so that any and all information in these media can be found, analyzed and exploited in a timely manner. So, we can classify as an analysis of large amount of information transformed and simplified into a smaller volume, for a better understanding decision making. Information is power, so if a company/manager knows how to use the data in hand, they can understand how to improve a product, how to create a more efficient marketing strategy, how to cut expenses, how to produce more in less time, how to avoid wasting resources, overcoming a competitor, how to provision services to a particular customer satisfactorily, and so on. Today's technologies have allowed and allowed us to exponentially increase the amount of information in the world, and now companies, governments and other institutions need to be able to deal with this "data explosion". Big Data proposes to help in this task, since the computational tools used until then for data management, by itself, can no longer do it satisfactorily.

Keywords: Big data. Benchmark. Information management. Data. Decision.

LISTA DE ILUSTRAÇÕES

Figura 1: Centro de operações <i>big data</i> Obama	20
Figura 2: Análise dos dados 3.0	21

Figura 3: Caracterizando o <i>big data</i> de acordo com os 3 Vs: Volume, Velocidade e Veracidade	24
Figura 4: Os cientistas de dados da IBM dividem grandes dados em quatro dimensões: Volume, variedade, velocidade e veracidade. Infográfico explicando exemplos de cada	25
Figura 5: Analisando os <i>insights</i> comerciais específicos	40

LISTA DE TABELAS

Tabela 1: Conceitos de <i>big data</i>	18
---	----

Tabela 2: Evolução da utilização dos dados e informação..... 42

LISTA DE SIGLAS

ABNT – Associação Brasileira de Normas Técnicas

CRM - *Customer Relationship Management*

HDFS - *Hadoop File System.*

HDFS - *Hadoop Distributed File System*

NASA - *National Aeronautics and Space Administration*

RFID - *Radio frequency identification*

TI – *Tecnologia da Informação*

SUMÁRIO

1 INTRODUÇÃO	12
1.2 OBJETIVOS	14
1. 2.1 OBJETIVO GERAL	14
1. 2.2 OBJETICO ESPECIFICOS	14
2 METODOLOGIA DO TRABALHO	15
3 <i>BIG DATA</i>	17
3.1 OS Vs DO <i>BIG DATA</i>	23
3.2 IMPACTO DO <i>BIG DATA</i>.....	26
4 <i>BENCHMARK</i>: AS ABORDAGENS CONCEITUAIS	28
4.1 TIPOS DE <i>BENCHMARKING</i>.....	29
4.2 NECESSIDADES E TÉCNICAS DO <i>BENCHMARK</i>	31
5 O <i>BIGA DATA</i> COMO FORMA DE <i>BENCHMARK</i>	36
5.1 VANTAGENS DA UTILIZAÇÃO DO <i>BIG DATA</i> COMO <i>BENCHMARK</i> ...	39
6 CONSIDERAÇÕES FINAIS	43
7 REFERENCIAS	45
GLOSSÁRIO.....	52

1 INTRODUÇÃO

“Quanto mais ampla e profunda a difusão da tecnologia da informação avançada em fábricas e escritórios, maior a necessidade de um trabalhador instruído” (CASTELLS, 1999, p. 306). Com a popularização da informática uma grande quantidade de dados tem expandido a necessidade de analisar grandes conjuntos de dados e esta competência se tornou uma base fundamental da competição que só aumenta, sustentando novas ondas de crescimento da produtividade, inovação e superávit de consumidores. Empresas de todos os setores terão que lidar com as implicações dos grandes dados, e não apenas alguns gerentes de dados orientados ou algum setor específico.

Berardi (2005) chama atenção para o perfil desse novo trabalhador:

“[...] enquanto o trabalhador industrial colocava no serviço assalariado suas energias mecânicas, seguindo um modelo repetitivo, despersonalizado, o trabalhador *high tech* empenha na produção sua competência singular, suas energias comunicativas, inovadoras, criativas, em suma, o melhor de suas capacidades intelectuais.” (BERARDI, 2005, p.41).

Assim, é preciso que o mercado se mantenha atualizado, inovador e sempre buscando novos conhecimentos para se reestruturar e manter sempre os bons desempenhos. Isso cada vez mais se passa em um ambiente onde o grande volume de dados, bem como pequenos detalhes que as envolvem, são continuamente obtidas por empresas. Diversos são os formatos para estas informações, sejam multimídias, dados oriundos de mídias sociais e até mesmo de dispositivos conectados pela *internet* das coisas, contribui com o enorme crescimento dos dados coletados para as análises futuras.

Porém, para a transformação destes grandes volumes de dados em *insights* informacionais, é exigido uma extração em profundidade de seus valores, que depende fortemente de um conhecimento prévio e de outros sistemas. Esse ciclo torna a própria análise de dados como insumo, e não como fim, para o aumento de

implantações de grandes sistemas que suportam grandes volumes de dados, tornando caótico o crescimento informacional. A atuação das organizações deve ser bastante intensificada pois a capacidade de processamento dos computadores cresce nove vezes a cada década e o crescimento não para, “A quantidade de dados no mundo cresce rapidamente e supera não apenas nossas máquinas como nossa imaginação” (MAYER-SCHONBERGER, CUKIER, 2013, p. 5).

À medida que as comunidades de gerenciamento de dados prestam maior atenção aos sistemas e arquitetura inovadoras de dados grandes, a pressão de medir, comparar e avaliar esses sistemas aumenta e conseqüentemente, a quantidade numerosa de sistemas impõem grandes desafios na escolha da melhor solução fazendo com que os *benchmarkings* ganhem importância. Neste contexto, o *big data* tem um papel ímpar neste processo de formação de *benchmarks* pois ele tem a capacidade de prover a percepção e compreensão entre informações, que até há pouco tempo havia dificuldade de entender. (MAYER-SCHONBERGER, CUKIER, 2013, p. 23).

É a partir desta visão que o *benchmarking* surge em um contexto organizacional de competição e dinamicidade em que as empresas precisam se adaptar. É uma forma de inovação na busca de estar sempre à frente de seus concorrentes, e assim, o *benchmarking* é uma técnica de extrema importância para o entendimento sobre como e onde se origina e como pode garantir o seu uso de modo eficaz visando melhorias organizacionais. Davenport (2013, p. 36) lembra o uso de dados não é algo de hoje, mas sim tão antigo quanto a própria tomada de decisão em si. Porém, o que deve se levar em consideração nos dias de hoje é que a inteligência dos negócios baseado na análise dos dados torna-se condição de competitividade quando ocorre investimento adequado e eficaz nos dados produzidos por clientes internos e externos das organizações, criando uma relação econômica.

Investimentos feitos para aumentar a competitividade são primordiais para o desenvolvimento econômico empresarial, seja qual for o ramo e empresa. Aprender a gerenciar os negócios e conhecer os seus concorrentes é a melhor prática de desenvolvimento de inovação e concorrência, “ Usar a participação de mercado relativa para medir os recursos fortes competitivos é analiticamente superior do que

adotar diretamente a participação de mercado em termos percentuais”. (THOMPSON, 2011 p.290).

A visão de Thompson (2011) apresentada no parágrafo anterior é que motiva o entendimento de como o *benchmark* pode ser aprimorado pelo *big data*. Se o *benchmark* tem como principal função prover uma base de comparação, a utilização de elementos de *big data* pode estender o uso de tais *benchmarks* para os níveis de eficiência proporcionados por estes.

1.1 OBJETIVOS

1.1.1 OBJETIVO GERAL

Identificar como o *big data* pode servir como *benchmark* de grande volume de dados de serviços de *internet*.

1.1.2 OBJETIVOS ESPECIFICOS

- Analisar o *big data* e a técnica do *benchmark* a partir de abordagens conceituais, históricas e tipológicas;
- Explicar como o *big data* pode ser aplicado como *benchmark*;
- Identificar outros métodos, técnicas e ferramentas de *big data* que podem potencializar o *benchmark*;
- Entender como *big data* pode auxiliar no *benchmark*, identificando qual o melhor tipo e como este se ajusta ou atende a necessidade da organização.

2 METODOLOGIA DE TRABALHO

Esta pesquisa foi realizada a partir de um levantamento bibliográfico cujo objetivo era recolher informações ou conhecimentos prévios sobre a relação entre *benchmarks* e *big data*. Esta revisão não seguiu nenhum método sistemático formal adota em estudos secundários e foi realizada de maneira *ad-hoc* a partir de artigos, livros e sites especializados. Por se tratar de um tema relativamente recente, parte considerável das fontes de informação utilizada neste trabalho veio de blogs e sítios de notícias especializados no assunto.

Esta pesquisa é de abordagem qualitativa uma vez que, segundo Michel (2009) pesquisas qualitativas auxilia nas interpretações de fenômenos a luz de um contexto e é baseada na interpretação do pesquisador sobre o fenômeno investigado.

Também é possível caracterizar esta pesquisa como exploratória que de acordo com Michel (2009, p.40) “visam proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a construir hipóteses”. Gil (2010, p.35) aponta que objetivo principal do estudo exploratório é o aprimoramento de ideias ou descoberta de intuições.

Michel (2009, p.41) afirma que pesquisas deste tipo podem ser feitas isoladamente na forma de revisão de literatura cujo objetivo é aumentar o conhecimento do pesquisador sobre o tema a partir de um olhar teórico sobre o tema observando novas abordagens e visões do assunto.

Como o objetivo geral desta pesquisa é analisar a relação entre *big data* e *benchmark* podemos afirmar que esta pesquisa possui esse caráter exploratório. Uma parte considerável desta pesquisa foi realizada a partir da revisão da literatura realizada.

Esta pesquisa também pode ser considerada como básica que segundo Gil (2010) é aquela que não tem aplicabilidade imediata e generalista. A pesquisa básica procura os princípios e fundamentos do mundo, das coisas, seu

funcionamento e sua intenção e desvendar características, propriedades básicas do fenômeno.

Segundo Michel (2009, p. 52), os métodos específicos têm por objetivos proporcionar ao investigador os meios técnicos de garantir a objetividade e a precisão do estudo e esse método também fornece a orientação necessária à realização da pesquisa social, sobretudo no que se refere à obtenção, processamento, análise e validade dos dados pertinentes à problemática que está sendo investigada. E aqui será realizada uma revisão de literatura ad-hoc onde será feito um fichamento para se coletar as informações pertinentes e uma análise do texto agrupando partes correlatas do texto.

3 **BIG DATA**

Uma definição para *big data* é dada por Mayer-Schonberg e Cukier (2013) que se refere a este como o processo de extração da informação de qualidade a partir de um grande volume de informação. Para extrair novas ideias e criar outra forma de valor em auxílio as empresas, esse processo marca o início de uma importante transformação, pois este vem se mostrando como uma nova forma de economia.

Mcafee e Brynjolfsson (2012) conduziram estudos que levaram à conclusão de que as empresas que efetivamente utilizam *big data* são 5% mais produtivas e 6% mais lucrativas que seus competidores, assim é visível que esses resultados se mostram como estímulos para a utilização dessa abordagem.

Dumbill (2012), relata em seu livro que o surgimento de grande volume de dados na empresa resulta em uma importante necessidade: A agilidade. Explorar com sucesso o valor de grande volume de dados requer experimentação e exploração, assim a criação de novos produtos ou a procura de inúmeras maneiras de ganhar vantagem competitiva invariavelmente passa pelo uso do *big data*.

“*Big data* é um termo relativamente recente (ou, ao menos, começou a aparecer na mídia recentemente) [...] Há tempos que departamentos de TI contemplam aplicações de *Data Mining*, *Business Intelligence* e CRM (*Customer Relationship Management*), por exemplo, para tratar justamente de análise de dados, tomadas de decisões e outros aspectos relacionados ao negócio.” (Emerson Alecrim, 2015). *Big data* vem ganhando força na atualidade devido ao grande volume de informações produzido pelos sistemas atuais e pelo progressivo aumento da capacidade de armazenamento dos dados. “A história de como os dados se tornaram grandes começa muitos anos antes da atual febre em torno do conceito *big data*. Há 70 anos atrás, encontramos as primeiras tentativas de quantificar a taxa de crescimento do volume de dados ou o que popularmente é conhecida como a explosão da informação.” (Gil Press da Forbes Magazine, 2013).

“Não há consenso do conceito de *big data* utilizado no mercado, pois cada organização considera esse fenômeno por alguma perspectiva, no entanto, a essência do termo é contemplada por muitas organizações referências” (Franks, 2012). A tabela seguinte apresenta algumas dessas definições:

Tabela 1 – Conceitos de *big data*

Organização / Autor	Descrição do conceito
GartnerGroup(2012)	“Big Data, em geral, é definido como ativos de alto volume, velocidade e variedade de informação que exigem custo-benefício, de formas inovadoras de processamento de informações para maior visibilidade e tomada de decisão”.
IDC – <i>Internacional Data Corporation</i> (2011)	“...as tecnologias de Big Data descrevem uma nova geração de tecnologias e arquiteturas projetadas para extrair economicamente o valor de volumes muito grandes e de uma variedade de dados, permitindo alta velocidade de captura, descoberta, e/ou análise”.
Cezar Taurion (2013)	“Big Data = volume + variedade + velocidade”.
IBM (2014)	“Big Data é o termo utilizado para descrever grandes volumes de dados e que ganha cada vez mais relevância à medida que a sociedade se depara com um aumento sem precedentes no número de informações geradas a cada dia”.

Fonte: Tabela criada pelos autores Tiago Volpato, Ricardo Ribeiro Rufino, Jaime William Dias da UNIPAR no artigo: *BIG DATA – TRANSFORMANDO DADOS EM DECISÕES*

Até os anos 2000 essas áreas tinham conhecimento de grandes estruturas de dados porém não possuíam meios (*hardware*) para armazenar ou manipular. Um exemplo disto são projetos da NASA para o estudo do Universo. O termo *big data* nasceu no início da década de 1990, na NASA, para descrever grandes conjuntos de dados complexos que desafiam os limites computacionais tradicionais de captura, processamento, análise e armazenamento informacional.” (Rodrigo Arrigoni, 2013).

Por alguns anos, o *big data* foi um conceito difícil de ser colocado em prática devido a restrições técnicas. Somente empresas ligadas a projetos governamentais tinham algumas implementações do conceito. Hoje com a evolução do processamento e armazenamento, e com a redução de custos, o conceito *big data*

pode ser aplicado em diversas áreas com destaque para a determinação de perfis de clientes para todas as fases do ciclo de crédito.

Para compreender melhor o *big data*, é proposto um detalhamento de suas características através de seus pilares e dimensões. O conceito de *big data* veio para estabelecer um novo patamar no processo de análise de dados. A atual tecnologia propicia, a criação de um grande número de informações, baseadas no comportamento de uma empresa ou indivíduo. Informações como: Operações com o cartão de crédito, fluxo de caixa/conta corrente, comportamento na *internet*, podem determinar escolhas, padrões de consumo, fraude, entre outros.

Moraes (2012), informa e explica como a aplicação de *big data* ajudou na campanha de reeleição do presidente norte-americano, Barack Obama, ajudando a orientar os voluntários, indicando as melhores formas de arrecadar fundos e apontar quem poderia ser convencido a apoiar a reeleição do presidente, sendo de imensa utilidade todas as informações obtidas e manuseadas. Assim, é visível as possibilidades de aplicações e impactos, como a análise de falhas em tempo real em um processo de fabricação de qualquer produto ou serviço, ou de como o exemplo informado: As estratégias empregadas na última campanha eleitoral à presidência norte-americana como pode ser visto na Figura 1.

Em departamentos separados, em um tecnológico e analítico seguimento, comandado pelo diretor de Inovação e Integração Michael Slaby. Ele informa sobre a campanha realizada da eleição do ex presidente Barack Obama, assim conta que:

“[...] Um dos mais interessantes foi uma plataforma chamada “dashboard”, uma maneira de conectar e unificar as atividades *off-line* e *on-line* de nossos voluntários. Todo o trabalho que cada um estava fazendo era registrado em um só lugar, como um escritório de campanha *on-line*. Os voluntários podiam se conectar, conhecer seus vizinhos que também faziam parte da campanha, saber quem eram os líderes locais e trabalhar juntos como um time. Também fizemos ferramentas *on-line* de registro de eleitores e manuais para ajudar nossos voluntários a usar o *Facebook* de maneira mais eficiente para atingir mais pessoas. [...]” (Trecho da entrevista de Slaby concedida a revista IstoÉ em 11 de junho de 2013 para a jornalista Juliana Tirobaschi)

Figura 1: O centro de operações do *big data* Obama



Fonte: *Inside the Cave* (2012)

Slaby ficou conhecido como o “guru” digital de Obama, chefe de tecnologia da campanha que levou o então senador (Barack Obama) a seu primeiro mandato presidencial, em 2008. Em 2011, atuou como líder do departamento de inovação e integração da campanha de reeleição de Obama.

A tecnologia e a análise de dados tornaram nossos voluntários mais eficientes, pois permitiu que se comunicassem melhor uns com os outros e evitassem perder tempo indo atrás dos alvos errados. Não tínhamos a intenção de construir a tecnologia mais avançada, mas usá-la para tornar a campanha mais eficiente. (SLABY, entrevista para a Revista Isto É, 2013).

Assim foi possível, por exemplo, prever com 15 dias de antecedência que o desempenho de Obama no primeiro debate não seria satisfatório, permitindo que a equipe planejasse uma ação *on-line*, que foi ao ar uma hora depois do programa.

No *Facebook*, a campanha de Obama teve 3 milhões de curtidas, pulverizando todos os recordes da maior rede social do mundo (Stivanin, 2012). Segundo a avaliação do *The Guardian* (2012) a campanha de Obama usou o poder do *Facebook* para direcionar os eleitores individuais. A rede estava com isso se transformando pela primeira vez numa grande ferramenta de campanha que tinha o potencial de transformar a amizade em uma arma política.

Com a utilização do *big data* para a melhoria dos processos é importante interpretar os dados, e se o que eles estão informando faz sentido para o estudo. Assumir todas as respostas como verdades é um risco, pois é preciso entender o motivo do comportamento. Nas duas últimas décadas o principal desafio do *big data* era seu armazenamento e processamento. Hoje vivemos em uma época em que os recursos tecnológicos permitem essas tarefas a um custo acessível e a um grande número de empresas. “É a ponta do *iceberg* de uma tendência maior que a gente está vivendo na sociedade hoje. É a tendência de automação do processo de produção de bens de informação... desde os dados brutos até os resultados das análise [...] *Big data* é uma parte deste contexto.” (Thoran).

Na Figura 2, Davenport relaciona e informa as características aos dados da era 3.0 enquadrando as empresas:

Figura 2 – Análise de Dados 3.0

Ideia em resumo		
Uma nova era	Por que é importante	O que fazer
As empresas que lidam com a análise de dados viram tudo mudar quando os grandes volumes de dados chegaram. Agora, outra grande mudança está em curso, como a ênfase na construção de poder de análise em produtos e serviços ao cliente.	A análise dos dados é usada para adicionar mais valor ao permitir que gestores tomem melhores decisões internas. O novo foco estratégico na entrega de valor para os consumidores tem profundas implicações sobre onde as funções de análise de dados se encontram dentro das organizações e o que devem fazer para ter sucesso.	Na era da análise de dados 3.0, os gestores devem dirigir esforços em pelo menos dez frentes, de combinar de forma criativa estratégias de gerenciamento de dados até moldar novos cargos de análise e definir diretrizes para responder a “sinais de fumaça digitais”.

Fonte: Davenport (2013, p. 37)

Davenport (2013, p. 36) informa que deve valer a partir de agora é a inteligência dos negócios na análise dos dados tornando assim uma condição de competitividade no mercado. Davenport (2013, p. 39) pontua dez requisitos para capitalizar a análise de dados 3.0 (a era da oferta dos dados enriquecidos, segundo o autor):

1. Vários tipos de dados, muitas vezes combinados;
2. Um novo conjunto de opções de gerenciamento de dados;
3. Tecnologias mais rápidas e métodos de análise;
4. Análise de dados incorporada;
5. Descoberta de dados;
6. Equipes interdisciplinares de dados;
7. Diretores de análise;
8. Análise prescritiva;
9. Análise de dados em escala industrial e;
10. Novas formas de decisão e de gestão.

O novo desafio do *big data* atualmente é o de como extrair informação de qualidade desses volumes. Isto implica a aplicação e análise de mineração de dados por especialistas. A utilização da informação dos dados provenientes de ambientes de *big data* deve ser continuamente analisada para garantir a qualidade dos processos. Assim, a maneira utilizada para pensar a questão, proposta por Schonberger-Mayer e Cukier (2013) é que *big data* se refere a trabalhos em grande escala que não podem ser feitos em escala menor, para extrair novas ideias e criar novas formas de valor de maneira que se alterem os mercados, as organizações, a relação entre cidadãos, governos e etc.

3.1 OS VS DO *BIG DATA*

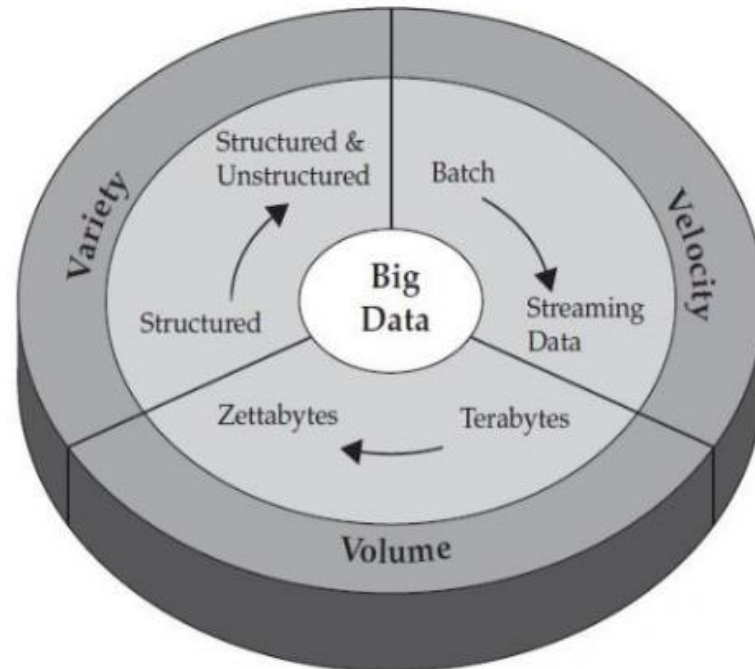
A necessidade de vencer o desafio, reunindo e analisando fontes de diversas naturezas, deu origem a pesquisas que nos levaram ao tema “*Big Data*”. Estas pesquisas foram desenhadas a partir de três aspectos iniciais (Davenport, 2014):

1. A múltipla natureza dos dados – Aspecto relacionado com as diferentes fontes disponíveis;
2. O uso de processamento em nuvem – Aspecto relacionado ao uso ilimitado de recursos computacionais e com processamento em larga escala, com a possibilidade de redução de custos (Economia de escala – É o aspecto econômico/financeiro);
3. Uso de tecnologias específicas, tais como processamento de rotinas em paralelo e ferramentas para otimização como Hadoop e MapReduce, HDFS, além de abordagens de *Machine Learning* e *Analytics*.

César Taurion definiu *big data* inicialmente como sendo a soma de 3‘Vs: Volume, Variedade e Velocidade. O volume significa a quantidade de informações que são armazenadas, já a variedade é caracterizada pela grande não padronização dos dados, ou seja, podendo ser de qualquer tipo e formato em um só lugar, como: vídeos, fotos, músicas e etc, e velocidade é o tempo que essa informação pode ser processada. Posteriormente o mesmo autor acrescentou os ‘Vs’ Veracidade e Valor, que são caracterizados respectivamente pela credibilidade da informação e pelo valor agregado a ela e ressaltou que as oportunidades que esses 5 ‘Vs’ trazem não podem ser desperdiçados. Já Dumbill (2012) informa que a abordagem de *big data* está apoiada em quatro valores, conhecidos como os 4 Vs do *big data*: Volume, Variedade, Velocidade e Veracidade.

As Figuras 3 e 4, apresentam os 3 conceitos interligados de *big data* sobre volume, variedade e velocidade de dados de acordo com a IBM:

Figura 3 – Caracterizando o *big data* de acordo com os 3 Vs: Volume, Velocidade e Veracidade.



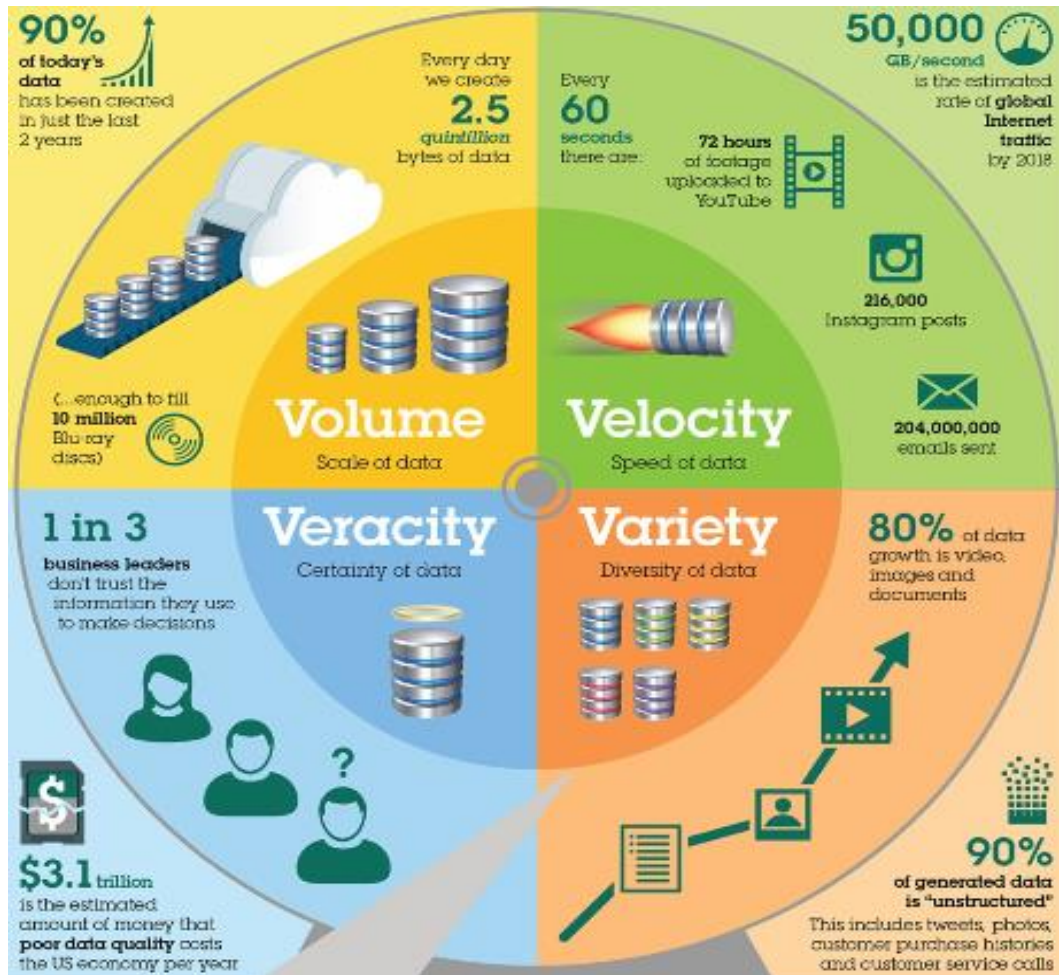
Fonte: IBM

Em relação a mudança dos dados, Betser e Belanger cita que:

Esta combinação da *internet* e o *WWW* em meados dos anos 1990, no entanto, sinalizaram uma mudança radical não só na quantidade de dados – o volume, mas também na taxa com que os dados chegavam — a velocidade; na diversidade de fontes de onde eles chegavam e nos tipos de dados e sua disponibilidade – a variedade. (Betser e Belanger, 2013, s.p.).

A velocidade significa tanto o quão rápido os dados estão sendo produzidos quanto o quão rápido os dados devem ser tratados para atender a demanda. Portanto, é de extrema importância que gestores e organizações tomem decisões rápidas se quiserem criar um diferencial competitivo.

Figura 4 – Os cientistas de dados da IBM dividem grandes dados em quatro dimensões: volume, variedade, velocidade e veracidade. Infográfico explicando exemplos de cada



Fonte: IBM

Para que seja possível entender os efeitos das dimensões do *big data* no processo de decisão das empresas, é necessário entender o processo decisório e então associá-los aos conceitos já apresentados. De acordo com Gomes (2007), decisão é o processo de escolha, direta ou indiretamente, de pelo menos uma das diferentes alternativas apresentadas, que atenderiam à resolução do problema existente.

Segundo Choo (2006), para tomar uma decisão racional, deve-se elencar todas as alternativas possíveis e identificar suas consequências. Para que isso seja possível, é preciso verificar os dados obtidos e extrair as informações fundamentais.

Para o autor, as empresas e organizações criam e utilizam as informações sob três aspectos:

- Interpretação da informação sobre o ambiente de modo a dar significado ao que está acontecendo à organização e ao que ela está fazendo;
- Criação de novos conhecimentos combinando a experiência de seus membros, de modo a aprender e inovar;
- Processo e análise da informação de modo a escolher empreender cursos de ação apropriados.

3.2 IMPACTO DO *BIG DATA*

Além do impacto no negócio, o *big data* pode afetar a empresa em vários aspectos, como governança e gerenciamento do planejamento, utilização, garantia e privacidade (ISACA, 2013):

1. **Governança:** A governança garante que as necessidades, condições e opções das partes interessadas sejam avaliadas para determinar que metas devem ser alcançadas pela empresa. Sem um processo de governança adequado, os projetos de *big data* podem desencadear uma série de problemas, incluindo dados enganosos e custos inesperados. Um desafio significativo no processo de governança de *big data* é categorizar, modelar e mapear os dados à medida que são capturados e armazenados
2. **Planejamento:** Envolve o processo de coleta e organização dos resultados obtidos pela análise dos dados.
3. **Utilização:** O uso de *big data* pode variar de uma empresa para outra, dependendo de seu nível de maturidade e cultura. Uma empresa menor pode demorar mais tempo para adotar o *big data* por não dispor de uma infraestrutura necessária para suportar os novos processos envolvidos. Por outro lado, empresas de porte como a IBM, HP e Amazon.com, fizeram os investimentos

necessários e adotaram o *big data*, utilizando as informações obtidas por essa tecnologia para tomar decisões de negócio e mudar a direção de suas vendas.

4. **Garantia:** A maior preocupação de garantia para uma empresa é a qualidade dos dados. Uma informação é eficaz se atende às necessidades do consumidor de informações. Quanto melhor a qualidade do dado, melhores serão as decisões com base no dado, criando valores para a empresa. A gestão de *big data* deve garantir a qualidade do dado ao longo do seu ciclo de vida.
5. **Privacidade:** Leis protegem a privacidade dos indivíduos e todas as informações coletadas sobre eles, mesmo que as pessoas compartilhem informações confidenciais inadequadamente. Mesmo com a autenticidade da informação coletada das mídias sociais, ela precisa ser protegida de usuários com má-fé e de governos muito controladores.

As empresas estão cada vez mais dependentes dos dados para a melhoria na tomada de decisões. Dados imprecisos, incompletos ou manipulados de maneira fraudulenta representam um risco crescente para os tomadores de decisões estratégicas (Feijó, 2013). Segurança e privacidade desempenham um papel cada vez mais importante em *big data*, e todas as partes interessadas devem estar cientes das implicações de armazenamento e análises cruzadas de grandes quantidades de dados confidenciais.

4 BENCHMARK: AS ABORDAGENS CONCEITUAIS

A essência do *benchmarking* baseia-se na ideia de que nenhuma organização é melhor em todas as áreas, o que implica reconhecer que existe alguém no mercado que consegue fazer melhor, assim o objetivo é, portanto, estimular e facilitar as mudanças organizacionais e a melhoria de desempenho através da aprendizagem com os outros (Keegan, *et al.*, 2006).

Para Spendolini (1993, p. 10), o *benchmarking* pode ser definido como “um processo contínuo e sistemático para avaliar produtos, serviços e processo de trabalho de organizações que são reconhecidas como representantes das melhores práticas, com a finalidade de melhoria organizacional”. Essa definição parece bem completa, quando mostra ou mesmo apresenta o *benchmark* como um processo contínuo e sistemático, enfatizando assim o caráter metodológico que tal técnica apresenta. Conforme Araújo (2000, p. 185), a técnica do *benchmarking* é:

Centrada na premissa de que é imperativo explorar, compreender, analisar e utilizar as soluções de uma empresa, concorrente ou não, diante de determinado problema. O *benchmarking* é uma excelente tecnologia de gestão organizacional e oferece aos que acreditarem corretamente em seu potencial, alternativas que aperfeiçoam processos organizacionais, produtos e serviços. (Araújo, 2000, p. 185)

Camp (1998, p. 10) apresenta uma definição a que chama de operacional quando diz que “*Benchmarking* é a busca das melhores práticas na indústria que conduzem ao desempenho superior”. Esta definição tem seu caráter operacional pois está bem centrada nas operações realizadas por uma empresa, bem como subentende que a empresa tenha um perfeito conhecimento a respeito de suas operações, ao utilizar da técnica do *benchmarking*. Assim é caracterizada como uma ferramenta de gestão organizacional que visa principalmente a praticidade e facilidade na busca de melhorias, sendo essas importantes vantagens para uma organização.

A concorrência está no âmago do sucesso ou do fracasso das empresas, determinando a adequação das atividades que podem contribuir para seu desempenho, como inovações, uma cultura coesa ou uma boa implementação. A estratégia competitiva é a busca de uma competição favorável para uma empresa no mercado, arena fundamental onde ocorre a concorrência. A estratégia competitiva visa estabelecer uma posição lucrativa e sustentável contra as forças que determinam a concorrência (Porter, 1990).

O aparecimento de tensões nos negócios e economia oferece oportunidades para mudanças e pesquisas, à medida que os participantes se engajam na interpretação coletiva e na negociação, com o objetivo de chegar a novas maneiras de conhecer e fazer. Para Chiavenato (2004), o *benchmarking* significa um padrão de excelência que deve ser identificado, conhecido, copiado e ultrapassado, que pode ser interno ou externo.

4.1 TIPOS DE *BENCHMARKING*

Benchmarking é um processo contínuo e sistemático para avaliar produtos e processos de trabalho de organizações que são reconhecidas como representantes das melhores práticas, com a finalidade de melhoria organizacional (Spendolini, 1992).

Alguns autores como Camp (1998), Dias (2008) e Amaral Júnior (1993) concordam que existe quatro tipos de *benchmarking*: O interno, o competitivo, o funcional e o genérico. Na verdade, o ponto de divergência entre estes autores está entre o *benchmarking* funcional e o genérico, ou seja, uns colocam tais tipos de *benchmarking* de forma separada, enquanto outros os colocam como único. Para efeito deste trabalho, será considerado apenas os três tipos citados por Spendolini (1993), Araújo Junior (2001) e Araújo (2000), visto que, de fato há grande semelhança entre o *benchmarking* funcional e o genérico. As definições de acordo com os autores são as seguintes:

1. **Benchmarking interno:** Realizado dentro da própria organização, entre suas várias unidades ou departamentos. É mais um processo de conhecimento das práticas internas;
2. **Benchmarking competitivo:** Realizado pela comparação com empresas que são competidoras diretas da organização. É muito útil para posicionar o desempenho frente à concorrência;
3. **Benchmarking funcional/ genérico:** Realizado através da identificação das melhores práticas em qualquer tipo de organização que estabeleceu uma reputação de excelência na área específica sujeita ao *benchmarking*. Amaral Júnior (1993, p. 9) caracteriza como “Descobrir como se tornar o melhor em um determinado processo. Ele por si só não melhora o desempenho. É preciso que as informações coletadas sejam analisadas e transformadas em ações estratégicas para se obter um vantagem competitiva.”

O que se espera em todos os tipos citados, é identificar dentro da organização as melhores práticas com o objetivo de disseminá-las por todas as áreas da empresa. É tanto uma forma de busca de melhorias, quanto uma oportunidade de autoconhecimento para a empresa. Assim, de um modo genérico todos os processos de *benchmarking* se resumem em cinco etapas básicas (Shetty, 1993):

1. Determinar do que fazer *benchmarking*;
2. Identificar parceiros;
3. Coletar e analisar dados;
4. Estabelecer metas;
5. Implementar ações e monitorar progresso.

O sucesso de um projeto de *benchmarking* depende do total envolvimento da alta gerência, dando o suporte e fornecendo os recursos necessários para sua implementação e seu desenvolvimento de uma mudança cultural na organização,

que deve reconhecer que pode aprender com terceiros e da disponibilidade de informações sobre o objeto a ser estudado, fator este que pode ser dificultado quando se realiza o *benchmarking* competitivo (Elmuti&Kathawala 1997).

Estes tipos de *benchmarkings*, quer seja utilizados individuais, quer seja utilizado de forma conjunta, devem na verdade buscar soluções para a organização. Eles precisam cumprir o objetivo do *benchmarking*, que como foi informado em parágrafos anteriores, é caracterizado para Amaral Júnior (1993, p. 9) como: “Descobrir como se tornar o melhor em um determinado processo. Ele por si só não melhora o desempenho. É preciso que as informações coletadas sejam analisadas e transformadas em ações estratégicas para se obter um vantagem”.

4.2 NECESSIDADE E TECNICAS DO BENCHMARK

Tyson (1998) define a Inteligência Competitiva, no contexto empresarial, como um processo sistemático que transformam dados e informações dispersas em conhecimento estratégico. É o conhecimento da posição competitiva atual e os planos futuros dos concorrentes. É o conhecimento das forças que dirigem os mercados. É o conhecimento sobre os produtos e as tecnologias. São também as informações externas ao mercado, tais como as influências econômicas, regulatórias, políticas e demográficas que tenham um impacto no mercado. Dias (2008) aponta três razões fundamentais para se praticar o *benchmarking*:

1. Colocando-o como uma forma prática para se otimizar a busca por melhorias;
2. Uma boa alternativa para se eliminar o antigo processo de aprendizagem por meio de tentativas e erros;
3. E uma forma de se utilizar processos que outros já utilizaram logrando êxito.

Camp (1998, p. 10) apresenta uma definição a que chama de operacional quando diz que “*Benchmarking* é a busca das melhores práticas na indústria que conduzem ao desempenho superior”. Esta definição subentende que a empresa tenha um perfeito conhecimento a respeito de suas operações, ao utilizar da técnica do *benchmarking*.

Davenport & Prusak (1998) salientam que: “A única vantagem sustentável que uma empresa possui, é aquilo que ela coletivamente sabe, a eficiência com que ela usa o que sabe e a prontidão com que ela adquire e usa novos conhecimentos”, ou seja, o modo de obtenção de novos conhecimentos é um fato bastante valioso. Le Coadic (2005) afirma que o conhecimento possui duas características importantes: Sua explosão quantitativa e a implosão do tempo para sua comunicação, e a define como um conhecimento inscrito (registrado) em forma escrita, seja ela em meio impresso ou digital, oral ou audiovisual, em um suporte, possuindo um elemento de sentido, e tendo por objetivo, a transmissão de conhecimento, que independentemente da forma que é repassada, é de extremo valor. Vasconcelos e Cyrino (2000) descrevem quatro principais correntes teóricas que tratam o conhecimento como estratégia empresarial, são as teorias: Das “Organização Industrial”, dos “Recursos”, dos “Processos de Mercado”, e das “Capacidades Dinâmicas”.

O *benchmarking* são processos de pesquisas e melhorias através das informações obtidas, para Spendolini (1993, p. 10), o *benchmarking* pode ser definido como “Um processo contínuo e sistemático para avaliar produtos, serviços e processo de trabalho de organizações que são reconhecidas como representantes das melhores práticas, com a finalidade de melhoria organizacional”. A definição de Spendolini é bem clara, e é entendido que este processo requer processos sistemáticos e contínuos para geração e mantimento da qualidade total na organização.

De acordo com Dolor Barbosa Xidieh (2000, p. 27- 43), o processo de *benchmarking* pode ser executado através de cinco fases principais:

1º Fase: Planejamento: É o passo mais crítico da condução do *benchmarking*, por isto, devem ser dedicados tempo e recursos na discussão dessa etapa. Ciente disso é necessário:

- Definir objeto e equipe de estudo e quais departamentos ou problemas serão submetidos a estudo e quem serão os responsáveis pela coleta dos dados.
- Entender o objeto de estudo. Deve-se definir os critérios sob os quais as organizações podem ser comparadas e o objetivo - Se são mudanças radicais ou incrementais, e que tipo de *benchmarking* utilizar.
- Selecionar organizações parceiras (Caso trate de um *benchmarking* funcional / genérico).

2ª Fase: Coleta de Dados

- Definir métodos de coleta: Se funcionará através de visitas a organizações parceiras (*benchmarking* funcional), ligações, demonstrações financeiras e etc.;
- A preparação de um questionário completo, contendo as perguntas essenciais sobre o processo do parceiro representa um grande passo;
- Coletar os dados.

3ª Fase: Análise

- Identificar as diferenças existentes em termos de desempenho do objeto do estudo, entre a entidade e o parceiro de *benchmarking* e, principalmente, compreender as razões pelas quais esses diferenciais existem;
- Projetar o desempenho futuro.

4ª Fase: Adaptação

- Após serem identificadas as áreas que precisam ser adequadas a realidade da empresa estudada, deve-se buscar alternativas para efetivamente incorporar esses diferenciais desenvolvendo um plano para implantar melhorias que gerem desempenho igual ou melhor do que o do parceiro;
- Definir um plano detalhado de mudanças e estabelecer metas concretas a serem atingidas.

5ª fase: Melhorias

- Executar planos de melhorias. Definir o plano de atuação de acordo com a estratégia da empresa. Recomenda-se observar alguns requisitos importantes como: recrutar a equipe de atuação, executar os planos, monitorar os resultados e por fim reavaliar as metas a partir de novos referenciais.

A adoção dos critérios de excelência serve como importante instrumento de auto avaliação das organizações e contribui fortemente para o vigoroso salto de especialmente por aquelas que utilizam o *benchmark* como um dos métodos de referencial para o seu planejamento estratégico para a implementação de melhorias, César Taurion afirmou em 2012 que o volume de dados do planeta dobra a cada 18 meses e que a geração diária chega a *petabytes* de dados. Ele ressaltou também o ‘V’ variedade ao escrever que estes dados são gerados a partir de sistemas estruturados e não estruturados tais como: *E-mails*, mídias sociais (*Facebook*, *Twitter*, *YouTube* e outros), documentos eletrônicos apresentações estilo *PowerPoint*, mensagens instantâneas, sensores, etiquetas RFID e câmeras de vídeo.

Além das inúmeras vantagens apresentadas pela aplicação do *benchmarking* nas organizações, são verificadas algumas dificuldades na utilização desta ferramenta. Segundo Spendolini (1999), existem desvantagens que impedem a implantação do *benchmarking* de forma eficiente e eficaz, entre elas destacam-se:

1. Ausência de recursos financeiros e de tempo, para a recolha de informações;
2. Os colaboradores da organização não se envolvem no processo;
3. O nível de confiança relativamente à empresa que fornece os dados é baixo;
4. O não acompanhamento de todo o processo de *benchmarking*, deixa de garantir bons resultados.

Para apoiar as organizações na estruturação do seu sistema de gestão da qualidade o envolvimento do gestor da informação na busca de informações e soluções de qualidade constitui um aspecto fundamental nos processos de melhoria da qualidade. A ênfase na melhoria contínua através do *benchmarking* reflete a tentativa de manter uma vantagem de qualidade ao longo do tempo, sempre buscando novos meios para melhoraria do desempenho atual. A filosofia básica da melhoria contínua é que cada pessoa/empresa nunca deve estar satisfeita com o que faz, mas estar sempre na busca constante do aperfeiçoamento

5 O *BIG DATA* COMO UMA FORMA DE *BENCHMARK*

A técnica do *benchmarking* surge no ramo dos negócios e das empresas e a partir da busca pela qualidade total, visa gerar lucros, status e atualização das empresas que a utiliza. Ou seja, ele já nasce revestido com esse objetivo empresarial de ser mais uma ferramenta de competitividade no ramo dos negócios. Além das mudanças com o usuário ou o manuseador da informação, também ocorre aos suportes e aos tomadores de decisões. Esse pessoal, que vem sendo chamado cientistas de dados, precisam ter raciocínio lógico apurado, conhecimento profundo de estatística e *software* dessa área, modelagem, domínio de sistemas computacionais e conhecimento do negócio e do mercado em que atuará. Além disso, será necessária a capacidade de comunicar seus *insights* a um pessoal com uma formação não técnica (Davenport; Barth; Bean,2012). Assim é bastante visível e viável a utilização do *big data* como uma forma de melhoria empresarial e econômico de um negócio.

Marcos Vieira (2014) descreve o processamento e análise do *big data* da seguinte forma:

Uma vez que os dados já estão todos armazenados e disponíveis de forma adequada, passamos para a etapa de tratamento e enriquecimento. A matéria de tratamento algorítmico abrange: Criação de índices, aplicação estatística indutiva (Análise de comportamento e clusters), modelagem preditiva (Estimativa resultados futuros) e de algoritmos relativos a natureza de dados (Processamento de imagens, análise semântica de textos, etc.). Esses processos têm grande dependência do resultado final que se deseja. Cada empresa possui uma necessidade e/ou uma capacidade técnica de aplicar alguns ou vários dos métodos citados. As habilidades necessárias nessa etapa se dividem sobretudo entre conhecimentos de economia (Econometria), computação, matemática e estatística. Como se pode imaginar é muito difícil que uma empresa possua uma equipe que domine todas essas técnicas, tanto mais difícil é encontrar um profissional que, sozinho, consiga resolver todas elas. O padrão no mercado é terceirizar projetos com empresas especializadas que já acumulam expertise nas soluções desejadas. (2014, s.p.)

A técnica vem se aperfeiçoando com o passar do tempo e se tornando cada vez mais conhecida através das experiências de sucesso, porém ainda falta

bastante para que as empresas possam investir com muita dedicação nesse formato de obtenção de informação para crescimento empresarial. Com isso, muitos foram ou são os exemplos de instituições que se utilizaram da prática do *benchmarking* em busca também de uma qualidade total, porém no meio do processo desistem por variados motivos.

Segundo Bovo e Balancieri (2001), abordando a questão da transformação "dado — informação — conhecimento", para se chegar ao conhecimento, um longo e árduo caminho deve ser percorrido. Os dados são elementos que mantêm a sua forma bruta (Texto, imagens, sons, vídeos, etc.), ou seja, eles sozinhos não ajudam a compreender determinada situação, enquanto a informação é o dado cuja forma e conteúdos são apropriados para um uso específico. Assim, a utilização de ferramentas para análise de informações estratégicas é empregada para auxiliar na transformação desses dados em informação e conhecimento, ocorrendo no final de todo o processo a obtenção de dados valiosos para o negócio e conseqüentemente a tomada de decisão. Conforme Davenport e Prusak (1998):

Conhecimento não é dado nem informação, embora estejam relacionados com ambos e a diferença entre esses termos seja normalmente uma questão de grau. Por mais primário que possa soar, é importante frisar que dado, informação e conhecimento não são sinônimos. O sucesso ou fracasso organizacional, muitas vezes pode depender de saber qual deles precisamos, com qual deles contamos e o que podemos ou não fazer com cada um deles.

Para gestor da informação o *benchmarking* permite um aperfeiçoamento do seus produtos e serviços na organização, identificar em outras organizações, pesquisar processos ou atividades rotineiras que são realizadas com sucesso para uma futura adaptação, aumentando sua vantagem competitiva e contribuir de forma significativa para a atividade de gestão com novas informações.

O conceito de *benchmarking* é relacionado ao cenário competitivo, este processo é usado para aprender como nos adaptar a esse padrão de excelência competitiva imposta, nos ajudando a entender como transformar nossos próprios processos internos ou externos. Para Manyika, o grande benefício do *big data* é

permitir a análise de qualquer tipo de informação digital em tempo real, gerando base para a tomada de decisões e conseqüentemente melhorando a competitividade, produtividade, inovação e benefícios ao consumidor, que no mercado econômico que estamos, é um grande diferencial.

No contexto organizacional, a estratégia corresponde à capacidade de se trabalhar contínua e sistematicamente o ajustamento da organização às condições ambientais que se encontram em constante mudança, tendo sempre em mente a visão de futuro. É a procura e implementação das melhores práticas a partir do eficiente gerenciamento dos dados recuperados dos concorrentes e com a eficiência na utilização estrategicamente. Jefferson Stela afirma que:

Não há uma indústria ou área específica que não se possa aplicar este novo conceito. Basicamente, em todo lugar onde há informação, trabalhada de uma forma criativa e estratégica, e que possa trazer um valor agregado como objetivo final, terá crédito na utilização de *big data*. Temos desde casos complexos como a utilização na análise de sedimentos coletados no fundo de oceanos identificados possíveis melhores pontos de perfuração de poços petrolíferos, como na utilização de lojas que querem entender se um consumidor que comprou um determinado produto pode ser o futuro consumidor de um outro produto ou serviço específico.

É notado a transformação dos dados recuperados através da mudança, no propósito de melhorar continuamente os produtos e serviços. Através dessas mudanças a empresa poderá entrar no mercado de forma mais eficiente e eficaz. A informação bem gerenciada pode mudar o cenário econômico organizacional refletindo positivamente no resultado empresarial.

Conforme Araújo (2000, p. 185) é acrescentado que, a técnica do *benchmarking* é: Centrada na premissa de que é imperativo explorar, compreender, analisar e utilizar as soluções de uma empresa, concorrente ou não, diante de determinado problema, o *benchmarking* é uma excelente tecnologia de gestão organizacional e oferece aos que acreditarem corretamente em seu potencial, alternativas que aperfeiçoam processos organizacionais, produtos e serviços.

Camp (1998, p. 10) apresenta uma definição a que chama de operacional quando diz que “*Benchmarking* é a busca das melhores práticas na indústria que conduzem ao desempenho superior”. Spendolini (1993) e Camp (1998) afirmam que a técnica do *benchmarking*, por apresentar certa flexibilidade, é perfeitamente adaptável à qualquer tipo de organização ou empreendimentos empresariais. Porém o percebemos é que a pratica é um pouco difícil a implementação como qualquer outro dentro de uma organização competitiva.”

5.1 VANTAGENS DA UTILIZAÇÃO DO *BIG DATA* COMO *BENCHMARK*

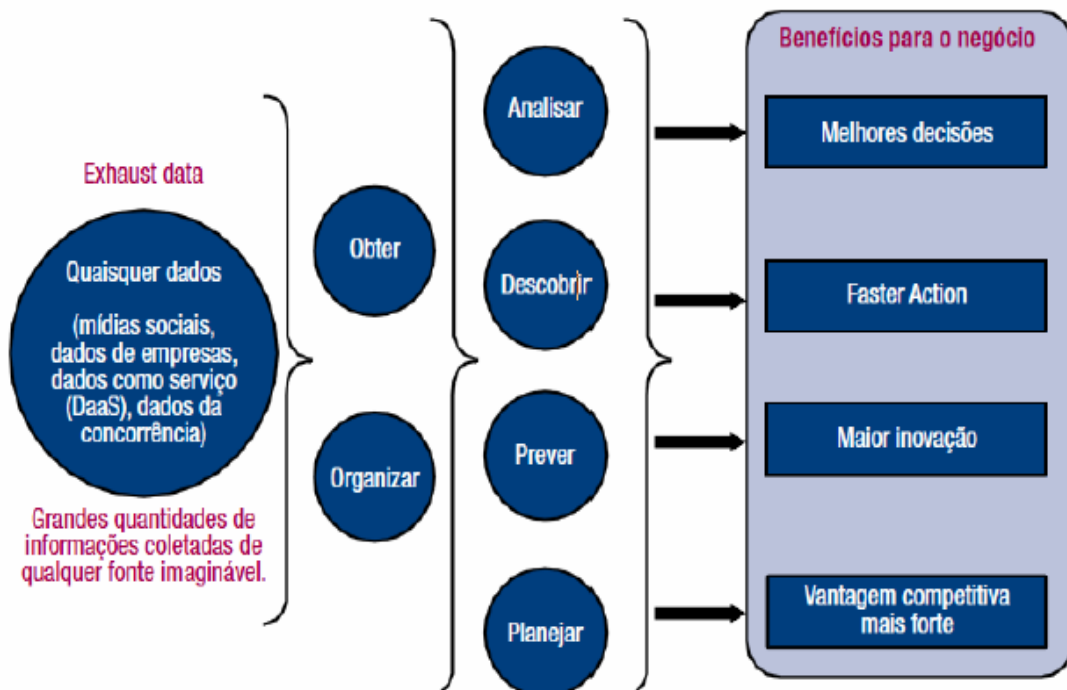
As oportunidades de negócios utilizando o *big data* são significativas, assim como os desafios. Uma pesquisa feita por Erik Brynjolfsson, economista da *Sloan School of Management* do Instituto de Tecnologia de *Massachusetts* (EUA), mostra que as empresas que utilizam a "Tomada de decisão com base em dados" obtêm um aumento de cinco a seis por cento em produtividade. O uso adequado de *big data* vai além de coletar e analisar grandes quantidades de dados, também exige a compreensão de como e quando usar os dados ao tomar decisões cruciais (ISACA, 2013).

Realizar uma lista de todos os benefícios que o processo de *benchmarking* proporciona é uma tarefa bastante difícil, pois esta é uma ferramenta de melhorias contínuas, assim visa capacitar os profissionais a um aprendizado rápido, onde a busca de melhores ideias deve ser uma constante dentro das organizações. (Luis César G. de Araujo, 2009, p. 196-203). O uso do *big data* entrega uma vantagem muito rápida e segura no benchmarking empresarial, como diz Spendolini (1993) e Camp (1998), a pratica de melhoria é bastante difícil, porém quando se torna algo ‘rotineiro’ tudo pode se tornar melhor, conseguindo assim o objetivo empresarial, seja qual for em relação as informações obtidas.

O *big data* permite que se gerenciem muito mais dados que antes, e os dados não precisam ser alocados em fileiras ou nas clássicas tabelas (Schonberger-Mayer e Cukier, 2013). Assim o efeito da utilização do *benchmarking* no desempenho das organizações tem sido bastante estudado. Em virtude das rápidas mudanças no mercado, definir metas olhando para dentro da empresa geralmente não responde às expectativas dos clientes. As expectativas dos clientes são guiadas pelos padrões estabelecidos pelos melhores fornecedores do setor e também pelas boas experiências de fornecedores dos outros segmentos de negócios. O maior benefício do *benchmarking* é, portanto, o de ajudar a alcançar níveis de desempenho que atendam plenamente às crescentes expectativas dos consumidores.

A Figura 5 apresenta o processo para acesso de "*insights*" comerciais específicos da empresa a partir de *big data* (ISACA, 2013).

Figura 5 – Analisando os *insights* comerciais específicos



Fonte: COBIT 5

Em termos de tecnologia, Cukier (2010) indica que os negócios em gestão da informação estão contribuindo para facilitar o processo de construção de sentido nos grandes volumes de dados, ou seja, *big data*. Por outro lado, as contribuições de *big data* são possíveis de serem realizadas desde que a matéria informacional gerada pela exploração dos dados, seja transformada em inteligência, por meio dos processos de análise. A complexidade do ambiente das corporações foi incrementado com a globalização, tornou-se ágil e orientado a rede e as organizações puderam passar a pesquisar com inteligência suas bases de clientes. Estas organizações tiveram que organizar suas bases, analisar padrões de e-mail, gravações telefônicas, mensagens instantâneas e outras evidências para determinar o modelo organizacional emergente. Assim, a própria forma da empresa começou a se modificar, habilitada pela tecnologia e estimulada pela necessidade de complexidade.

Muitos dos ganhos com *big data*, acontecem não por causa de *chips* mais rápidos ou melhores algoritmos, mas sim pela existência de mais dados (Schonberger-Mayer e Cukier, 2013). Mineli, Chambers e Dhiraj (2013) exploram melhor a questão e identificam as razões para esta nova era:

1. **Tempestade perfeita da computação:** *Big data* é o resultado natural das quatro maiores tendências globais, a lei de Moore, computação móvel (com *smartphones* e *tables*), redes sociais (*Facebook*, *Four Square* e outros), e a computação em nuvem (*cloud computing*- possibilidade de alugar ou arrendar *hardware* ou *software* para utilização);
2. **Tempestade perfeita de dados:** Volumes de dados transacionais foram por décadas utilizados pela maioria das grandes empresas, mas o que se apresenta atualmente é mais volume, velocidade e variedade - os Vs - de dados que chegaram de forma inédita. Esta tempestade perfeita dos Vs, torna extremamente complexo e pesado o gerenciamento de dados atual e análise de tecnologias e práticas;

3. **Tempestade perfeita de convergência:** Gerenciamento de dados tradicionais, análise de *software*, tecnologias de *hardware*, tecnologia de código aberto e *commodities* de *hardware* estão se fundindo para criar novas alternativas para TI e executivos de negócios para enfrentar *big data*.

Tabela 2 – Evolução da utilização dos dados e informação

Ano ou Período	Fato
8000 a.C.	Comerciantes sumérios usavam tábuas e fichas de barro para denotar os bens comercializados.
1086	Livro Domesday, um dos mais venerados tesouros britânicos, foi uma contagem abrangente - e sem precedentes - dos ingleses, suas terras e propriedades.
1439	Primeira revolução da informação: Impressora de Gutenberg.
1453 e 1503	De acordo com a historiadora Elizabeth Eisenstein, oito milhões de livros foram impressos. Volume de informações dobra na Europa, em 50 anos.
1662	O comerciante britânico chamado John Graunt queria saber a população de Londres na época da peste negra. Em vez de contar as pessoas, ele inventou uma abordagem - que hoje chamamos de "estatística"- que lhe permitiu estimar a população.
1880	O Census Bureau dos Estados Unidos contratou o inventor Herman Hollerith, que desenvolveu um sistema de cartões perfurados e máquinas de tabulação para o censo de 1890. Ele conseguiu diminuir o tempo de tabulação do censo de oito anos para um.
1920	Descobertas da mecânica quântica abalaram para sempre o sonho das medições abrangentes e perfeitas.
1934	Jerzy Neyman, estatístico polonês, demonstrou que a abordagem da amostragem levava a erros e que para isto deveria ser considerada a aleatoriedade na escolha da amostra. A amostragem tornou-se a solução para o problema da profusão de dados.
1950-1960	Implementação de um sistema de folha de pagamento para General Electric Corporation, por Joe Glickauf e Arthur Andersen em um computador eletrônico digital. Início das ideias de "Revolução da informação" e "Era Digital" ou "Era da informação".
1980-1990	Nesta época, Leonard Kleinrock, Lawrence Roberts, Robert Kahn, e Vint Cerf inventaram a internet. 40% de toda a potência computacional do mundo existia na forma de calculadoras de bolso. "Era da rede".
2000	Dados: 75% dados analógicos (papel, filme, vinil, fitas magnéticas, livros, fotografia); 25% dados digitais.
2007	Dados: 7% dados analógicos; 93% dados digitais.
2010	Ciências como a astronomia e a genômica, vivenciaram uma explosão informacional, e cunharam o termo <i>big data</i> , que representa uma transição em termos de armazenamento e análise. "Era <i>big data</i> ".
2013	Dados: 2% dados analógicos; 98% dados digitais - 1200 <i>exabytes</i> . Volume de informações dobra a cada 3 anos.

Fonte: Definição de Schonberger-Mayer, Cikier k.(2013) e Milneli M., Chambers M.,Dhiraj A. (2013)

Por outro lado, e acordo com Cukier (2010), os efeitos do fenômeno *big data* vem sendo percebidos em todos os lugares, nas empresas, na ciência e no governo. Seu conceito está migrando para todos os campos do conhecimento humano pois em essência seu avanço é uma continuação da antiga busca da humanidade em medir, registrar e analisar (benchmarking) o mundo (Schonberger-Mayer e Cukier, 2013).

6 CONSIDERAÇÕES FINAIS

Taurion (2013) indica que antes, no ano 2000, apenas 25% dos dados disponíveis no planeta estavam em formato digital e, em 2013, esses dados somavam 98%. Portanto, o mercado de TI e gerenciamento tem nas mãos o desafio de analisar e entender grandes quantidades de dados que está ficando cada vez maior por causa do armazenamento mais barato e evolução dos dados digitais e dispositivos de coleta de informações.

A utilização do conceito *big data* permite que as empresas analisem grandes volumes de dados de variados tipos gerando informações mais precisas dos seus clientes. Estas informações são fundamentais para antecipar necessidades, negócios, oferecer serviços melhores e mais rápidos. O *big data* pode contribuir grandemente para o futuro das corporações, que precisam decidir seus processos em informações de qualidade num menor tempo possível, garantindo competitividade no âmbito que estão inseridas. Pelo exposto no decorrer deste trabalho, compreende-se que estamos passando por uma grande transformação social e tecnológica, caracterizada pela quantidade de dados que surgem exponencialmente a cada segundo. Cabem às empresas, organizações e até governos aproveitarem a grande oportunidade pra entender o que a sociedade quer (realização de *benchmarking*) e saber que a velocidade significa tanto o quão rápido os dados estão sendo produzidos quanto o quão rápido os dados devem ser tratados para atender a demanda destas informações. Portanto, é de extrema importância que gestores e organizações tomem decisões rápidas se quiserem criar e estar em um cenário diferencial e competitivo.

Segundo a SAS (2013) apesar da promoção envolvendo o *big data*, grande parte das organizações ainda não se organizaram para implantar uma estratégia de *big data*. Líder de mercado em soluções e serviços de inteligência analítica, o SAS e a SourceMedia entrevistaram 339 gestores de TI sobre o uso de tecnologias de gerenciamento de dados em suas organizações. Os resultados mostram que poucas empresas tiram vantagens das informações sobre seus produtos, clientes ou de outras fontes de dados. A pesquisa demonstra que quem não conhece de fato o seu público/cliente, as chances de crescer no mundo dos negócios se torna bastante

complexa. Como se não bastasse, o dinamismo do mundo moderno também exige dos gestores altíssimo conhecimento e de realização constante de *benchmarking* e a obtenção de um elevado nível de informações valorosas, tudo com antecedência, antes da eclosão do fato. Ou seja, é preciso ter visão 360º sobre todos os *stakeholders*, prevendo comportamentos, inovações dos rivais e aumento de custos para se manter no atual cenário e prosperar continuamente. Assim emerge a necessidade de uma solução que permita a recolha e integração de dados, provenientes de várias fontes mas para além disso que possibilitem a análise de dados em qualquer lugar e momento pois o rápido acesso aos dados é crucial para que os processos de tomada de decisão sejam mais rápidos, práticos e eficientes possíveis (Russom, 2013).

Por fim da pesquisa, as empresas devem entender o conceito de *big data* e perceber que a utilização do *big data* como processo de *benchmarking* nos negócios é de extrema importância para o desenvolvimento e crescimento, pois o que se verifica é que as empresas reconhecem o *big data* e acreditam que o mesmo pode ajudar no seu negócio, porém não percebem de que forma e como o *big data* pode ser utilizado para gerar a inovação.

7 REFERENCIAS BIBLIOGRAFICAS

ALECRIM, Emerson. **O que é Big Data?** Atualizado em 13/01/2015. Disponível em: < <https://www.infowester.com/big-data.php> > Acesso em: 15/04/2017

ALMEIDA, Patrícia de Fatima. **B I G D A T A : E A SUA APLICABILIDADE EM INSTITUIÇÕES FINANCEIRAS.** Disponível em:< <https://pt.scribd.com/document/328445134/Tcc-Big-Data-e-a-Sua-Aplicabilidade-Em-Instituicoes-Financeira>> Acesso em 22/03/2017.

AMARAL JÚNIOR, Geraldo. **Benchmarking: a arte de vencer com o acerto dos outros e vencer a competição.** Enfoque, São Paulo, v.3, ano XX, p. 8-11, jul. / set., 1993

ARAÚJO JÚNIOR, Rogério Henrique de. **Benchmarking.** In: TARAPANOFF, Kira (Org.). **Inteligência organizacional e competitiva.** Brasília: EDUNB, 2001. p. 241-263.

ARAÚJO JÚNIOR, Rogério Henrique de. **Benchmarking.** In: TARAPANOFF, Kira (Org.). **Inteligência organizacional e competitiva.** Brasília: EDUNB, 2001. p. 241-263.

ARAÚJO, Luís César G de. **Organização, Sistemas e Métodos e as Tecnologias de Gestão Organizacional.** Volume 2. São Paulo: Ed. Atlas SA, 2007.

CHIAVENATO, Idalberto. **Introdução à Teoria Geral da Administração.** Rio de Janeiro: Elsevier, 2004.

ARAUJO, Luis César G. De. **Benchmarking: ser o melhor entre os melhores. Organização, sistemas e métodos e as tecnologias de gestão organizacional.** São Paulo: Atlas, 2001.

ARAUJO, Luis César G. de. **Organização, sistemas e métodos e as tecnologias de gestão organizacional.** 4. ed., rev. e atual. São Paulo: Atlas, 2009

ARRIGONI, Rodrigo. **Uma entrevista didática sobre o *Big Data***. 11/02/2014. Disponível em:< <https://www.ecommercebrasil.com.br/artigos/entendendo-big-data/>> Acesso em: 21/02/2017

ARRIGONI, Rodrigo. **Uma revista didática sobre Big Data**. Postado em 25/07/2013. Disponível em: < <http://adnews.com.br/tecnologia/uma-entrevista-didatica-sobre-big-data.html> > Acesso em: 18/04/2017

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 10520: informação e documentação: Citações em documentos: apresentação. Rio de Janeiro: ABNT, 2002

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBR 6023: informação e documentação: referências: elaboração. Rio de Janeiro: ABNT, 2002.

BERARDI, Franco. **A fábrica da infelicidade: trabalho cognitivo e crise da *new economy***. Rio de Janeiro: DP&A, 2005

Betser, J.; Belanger, D. (2013) **Architecting the enterprise via big data analytics**. In: Liebowitz, Jay. **Big data and business analytics**. Boca Raton: Taylor & Francis Group, LLC, 2013. Disponível em <<http://goo.gl/D0C0h9>> Acesso em 03/04/2017

BRUM, Analisa de Medeiros. **Um olhar sobre o Marketing interno**. 3.ed. Porto Alegre: L&PM, 2000.

CAMP, R.C, **Benchmarking dos processos de negócios; descobrindo e implementando as melhores práticas**. 1ª ed., Rio de Janeiro, RJ, Quality Mark Ed., 1996.

CAMP, Roberto C. **Benchmarking: o caminho da qualidade total**. 3. ed. São Paulo: Pioneira, 1998.

CASTELLS, Manuel. **A sociedade em rede**. São Paulo: Paz e Terra, 1999. v. 1.

Choo, C. W. (2006) **A Organização do Conhecimento**: com as organizações usam a informação para criar significado, construir conhecimento e tomar decisões. 2 ed. São Paulo. Senac São Paulo, 2006.

CHOO, Chun Wei. **A Organização do Conhecimento – como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões**. São Paulo: Editora SENAC São Paulo, 2003. 421 p.

DAVENPORT, T. H; BARTH, P; BEAN, R. **How Big Data is different**. MIT Sloan Management Review. Vol 54, nº 1, 2012.

DIAS, Flávio. **Benchmarking**. Disponível em: <<http://www.coladaweb.com/admmaterial/benchmarking.htm>>. Acesso em: 03 nov. 2008.

DRUCKER, Peter Ferdinand. **O melhor de Peter Drucker: obra completa**. São Paulo: Nobel, 2002.

DUMBILL, E. **What is Big Data?** In: O'Reilly Media Inc. **Big Data Now: current perspectives**. O'Reilly Media: Califórnia. 2012. Disponível em: <<http://www.oreilly.com/data/free/files/big-data-now-2012.pdf>>. Acesso em: 09/03/2017.

ELMUTI, D., KATHAWALA, Y., **An Overview of Benchmarking Process: a Tool for Continuous Improvement and Competitive Advantage, Benchmarking for Quality Management & Technology**, vol.4, no.4, 1997, pp. 229-245.

FEIJÓ, Bruno Vieira. **A Revolução dos Dados**. **Revista Exame PME** – Pequenas e

FERREIRA, Ademir A. FONSECA, Ana Carla Reis e PEREIRA, Maria Isabel. **Gestão Empresarial**. São Paulo - Ed. Pioneira, 1997.

Franks, B.; Davenport, T. (2012) **Taming the big data tidal wave**. John Wiley & Sons. 2012

GIL, A. **Métodos e técnicas de pesquisa social**. In: Métodos e técnicas de pesquisa social. Atlas, 2010.

HERRING, Jan P. **Tópicos fundamentais de inteligência: processo para identificação e definição de necessidades de inteligência**. In PRESCOTT, John E.; MILLER, Stephen H. *Inteligência Competitiva na Prática*. Editora Campus, São Paulo, 2002. p.274-291. Inc. Hoboken, Nova Jersey, 2012.

ISACA, **Big Data – Impactos e Benefícios**. 2013. Disponível em: <http://www.isaca.org/Knowledge-Center/Research/Documents/BigData_whp_Por_0413.pdf> Acesso em: 20 Maio. 2017.

ISACA, **Privacidade & Big Data**. 2013. Disponível em: <http://www.isaca.org/Knowledge-Center/Research/Documents/Privacy-and-Big-Data_whp_Portuguese_0913.pdf> Acesso em 20 Maio.2017.

Keegan, Richard e O'Kelly, Eddie. 2006. **Aplicar o benchmarking para a competitividade: guia prático para PME**. Lisboa: IAPMEI - Instituto da Empresa

KOTLER, Philip. **Administração de marketing: análise, planejamento, implementação e controle; tradução Ailton Bomfim Brandão**. 5.ed. São Paulo: Atlas, 1998.

LAKATOS, E. M.; MARCONI, M. A. **Fundamentos de metodologia científica**. 3. ed. São Paulo: Atlas, 2001.

M. J. Harry and J. R. Lawson, **Six Sigma Productibility Analysis and Process Characterization**, Addison-Wesley, 1992.

MANYIKA, J. **Big data: The next frontier for innovation, competition, and productivity.** Postado em 2011. Disponível em: <http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation> Acesso em: 05/04/2017

MCAFEE, A; BRYNJOLFSSON, E. **Big Data: The Management Revolution.** Harvard Business Review, edição de outubro de 2012. Médias Empresas, São Paulo, p. 30-43, set. 2013.

MICHEL, M. **Metodologia e pesquisa científica em ciências sociais.** São Paulo: Atlas, 2009.

MORAES, M. **Big Brother Obama.** InfoExame, edição de dezembro de 2012.

PORTER, M. **Vantagem competitiva: criando e sustentando um desempenho superior.** Rio de Janeiro: Campus, 1996.

PRAHALAD, C. K. **A atividade dos gerentes da nova era no emergente panorama competitivo.** In DRUCKER, Peter F. Foundation. **A Organização do Futuro: Como Preparar Hoje as Empresas de Amanhã.** sac Paula: Futura, 1997.

PRESS, Gil. **A Very Short History Of Big Data.** Postado em maio de 2013. Disponível em: <<http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/>> Acesso em: 21/02/2017

QUEIZOZ, Denilton. SANTOS, Renata. DA SILVA, Wanderson. **Benchmarking: A pesquisa que leva ao sucesso.** Disponível em: <<http://www.webartigos.com/artigos/benchmarking/53770/>> Acesso em: 29/03/2016

ROBLES JR, Antônio. BONELLI, Valério Vitor. **Gestão da Qualidade e do Meio Ambiente - enfoque econômico, financeiro e patrimonial.** São Paulo:Ed. Atlas SA, 2006.

Russom, P. (2013). **Managing Big Data**. *TDWI Best Practices Report Fourth Quarter 2013*.

SAS. *Homepage da Instituição*. Disponível em:
<<http://www.sas.com/offices/latinamerica/brazil/news/preleases/pesquisa-sasbigdata>.

SCHONBERGER-MAYER, Viktor; CUKIER Kenneth. Tradução Paulo Palzonoff Junior. **Big data: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana**. Rio de Janeiro: Elsevier, 2013

SHETTY, Y. K., **Aiming high: competitive benchmarking for superior performance, Long Range Planning**, vol.26, no.1, 1993, pp. 39-44.

Site Oracle. **A Base de inovação de dados: Big Data Corporativo**. Disponível em
< <https://www.oracle.com/br/big-data/index.html> > Acesso em: 21/03/2017

Site SAS. **Big data: O que é e porque é importante?** Disponível em:<
https://www.sas.com/pt_br/insights/big-data/what-is-big-data.html> Acesso em:
07/03/2017

Spendolini, M. J.. **Benchmarking**. Tr. Kátia Aparecida Roque, São Paulo: Makron Books. 1992.

SPENDOLINI, Michael J. **Benchmarking**. São Paulo: Makroon Books, 1993.

SPILLER, Eduardo Santiago et al. **Gestão de serviços e marketing interno**.1.ed. Rio de Janeiro: FGV, 2004.

Stela, Jefferson. **Tudo sobre Big Data: a revolução da análise de dados**. Disponível em: <<http://www.t-systems.com.br/imprensa-analistas/tudo-sobre-big-data-a-revolu-o-da-an-lise-de-dados/1303392>> Acesso em: 19/04/2017

THOMPSON, Arthur A; STRICKLAND A.J, GAMBLE, Jonh E. **Administração estratégica**. 15 ed. Porto Alegre: AMGH, 2011.

VICENTIN, David. GOLDFREIND, José. **Lean: Seis sigma e benchmarking**. Disponível em:

<http://www.setecnet.com.br/artigos/lean_seis_sigma_e_benchmarking.pdf>. Acesso em: 26/03/2016.

VIEIRA, Marcos. **Entendendo o Big Data**. 11/02/2014. Disponível em: <<https://www.ecommercebrasil.com.br/artigos/entendendo-big-data/>> Acesso em: 09/02/2017

VOLPATO Tiago. RUFINO, Ricardo Ribeiro. DIAS, Jaime William. **BIG DATA – TRANSFORMANDO DADOS EM DECISÕES**. UNIPAR. Disponível em: <http://web.unipar.br/~seinpar/2014/artigos/graduacao/Tiago_Volpato.pdf> Acesso em: 11/03/2017.

XIDIEH, Dolor Barbosa. **Benchmarking**. 2. ed. Campinas: Alínea, 2000.

ZABOT, J. M., SILVA, L. C. M. **Gestão do Conhecimento: aprendizagem e tecnologia: construindo a inteligência coletiva**. São Paulo: Editora Atlas, 2002.

GLOSSÁRIO

Ad hoc - Significa “para esta finalidade”, “para isso” ou “para este efeito”. É uma expressão latina, geralmente usada para informar que determinado acontecimento tem caráter temporário e que se destina para aquele fim específico.

Analytics - *Analytics* é a essência de *big data*. Trata a análise dos dados.

Business Intelligence - Inteligência empresarial. Refere-se ao processo de coleta, organização, análise, compartilhamento e monitoramento de informações que oferecem suporte a gestão de negócios.

Cloud computing - Refere-se à utilização da memória e da capacidade de armazenamento e cálculo de computadores e servidores compartilhados e interligados por meio da Internet, seguindo o princípio da computação em grade.

Customer Relationship Management - Gestão de Relacionamento com o Cliente. Criada para definir toda uma classe de ferramentas que automatizam as funções de contacto com o cliente.

Dashboard – É utilizado para indicar um painel de indicadores.

Hadoop Distributed File System - Sistema de arquivos distribuído que armazena dados em máquinas dentro do cluster, sob demanda, permitindo uma largura de banda muito grande em todo o cluster.

Hadoop File System - É uma estrutura de armazenamento de arquivos que utiliza blocos de 128 *Mbytes*, que são muito menores do que os blocos de particionamento tradicionais, utilizados em dispositivos de armazenamento.

Hardware – É a parte física de um computador, é formado pelos componentes eletrônicos, como por exemplo, circuitos de fios e luz, placas, utensílios, correntes, e

qualquer outro material em estado físico, que seja necessário para fazer com o que computador funcione.

High tech - De alta tecnologia, cuja tecnologia é muito avançada.

Lei de Moore - Até meados de 1965 não havia nenhuma previsão real sobre o futuro do *hardware* quando o então presidente da Intel, Gordon E. Moore, fez sua profecia, na qual o número de transistores dos chips teria um aumento de 100%, pelo mesmo custo, a cada período de 18 meses. Essa profecia tornou-se realidade e acabou ganhando o nome de Lei de Moore.

Machine Learning – Ou Aprendizado de Máquina é um método de análise de dados que automatiza o desenvolvimento de modelos analíticos. Usando algoritmos que aprendem interativamente a partir de dados, o aprendizado de máquinas permite que os computadores encontrem insights ocultos sem serem explicitamente programados para procurar algo específico.

Machine Learning - Trata o uso de algoritmos que identificam o melhor modelo para ser aplicado ao conjunto de dados.

MapReduce – Modelo de programação desenhado para processar grandes volumes de dados em paralelo, dividindo o trabalho em um conjunto de tarefas independentes.

Petabytes - É um múltiplo da unidade de informação byte. O prefixo *peta* indica a quinta potência de 1000 e significa 10^{15} no Sistema Internacional de Unidades (SI).

Radio frequency identification - A tecnologia de RFID nada mais é do que um termo genérico para as tecnologias que utilizam a frequência de rádio para captura de dados. Por isso existem diversos métodos de identificação, mas o mais comum é armazenar um número de série que identifique uma pessoa ou um objeto, ou outra informação, em um microchip.

Software - é uma sequência de instruções a serem seguidas e/ou executadas, na manipulação, redirecionamento ou modificação de um dado/informação ou acontecimento.

Stakeholders – Significa público estratégico e descreve uma pessoa ou grupo que tem interesse em uma empresa, negócio ou indústria, podendo ou não ter feito um investimento neles.

