

## **Big Data e Gestão do Conhecimento: Definições e Direcionamentos de Pesquisa.**

**JOSÉ CARLOS DA SILVA FREITAS JUNIOR**

Universidade Federal do Rio Grande do Sul  
freitas95@yahoo.com.br

**MIRIAN OLIVEIRA**

PUC-RS  
miriano@puers.br

**ANTONIO CARLOS GASTAUD MAÇADA**

Universidade Federal do Rio Grando do Sul  
acgmacada@ea.ufrgs.br

**RAFAEL ALFONSO BRINKHUES**

Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul  
rafabrink@gmail.com

## TECNOLOGIA DA INFORMAÇÃO

### **BIG DATA E GESTÃO DO CONHECIMENTO: DEFINIÇÕES E DIRECIONAMENTOS DE PESQUISA**

#### **Resumo**

O objetivo deste artigo é analisar como as pesquisas relacionam Big Data (BD) e Gestão do Conhecimento (GC). Há uma necessidade emergente das organizações buscarem novas formas de analisar e processar o grande volume de dados desestruturados que não podem ser analisados pelos meios tradicionais para gerar informações e conhecimento. Assim, para atingir o objetivo proposto foi realizada uma revisão sistemática de literatura, considerando artigos publicados em periódicos e congressos científicos. Nas bases pesquisadas, EBSCOhost e *Web of Science*, foram identificados 85 artigos. Os resultados foram apresentados em duas partes, a primeira onde se faz uma síntese dos dados da pesquisa, destacando as principais palavras-chaves, o número de publicações por fontes dos dados, dentre outras, e a segunda parte é relativa análise do conteúdo. Essa análise trouxe contribuições destacando a necessidade treinamento/capacitação de pessoas para analisar BD e gerar conhecimento, de desenvolvimento de ferramentas de GC para os ativos de Big Data. Foi verificada a tendência de pesquisas enfocando a área da saúde e o uso preponderante de metodologias de *Design Science* e experimento. Como sugestões de pesquisas propõem-se estudos sobre investimento na capacitação, desenvolvimento de cientistas de dados e sobre comportamento das pessoas no contexto de Big Data.

**Palavras-Chave:** Big Data; Gestão do Conhecimento; Revisão Sistemática.

#### **Abstract**

The aim of this article is to analyze research relating Big Data (BD) and Knowledge Management (KM). Organizations have been seeking new ways to analyze and process the large volume of unstructured data that cannot be analyzed by traditional means to generate information and knowledge. So, to achieve this purpose we developed a systematic literature review, considering articles published in journals and scientific conferences. We surveyed two databases, EBSCOhost and Web of Science, 85 articles were identified. The results were presented in two parts, the first is a synthesis of research data, highlighting the main keywords, the number of publications by the data sources, among others, and the second part is based on analysis of content. This analysis brought contributions highlighting the need of education and training people to analyze BD and generate knowledge. KM needs to develop tools for the assets of Big Data. The trend of researches hanging the health and the predominant use of methodologies and Design Science experiment has been verified. As suggestions of research we point out studies on investment in training, development of data scientists and the individual behavior in Big Data context.

**Key- Words:** Big Data; Knowledge Management; Systematic Review.

## 1. INTRODUÇÃO

As organizações estão vivenciando o crescimento do volume de dados e da informação. O termo Big Data (BD) tem sido usado para descrever o enorme volume de dados, que se caracteriza também por possuir um fluxo de desestruturado cuja infraestrutura existente nas organizações é incapaz de processar, tendo como suas principais características o alto volume, velocidade e variedade (DAVENPORT; BARTH; BEAN, 2012; ZIKOPOULOS et al., 2012).

O Big Data relacionado ao serviço de empresas de Internet tem crescido rapidamente. Por exemplo, o Google processa dados de centenas de petabyte (PB), o Facebook gera os dados de registro de mais de 10 PB por mês, Baidu, uma empresa chinesa, processa dados de dezenas de PB, e Taobao, uma subsidiária do Alibaba gera dados de dezenas de Terabyte (TB) para o comércio on-line por dia (CHEN; MAO; LIU, 2014). Como resultado, muitas empresas não estão sendo capazes de lidar com esse grande volume de dados e acabam perdendo dinheiro, conforme destaca a pesquisa realizada pela Oracle (2012) ao apontar que 93% dos executivos acreditam que suas organizações estão perdendo receitas - em média, 14% ao ano - como resultado de não ser capaz de alavancar totalmente as informações coletadas.

A relação entre conhecimento, informação e dados é destacada por Choo (2006) ao afirmar que o primeiro é a aplicação dos dois últimos, ou seja, o conhecimento tem como base a coleta de dados e informações adequadas, com a finalidade de ser útil à organização. AL-Ammary (2014) destaca que o conhecimento tem sido cada vez mais reconhecido como o ativo mais valioso nas organizações por ser um facilitador na formulação de uma estratégia e na busca por vantagem competitiva sustentável.

Logo, em um contexto de Big Data o grande desafio é transformar todo o volume de dados em informação a fim de gerar conhecimento com valor para as organizações, de forma a trazer vantagem competitiva para as mesmas (ERICKSON; ROTHBERG, 2014). Logo, o Big Data pode contribuir para com a Gestão do Conhecimento (GC) nas organizações. A GC pode ser entendida como um conjunto de processos que visam à criação, armazenamento, compartilhamento e utilização do conhecimento (CHYI LEE; YANG, 2000).

Assim sendo, o objetivo deste estudo é analisar como as pesquisas relacionam Big Data e Gestão do Conhecimento, visando identificar aspectos que venham a contribuir com a evolução dos estudos na área de Sistemas de Informação (SI) e Gestão do Conhecimento.

Para isso, o presente estudo está estruturado da seguinte forma: na seção 2 será apresentado o referencial teórico abordando a Gestão do Conhecimento e Big Data; na seção 3 será descrito o método adotado; na seção 4 é realizada a discussão dos resultados da pesquisa; e por fim, na seção 5 são trazidas as considerações finais.

## 2. REVISÃO DA LITERATURA

Esta seção destina-se a aprofundar o estudo teórico de dois campos da presente pesquisa, a Gestão do Conhecimento e Big Data.

### 2.1 GESTÃO DO CONHECIMENTO

Inicialmente, faz-se necessário um entendimento das relações básicas entre dados, informação e conhecimento. Segundo Batra (2014), esta relação hierárquica tem origens na área de tecnologia da Informação (TI), pois compreender a distinção entre dados e informações é um pré-requisito para a disciplina de sistema de informação.

Os dados em si, são a entrada para um sistema de processamento de dados (no qual a saída é a informação), é tipicamente gerada por meio de sistemas de processamento de transações. Sempre que uma decisão é tomada e que é cumprida, isso leva a uma transação ou uma série de operações que estão ocorrendo. Assim, o resultado secundário de qualquer transação concluída é a geração de dados, e dependendo da natureza do sistema de processamento de transações (seja on-line, em tempo real ou off-line) as informações como a saída também são atualizadas, e este processo cíclico continua (BATRA, 2014).

A informação é convertida para conhecimento uma vez que ela é processada na mente do indivíduo e o conhecimento se torna informação ao se articular e apresentar em forma de gráficos, textos, palavras e outros símbolos (ALAVI; LEIDNER, 2001). Assim sendo, a Gestão do Conhecimento tem se concentrado em aspectos do conhecimento que podem torná-lo mais fácil ou mais difícil de ser capturado e compartilhado como, por exemplo, por meio da transformação do conhecimento tácito em explícito (NONAKA; TAKEUCHI, 1997; POLANYI, 1967). Logo, para que as organizações sejam bem sucedidas elas devem estar capacitadas a capturar, integrar, criar e utilizar o conhecimento de maneira disciplinada, sistêmica e estratégica (SAMBAMURTHY; SUBRAMANI, 2005).

A criação de conhecimento organizacional reflete a capacidade de a organização gerar novos conhecimentos, distribuí-los e incorporá-los, Nonaka e Takeuchi (1997) defendem que esta abordagem compreende duas dimensões: uma epistemológica e outra ontológica. A dimensão epistemológica baseia-se na diferenciação entre o conhecimento tácito e o conhecimento explícito. Já a dimensão ontológica refere-se aos níveis de criação do conhecimento, como o individual, grupal, organizacional e interorganizacional (NONAKA; TAKEUCHI, 1997).

Davenport e Prusak (1999) definem conhecimento como uma mistura fluida de experiência, valores e informações que servem para gerar novas informações, ou seja, novas formas de conhecimento, dando ao conhecimento uma conotação similar a um sistema de feedback, ou seja, o conhecimento gera conhecimento. E que ao longo do tempo se desenvolve, seja por experiências, por livros, por formações curriculares ou mesmo por uma forma informal, como num diálogo entre amigos.

A Gestão do Conhecimento acontece quando a organização é capaz de reunir, armazenar e transferir conhecimentos, provendo o acesso às fontes de conhecimento por meio de mecanismos de identificação e recuperação de informações, visando à aplicação do conhecimento às necessidades específicas da organização (ALAVI; LEIDNER, 2001). Esses autores consideram uma classificação ampla contendo quatro processos básicos: criação, armazenamento/recuperação, transferência e aplicação de conhecimento. Tal proposta foi feita com o objetivo de investigar o papel da TI no suporte aos processos de gestão de conhecimento. Para estes autores, a criação de conhecimento ocorre de acordo com os quatro modos de conversão da teoria de Nonaka e Takeuchi (1997).

Hansen et al. (1999) sugerem que a estratégia de GC seja um mix de codificação (focada nas tecnologias) e de personalização (focada nos recursos):

- Estratégia de codificação: Visa capturar e codificar o conhecimento tácito em uma forma explícita (manuais, documentos, procedimentos) e deixá-lo disponível para a reutilização de toda a organização;

- Estratégia de personalização: Procura desenvolver as relações interpessoais de modo que estas promovam o compartilhamento do conhecimento tácito em toda a organização.

Todavia, o que se observa é o rápido crescimento dos vários esforços analíticos e de inteligência em disciplinas específicas indica que

Os ativos intangíveis potencialmente valiosos são encontrados em uma variedade de lugares, no interior e fora da empresa (ERICKSON; ROTHBERG, 2014). Isto pode indicar uma potencial ligação com o Big Data. Para ilustrar esse ponto, tem-se a hierarquia apontada

entre dados, informação e conhecimento, mas quando se fala em grande volume de dados, não está claro se essa relação se mantém. Logo, evidencia-se a necessidade de se avançar nos estudos sobre a relação entre GC e BD.

## 2.2 BIG DATA

Primeiramente, vale dizer que existem várias definições para o fenômeno Big Data, por isso e a fim de servir de base para a presente pesquisa o quadro 1 a seguir apresenta abordagens com definições de BD.

Quadro 1 – Definições de Big Data

Autores	Definições
Kim, Trimi, e Ji-Hyong (2014)	Trata-se de um termo geral para a enorme quantidade de dados digitais coletados a partir de todo os tipos de fontes.
Mahrt e Scharrow (2013)	Denotam um maior conjunto de dados ao longo do tempo, conjuntos de dados estes que são grandes demais para serem manipulados por infraestruturas de armazenamento e processamento regulares.
Davenport (2012) e Kwon (2014)	Dados demasiadamente volumosos ou muito desestruturados para serem gerenciados e analisados através de meios tradicionais
Di Martino et al. (2014)	Refere-se ao conjunto de dados cujo tamanho está além da habilidade de ferramentas típicas de banco de dados em capturar, gerenciar e analisar.
Rajesh (2013)	São conjuntos de dados que são tão grandes que se tornam difíceis de trabalhar com o uso de ferramentas atualmente disponíveis.
Mahrt e Scharrow (2013)	Denotam um maior conjunto de dados ao longo do tempo, conjuntos de dados estes que são grandes demais para serem manipulados por infraestruturas de armazenamento e processamento regulares.

Fonte: Adaptado de Freitas e Maçada (2014)

Baseando-se no Quadro 1, é possível notar uma similaridade nos conceitos, que podem ser sintetizados como sendo um grande volume de dados estruturados ou não, de fontes diversas, que devem ser gerenciados e analisados de forma peculiar.

Assim sendo, o desafio das organizações no gerenciamento dessa quantidade de dados é extrair informações estruturadas. As empresas que aprenderem a aproveitar oportunidades em “Big Data” poderão utilizar informações em tempo real a partir de sensores. A identificação por radiofrequência e outros dispositivos de identificação atuarão para que as organizações compreendam melhor seus ambientes de negócios em um nível mais granular, a fim de criar novos produtos e serviços, e/ou para responder às mudanças nos padrões de uso à medida que ocorrem (DAVENPORT; BART; BEAN, 2012).

Segundo Di Martino et al. (2014) as fontes de dados são as mais diversas, segundo o autor os dados são gerados por vários tipos de recursos, tais como dispositivos móveis, celulares, sensores, portais sociais, internet das coisas, dados da própria empresa, câmeras digitais e sistemas de monitoramento, registros de software, dentre outros. Logo, que as organizações terão de desenvolver processos contínuos para acessar essas fontes de dados, analisá-los e interpretá-los (DAVENPORT; BART; BEAN, 2012).

Vale ressaltar que o Big Data, na sua grande maioria, é um conjunto de dados que a cada milésimo de segundo são inseridos novos, assim sendo, o mundo atual dos dados digitalizados contém uma inimaginável quantidade de informações, base que está em crescimento permanentemente (DAVENPORT, 2012). Além disso, existem outras características apontadas por diversos autores, mas para o presente estudo serão consideradas a seguintes: volume, variedade e velocidade, denominada 3 Vs, que seriam as mais comumente citadas (DI MARTINO et al., 2014; MCAFEE; ZIKOPOULOS et al., 2012).

O processo gestão de Big Data tem como entrada as múltiplas fontes de dados, e suas características (volume, variedade e velocidade) são relevantes desde a geração até o

processamento dos dados. Assim, cabe às organizações o trabalho de selecionar, processar, analisar e escolher os dados que irão gerar as informações com veracidade e oportunidade (FREITAS JUNIOR, MAÇADA, BRINKHUES, DOLCI, 2015).

Nesse sentido, empresas, nos mais diversos setores, podem usar o Big Data como um recurso, uma vez que cada dispositivo eletrônico envia e o consumidor deixa um rastro, assim é possível analisar esses conjuntos de dados para beneficiar clientes e mercados, e otimizar as decisões de negócios (KIM; TRIMI; JI-HYONG, 2014). Entretanto, encontrar pessoas com habilidades para analisar Big Data é, talvez, o maior desafio para as organizações empresariais. As competências exigidas não são simples e nem exclusivamente orientada para a tecnologia. As empresas recebem muitos dados e tem pouco tempo de processá-los e transformá-los em informação e conhecimento. Cabendo a elas tornarem-se provedoras de percepção, usando a análise de dados para processar a informação e dizer-nos o que fazer com ela. (GARTNER, 2012).

Além das empresas prestadoras de serviços e das indústrias, uma área que vem aprofundando estudos sobre o Big Data é a área da Saúde. De acordo com Splendiani (2011) as pesquisas nessa área dependem cada vez mais de grandes coleções de conjuntos de dados e conhecimento, cuja geração, representação e análise muitas vezes exigem grandes esforços colaborativos e interdisciplinares. Com isso surge uma nova dimensão de pesquisa, pesquisa em Big Data, que visa o desenvolvimento de ferramentas computacionais para gerenciar uma grande quantidade de dados, bem como melhorar a colaboração e comunicação da comunidade em geral, tendo que por vezes iniciar um movimento de mudança na cultura da empresa.

Por fim, os dados estão amplamente disponíveis, são estrategicamente importantes, assim como o conhecimento que se pode extrair dos dados e, por conseguinte a sabedoria oriunda dele (MAYER-SCHÖNBERGER; CUKIER 2013). Logo, acredita-se que os estudos e pesquisas devem verificar se a presença ou não das características do Big Data se relacionam com a Gestão do Conhecimento, o que será feito a seguir.

### 2.3 BIG DATA e GESTÃO DO CONHECIMENTO

Ao considerar os conceitos de Gestão do Conhecimento e Big Data abordados anteriormente, percebe-se a conexão entre os mesmos. A Gestão do Conhecimento contempla processos para criação e compartilhamento do conhecimento, enquanto o foco do Big Data é gerar conhecimento a partir de um conjunto de dados. Esta intersecção pode trazer benefícios para as organizações, o que torna relevante investigar a relação entre eles, considerando que, segundo Davenport e Prusak (1999), o conhecimento é responsável pelo funcionamento das organizações.

O conhecimento originado do Big Data é um recurso único, valioso, para qualquer organização sendo capaz de gerar vantagem competitiva sustentável. Big Data também ajuda a transformar muitos dos métodos tradicionais de realização de atividades empresariais, gerando insights que se transformarão em conhecimento que virá a impulsionar capacidade da administração para a tomada de decisões (KABIR; CARAYANNIS, 2013).

Corroborando com a importância de avançar nos estudos que relacionam GC e BD, Manyika et al. (2011) apresenta um relatório da *McKinsey Global Institute* (MGI), que fornece uma estimativa da "Facilidade de Captura" do potencial de valor do Big Data para cada setor. A estimativa é baseada em quatro indicadores com uma relação com conceitos comuns de conhecimento.

- Talento (*Talent*), que estaria intimamente relacionada com a compreensão comum do capital intelectual humano, com ênfase no conhecimento tácito como o talento individual ou know-how que pode ser difícil para compartilhar;

- Mentalidade orientada por dados (*Data-driven mindset*) remonta ao capital humano, especificamente o conhecimento de gestores e líderes da empresa;
- Intensidade de TI (*IT intensity*), que tem uma conexão com o capital estrutural, ou seja, a estrutura de TI da empresa para o gerenciamento de dados, informações e conhecimento também é uma parte substancial do capital estrutural;
- A disponibilidade de dados (*Data availability*) é o indicador que não está diretamente relacionado com o conhecimento, mas sim de forma indireta, pois diz respeito aos precursores de conhecimento: os dados e informações.

Assim, com base nos aspectos elencados nesta seção, a seguir será exposta a operacionalização das etapas dessa pesquisa, buscando verificar as relações para a investigação entre Big Data e Gestão do Conhecimento.

### 3. MÉTODO

A presente pesquisa é de natureza qualitativa enfatizando “os aspectos subjetivos da atividade humana focando o significado, e não a mensuração, de fenômenos sociais” (COLLIS; HUSSEY, 2005, p. 59). Para atingir o objetivo proposto, qual seja, analisar como as pesquisas relacionam Big Data e Gestão do Conhecimento, este estudo seguiu os passos metodológicos propostos por Wolfswinkel et al. (2013) para trazer rigor à pesquisa.

A escolha desses passos se justifica pela dificuldade de estruturação e apresentação uma revisão da literatura dos resultados da pesquisa, de forma pertinente, trazendo relevância e qualidade aos resultados, conforme destacam Webster e Watson (2002). O quadro a seguir ilustra os cinco passos do método proposto.

Quadro 2 - método de cinco passos do método para a revisão da literatura

Ordem	Tarefas
1. DEFINIÇÃO	
1.1	Definir os critérios de inclusão / exclusão
1.2	Identificar as áreas de pesquisa
1.3	Determinar as fontes apropriadas
1.4	Decidir sobre os termos específicos de pesquisa
2. PESQUISA	
2.1	Pesquisa (coleta dos dados)
3. SELEÇÃO	
3.1	Refinar a amostra
4. ANÁLISE	
4.1	A codificação aberta
4.2	Codificação Axial
4.3	Codificação seletiva
5. APRESENTAÇÃO	
5.1	Representar e estruturar o conteúdo
5.2	Estrutura do artigo

Fonte: Adaptado de Wolfswinkel et al. (2013)

Com base nesses passos, esta seção explicita como a pesquisa foi realizada, expondo os procedimentos tomados pelo pesquisador a fim de que seja possível verificar o rigor desta pesquisa.

#### 3.1 DEFINIÇÃO

No passo “definição”, foram definidos os critérios de inclusão e exclusão; identificadas as áreas de pesquisa; determinadas as fontes apropriadas; e, decidido sobre os termos específicos de pesquisa.

Inicialmente, foi feita uma revisão bibliográfica que permitiu que o pesquisador pudesse verificar possíveis palavras-chave, que seriam mais relevantes para o tema, assim, conforme se pôde verificar no item 2.3 como o que se pretende é verificar aspectos que circundam as relações entre Big Data e Gestão do Conhecimento. Em seguida, foi realizada uma pesquisa piloto com esses temas no idioma em português, não houve um retorno efetivo, o que resultou foram vários artigos sobre Big Data no idioma inglês, mas não houve representatividade de artigos em português, assim, optou-se por utilizar os termos em inglês, entre aspas, com o modo booleano, separado pelo conectivo em inglês *and*.

Assim sendo, para a coleta de dados, foram utilizadas as seguintes palavras-chaves: “*Big Data*” and “*Knowledge Management*”. Como o referencial teórico indica a importância do compartilhamento da informação como sendo um tema presente na GC, buscou, também, e “*Big Data*” and “*Knowledge Sharing*”.

As buscas iniciais foram realizadas em várias bases, com os mesmos termos, vindo a optar pelas que apresentaram os resultados mais consistentes em número e qualidade que pudessem sustentar a presente pesquisa. Optou-se pela bases Elton Bryson Stephens Company (EBSCOhost) e *Web of Science* (WOS). Além do fato de terem apresentado mais resultados, a escolha da primeira se deve também ao fato de contemplar várias bases como a *Academic Source Complete* (ASC), *Academic Search Premier* (ASP), *Business Source Complete* (BSC), dentre outras, e a segunda, por proporcionar o acesso a periódicos múltiplos com altos fatores de impactos, conforme destaca Barata (2010).

Os links de acesso serão apresentados a seguir:

- EBSCO - <<http://web.a.ebscohost.com/ehost/search/selectdb?sid=d9c1110c-5419-4abb-a903-864917c378bf%40sessionmgr4001&vid=0&hid=4206>>
- WEB OF SCIENCE - <[http://apps.webofknowledge.com.ez45.periodicos.capes.gov.br/WOS\\_GeneralSearch\\_input.do?product=WOS&search\\_mode=GeneralSearch&SID=2DuJLb6Z5i31qEqiJtu&preferencesSaved=>](http://apps.webofknowledge.com.ez45.periodicos.capes.gov.br/WOS_GeneralSearch_input.do?product=WOS&search_mode=GeneralSearch&SID=2DuJLb6Z5i31qEqiJtu&preferencesSaved=>)>

Para a busca nas duas bases os termos e restrições foram os mesmos, período de 2004 a 2014, quanto ao idioma não houve restrição, entretanto todos os resultados obtidos foram em língua inglesa. As variações ocorreram devido à peculiaridade e modo de busca de cada base, o que será detalhado a seguir.

Na primeira base de dados, EBSCO Host, foram selecionadas as seguintes fontes de dados/bases de dados: “*Academic Search Elite*”, “*Library, Information Science & Technology Abstracts with Full Text*”, “*Academic Search Premier*” e “*Business Source Complete*”. O passo seguinte foi a seleção de pesquisa avançada, que permite que se insira mais de um termo de busca, utilizando-se o modo de pesquisa booleano/frase. Em seguida, para cada uma das fontes acima, foi selecionado a busca por *Journals*, periódicos e artigos de congressos. A próxima restrição foi o tipo de documento, tendo sido selecionado o termo “tudo” e marcado “texto completo em pdf” para as seguintes fontes “*Academic Search Premier*” e “*Academic Search Elite*”, pois eram as únicas que disponibilizaram esta última opção de seleção.

Na segunda base de dados, a Web of Science, foi selecionada a opção busca básica que já permitiu a inserção dos termos de busca, utilizando-se o modo de pesquisa booleano/frase. Selecionou-se o período, já destacado, não houve mais nenhuma restrição seguindo-se a busca com as configurações default padrão da base.

### 3.2 PESQUISA e SELEÇÃO

Os dois passos seguintes propostos por Wolfswinkel et al. (2013) ocorreram de forma integrada. A cada busca a amostra foi refinada para verificar possíveis duplicidades e outros problemas. A pesquisa foi realizada no dia 10 de janeiro de 2015.

Na primeira busca, na base EBSCOHost com os termos “Big Data” and “*Knowledge Management*” foram obtidas 40 (quarenta) publicações com as seguintes fontes, Revistas Acadêmicas (23), *Journals* (16), Publicações comerciais (14) Revistas (3). Já a busca com os termos “Big Data” e “*Knowledge Sharing*” foram obtidos 2 resultados, em revistas acadêmicas. Totalizando no primeiro momento 42 (quarenta e dois artigos). Entretanto, após verificar as duplicidades e excluir publicações que não eram científicas, permaneceram 29 publicações.

Na segunda busca, com a base do Web of Science, com os termos “Big Data” e “*Knowledge Management*” foram obtidos 55 (cinquenta e cinco) artigos, e com os termos “Big Data” e “*Knowledge Sharing*” foram obtidos 4 (quatro) artigos. Porém, dois desses já constavam na pesquisa realizada na outra base, assim, após excluir duplicidades e excluir publicações que não eram científicas, permaneceram 57 (cinquenta e sete) artigos. Desse total, 50 artigos são de conferências e as outras 7 (sete) são de periódicos. Portanto, somando as duas bases permaneceram 85 (oitenta e cinco) artigos para análise.

### 3.3 ANÁLISE

Os dados foram analisados seguindo os conceitos de análise de conteúdo, obedecendo às etapas de redução, de exibição e de conclusão e verificação (GIL, 2009). Assim, buscou-se analisar primeiramente cada uma das bases, e depois uni-las realizando uma terceira análise global. Para tanto, utilizou-se o software Weft QDA. Os dados foram categorizados inicialmente seguindo os indicadores apontados por Manyika et al. (2011) no relatório da McKinsey Global Institute, apresentado no item 2.3 deste estudo. Durante a análise dos dados outras categorias emergiram nesta fase.

### 3.4 APRESENTAÇÃO

Este último passo do método sugere a representação e estruturação do conteúdo e a estruturação do artigo. Logo, esta fase é o resultado da pesquisa em si, expresso neste artigo, sendo que a representação e estruturação do resultado da pesquisa em si serão apresentadas na seção 4 a seguir.

## 4. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Esta seção visa analisar os resultados obtidos na presente pesquisa que busca analisar como as pesquisas relacionam Big Data e Gestão do Conhecimento.

### 4.1 CONTEXTUALIZAÇÃO DOS ARTIGOS PUBLICADOS

O primeiro resultado a ser destacado é a evolução das publicações destes dois assuntos juntos ao longo do tempo, sendo o período de busca foi de 2004 a 2014. No Quadro 3 são apresentados os dados obtidos, verifica-se que é um tema recente, com o montante maior no ano de 2013, o que se justifica pelas conferências ocorridas naquele ano.

Quadro 3 – Frequência das publicações por ano

Ano	Frequência
2007	1
2008	0
2009	0
2010	0
2011	1
2012	5
2013	62
2014	16

Fonte: EBSCOHost e *Web of Science*

No Quadro 4 os artigos são agrupados segundo as conferências internacionais encontradas na pesquisa da base Web of Science. Observa-se que a maioria dos artigos, 38 (trinta e oito) dentre as 50 (cinquenta), referem-se a assuntos de saúde, o que indica o crescimento de importância do estudo da Gestão do Conhecimento e do Big Data para essa área, indo ao encontro do que foi destacado no referencial teórico.

Quadro 4 - Conferências

Títulos da conferência	Número de publicações
Conference On Ehealth - Health Informatics Meets Ehealth	38
Ieee International Conference On Big Data Big Data	2
10th International Conference On Intellectual Capital Knowledge Management And Organizational Learning Ickm	2
International Conference On Collaboration Technologies And Systems	1
Conference On Next Generation Analyst	1
9th International Conference On Hybrid Artificial Intelligence Systems	1
9th Ieee International Symposium On Applied Computational Intelligence and Informatics	1
7th International Conference On Knowledge Management In Organizations Kmo 2012	1
2nd International Symposium On Computer Communication Control And Automation 3ca	1
23rd International Business Information Management Association Conference On Visio 2020 Sustainable Growth Economic Development And Global Competitiveness	1
11th International Conference On Artificial Intelligence And Soft Computing Icaisc	1
Total	50

Fonte: *Web of Science*

No Quadro 5 estão relacionados os periódicos onde os artigos foram publicados, destaca-se a revista “*KM World*” que possui 12 (doze) artigos. Acredita-se que essa concentração se deve ao fato de a mesma ser específica para o tema gestão do conhecimento.

Quadro 5- Fonte dos Dados

Fonte ( Journal e Revistas )	Número de publicações por fonte
<i>KM World</i>	12
<i>Briefings in Bioinformatics</i>	01
<i>Computing And Informatics</i>	01
<i>CyberPsychology, Behavior &amp; Social Networking</i>	01
<i>El Profesional de la Información</i>	01
<i>Electronic Journal of Knowledge Management.</i>	01

<i>Feliciter</i>	01
<i>Financial Executive</i>	01
<i>Information Management Journal</i>	01
<i>Information Services &amp; Use</i>	01
<i>International Journal of Database Theory &amp; Application</i>	01
<i>International Journal of Information, Business &amp; Management.</i>	01
<i>International Journal of Pattern Recognition &amp; Artificial Intelligence</i>	01
<i>International Journal of Software Engineering &amp; Knowledge Engineering</i>	01
<i>IUP Journal of Knowledge Management</i>	01
<i>Journal Of Business Logistics</i>	01
<i>Marketing Insights</i>	01
<i>Multimedia Tools And Applications</i>	01
<i>Online Searcher</i>	01
<i>Profesional De La Informacion</i>	01
<i>Research Technology Management</i>	01
<i>Review of Management</i>	01
<i>Scientific Programming</i>	01
<i>Service Industries Journal</i>	01
Total	35

Fonte: EBSCOHost e Web of Science

As principais palavras chaves encontradas nas publicações pesquisadas estão apresentadas em ordem ocorrência no Quadro 6. Pode-se observar que os termos da busca principal da pesquisa possuem maior número de ocorrências.

Quadro 6 – Palavras-chave

Palavras-chave	Número de ocorrências	Palavras-chave	Número de ocorrências
<i>knowledge management</i>	22	<i>customer relationship management</i>	3
<i>big data</i>	13	<i>database searching</i>	3
<i>Information</i>	7	<i>Repositories</i>	3
<i>data mining</i>	6	<i>business intelligence</i>	3
<i>Knowledge</i>	7	<i>data integration</i>	3
<i>open data</i>	5	<i>decision making</i>	3
<i>Prediction</i>	5	<i>demand forecasting</i>	2
<i>big data analysis</i>	4	<i>Generators</i>	2
<i>Analytics</i>	4	<i>granularities of knowledge contente</i>	2
<i>Data</i>	4	<i>Inconsistencies</i>	2
<i>information technology</i>	4	<i>intellectual capital</i>	2
<i>big data integration</i>	3	<i>Prediction</i>	2
<i>business analytics</i>	3	<i>Wisdom</i>	1

Fonte: elaborado pelos autores

#### 4.1.1 Métodos Utilizados

No que diz respeito às metodologias de pesquisa utilizada nas publicações, verificou-se que em muitas publicações a mesma não foi explicitada. Dentre as que foram apresentadas, destaca-se a elevada ocorrência de revisões sistemáticas de literatura. Isso pode ser explicado pela emergência recente do tema.

Todavia, foi observado o preponderante uso de metodologias de *Design Science* e experimento, também justificado pela característica do tema já destacado. Sendo que a maior incidência foi o uso do *Design Science*, com 28 (vinte e oito) ocorrências. Novamente, destacaram-se as pesquisas na área da saúde com a utilização dessa metodologia, nas quais busca a construção e desenvolvimento de arquiteturas de base de dados e de pesquisa.

#### 4.1.2 BD, GC e Saúde

Esta categoria emergiu ao se verificar o surgimento de uma nova tendência de pesquisa na área de saúde, relacionando os três temas apresentados no título deste item. Este aspecto já havia sido apontado no referencial teórico, vindo a ser comprovada na análise dos resultados.

Verificou-se, ainda, que das pesquisas analisadas, as principais publicações sobre GC, BD e Saúde foram em conferências. Trata-se de 38 (trinta e oito) trabalhos que propõem sistemas de “eHealth architecture” (para prover ferramentas de pesquisa que visam reunir uma base de dados grande para a investigação na área das doenças raras). Como por exemplo o estudo denominado “*A Generic, Expandable Architecture for Data Acquisition for Medical Research for Example Haemophilia*” apresentado na conferência *Health Informatics Meets Ehealth*.

### 4.2 ANÁLISE DA RELAÇÃO ENTRE GC E BD

Inicialmente, os dados foram categorizados de acordo com os indicadores propostos pela *McKinsey Global Institute* (MANYIKA et al., 2011) que fornece uma estimativa da "Facilidade de Captura" do potencial de valor do Big Data devido a relação com conceitos comuns de conhecimento, são eles: o talento; intensidade de TI; mentalidade orientada por dados e a disponibilidade de dados. Além desses, observou-se a ocorrência do tema saúde, e buscou-se conhecer os principais métodos e sugestões de pesquisa, estes três últimos aspectos constituem três categorias que emergiram durante a análise dos dados.

#### 4.2.1 Talento

O talento está intimamente relacionado com a compreensão comum do capital intelectual humano, a ênfase está no conhecimento tácito, ou seja, como o talento individual ou *know-how* que pode ser difícil de ser compartilhado (MANYIKA et al., 2011). No que tange a essa categoria, pode-se destacar que os estudos apresentados no periódico *CyberPsychology, Behavior & Social Networking* apontam para a necessidade de investimento também nas pessoas, para capacitá-las a fim de que elas sejam capazes de ter insights, permitindo a integração de experiências sociais de larga escala e insights comportamentais (GAGGIOLI, 2013).

Além disso, os trabalhos analisados apontam para a possibilidade de se obter conhecimento a partir das experiências dos clientes, o que implica em pesquisas sobre predição, um dos principais exemplos é a análise do perfil do cliente de cartões de crédito e

com isso é possível conhecer as preferências individuais e, por conseguinte, oferecer serviços personalizados. Isto requer prática para o desenvolvimento do *know-how* (GHOCHANI, 2013).

Para isso, existe uma necessidade urgente de sistemas capazes de extrair informação útil a partir de uma massa de informações e grande volume de dados, e sistemas que permitam a visualização dos mesmos, tais sistemas também poderão auxiliar na transformação do conhecimento tácito para o explícito, bem como o desenvolvimento do capital intelectual humano para a geração do conhecimento necessário (LARA-NAVARRA; SERRADELL-LÓPEZ; MANIEGA-LEGARDA, 2013; LAMONT, 2012)

#### **4.2.2 Mentalidade Orientada por Dados**

A Mentalidade orientada por dados, assim como o talento já analisado, remonta ao capital humano, especificamente o conhecimento de gestores e líderes da empresa de acordo com Manyika et al. (2011).

Segundo Kabir e Carayannis (2013) as organizações devem considerar o conhecimento “escondido” no Big Data como o conhecimento tácito, tais autores argumentam que se deve aproveitar a experiência acumulada pelos membros das empresas para que sejam desenvolvidos processos e técnicas de armazenamento e análise de dados para a geração de conhecimento. Além disso, os referidos autores destacam a importância do investimento no desenvolvimento das pessoas responsáveis pelas decisões, ou seja, líderes e gestores capazes de extrair valor dos dados e informações, a fim de gerar conhecimento.

#### **4.2.3 Intensidade de TI**

Segundo o relatório da *McKinsey Global Institute* de 2011, a intensidade de TI tem uma conexão com o capital estrutural, ou seja, a estrutura de TI da empresa para o gerenciamento de dados, informações e conhecimento também é uma parte substancial do capital estrutural.

Esta categoria é corroborada pelos autores Kabir e Carayannis (2013), Sukumar e Ferrell (2013) e Batra (2014), que destacam que a importância das empresas desenvolverem ou adquirirem ferramentas, tecnologias e outras propostas de soluções para que o Big Data possa gerar informação e essa informação seja convertida em conhecimento para as organizações, uma vez que muitas organizações estão obtendo vantagem do *business analytics* e soluções de inteligência para ajuda-las a encontrar novos insights em seus processos melhorando assim sua performance.

Além das publicações em periódicos, as conferências também destacam a importância do investimento da estrutura de TI para a geração do conhecimento a partir do Big Data. Isto pôde ser observado pelos estudos que destacam técnicas emergentes de computação em nuvem e modelo de computação big data, como o MapReduce, conforme se pôde observar no artigo denominado “*IntegrityMR: Integrity Assurance Framework for Big Data Analytics and Management Applications*” publicado na conferência “*2013 IEEE INTERNATIONAL CONFERENCE ON BIG DATA*”.

Em síntese, os estudos apontam a necessidade de uma estrutura de TI capaz de armazenar um grande volume de dados, de processar e de recuperar dados e informações combinando com soluções de *Business Intelligence* (BI) e/ou *Business Analytics* (BA), com isso, acredita-se ser mais possível obter o conhecimento necessário para a organização (LAMONT, 2012; JOHNSON, 2012).

#### 4.2.4 Disponibilidade de Dados

A disponibilidade de dados é um indicador que não é diretamente relacionado com o conhecimento, mas com os antecedentes do conhecimento, ou seja, os dados e informações (MANYIKA et al., 2011). Os estudos apontam a necessidade da obtenção de valor a partir dos dados, segundo Kabir e Carayannis (2013) para que isso ocorra é necessário que a coleta seja realizada com objetivos específicos. Ainda segundo estes autores, atualmente, com o crescimento na quantidade de informação e a possibilidade da utilização dos dados no futuro, combinado com a diminuição de custos de recursos de armazenamento e conectividade, as organizações estão se concentrando no Big Data, mas para isso devem gerar valor a partir dos mesmos.

Erickson e Rothberg (2014) vão ao encontro das afirmações acima ao reconhecer o BD como um ativo valioso, sugerindo a necessidade de se estabelecer uma base conceitual de Big Data. Como ponto de partida, os autores sugerem a aplicação de métricas de capital intelectual e ferramentas de gestão do conhecimento para os ativos de Big Data.

Em síntese, pôde-se observar ainda que as publicações pesquisadas que buscaram desenvolver uma base teórica para o estudo de GC e BD apontam para a necessidade de se considerar o valor do Big Data. Além disso, os estudos indicam que as fontes de dados também têm aumentado assim como tamanho de dados. Logo, os dados em si tornam-se um recurso fundamental para uma empresa, pois oportunidades deliberadas são criadas para gerar dados dos próprios dados, de forma crescente, podendo criar valor e proporcionando assim uma vantagem competitiva para a empresa (SUKUMAR; FERRELL, 2013; BATRA, 2014).

Síntese da Relação entre BD e GC

Categorias	Relação Entre GC e BD	Implicações
Talento ( <i>Talent</i> )	Capital Intelectual	Treinamento/ capacitação de pessoas para analisar o BD e gerar conhecimento
Mentalidade orientada por dados	Necessidade de desenvolvimento do capital humano,	
Intensidade de TI	A estrutura de TI da empresa	Necessidade de investimento em TI desenvolvendo uma estrutura capaz
A disponibilidade de dados	Geração de Valor dos dados	Necessidade de desenvolvimento de ferramentas de gestão do conhecimento para os ativos de Big Data.

Fonte: elaborado pelos autores

#### 4.3 SUGESTÕES DE PESQUISA

Para as empresas, de uma forma geral, a GC em um contexto de BD ainda é um terreno incipiente. Muitas delas entendem que existem mais conhecimento e as informações que podem ser extraídos a partir de dados disponíveis em Big Data usando criatividade, recombinação e métodos inovadores, produzindo um valor substancial, porém há a necessidade de encontrar uma abordagem adequada para alinhar o Big Data com a estratégia da organização (KABIR; CARAYANNIS, 2013).

Assim, foi possível verificar nesse estudo várias oportunidades de novas abordagens e algumas sugestões de pesquisas futuras que possam contribuir para o avanço da pesquisa na área. A seguir serão listadas algumas dessas sugestões:

- Buscar o estado da arte a fim de se propor soluções para a gestão de grande volume de dados (KOSCIEJEW, 2013);
- Geração de valor dos dados no contexto de BD (ERICKSON; ROTHBERG, 2014);

- Big Data, gestão do conhecimento e estratégia da organização (SUKUMAR; FERRELL, 2013; BATRA, 2014);
- Como a granularidade pode impactar no conteúdo de Big Data, na extração das informações e geração do conhecimento (ZHANG, 2013);
- Pesquisas sobre investimento na capacitação, desenvolvimento de cientistas de dados e sobre comportamento das pessoas no contexto de Big Data (GAGGIOLI, 2013);
- Estudo sobre predição e gestão do conhecimento a partir do comportamento de clientes (GHOCHANI, 2013; LAMONT, 2012);
- Implicações de Big Data Analytics (KABIR; CARAYANNIS, 2013);
- Soluções de *Business Intelligence* (BI) para de gerenciamento de dados e gestão do conhecimento (LAMONT, 2012; JOHNSON, 2012).

Essas propostas sintetizam algumas oportunidades e sugestões que tiveram uma ocorrência maior nesta análise de dados e possuem um caráter mais genérico, sabe-se que existem outras mais específicas a cada estudo, como por exemplo, relacionadas a já destacada questão de saúde.

## 5. CONSIDERAÇÕES FINAIS

Esse estudo verificou as conexões naturais existentes entre Gestão do Conhecimento e Big Data. Kabir e Carayannis (2013) evidenciam que esta conexão é clara, mas por se tratar de um tema novo, o desafio está em encontrar as maneiras de relacionar a Gestão do Conhecimento em um contexto de grande volume de dados (ERICKSON; ROTHBERG, 2014).

Este desafio foi um dos norteadores desta pesquisa que teve como objetivo identificar as relações entre Big Data e a Gestão de Conhecimento, a fim de verificar as possíveis oportunidades de pesquisa. Tal finalidade foi alcançada a partir do direcionamento apresentado na revisão de literatura e por meio da análise da busca realizada nas bases de pesquisa científica de trabalhos sobre o tema.

Os resultados apresentados e a análise contribuem tanto para a pesquisa acadêmica quanto para a prática gerencial. Existem vários desafios ainda a serem vencidos, mas com o aumento de pesquisas nessa área as organizações poderão gerenciar o conhecimento com mais qualidade e agregar valor aos seus negócios.

Do ponto de vista gerencial, a sintetização dos trabalhos sobre o tema propõe reflexões de como obter valor dos dados neste contexto, bem como se pensar em tecnologias, ferramentas, enfim, soluções para que se possa gerar e gerenciar conhecimento no contexto de Big Data. Além disso, observa-se que além de contexto, o Big Data pode ser estudado como um recurso ou ativo que deve estar alinhado a estratégia da empresa para gerar vantagem competitiva.

Pelo lado acadêmico, por se tratar de um tema emergente, foram apresentadas várias oportunidades e sugestões de pesquisa relacionando GC e BD. Além desses direcionamentos, essa pesquisa contribui para a literatura de ambos os temas de apresentando uma abordagem conceitual dos mesmos.

## 6. REFERÊNCIAS

ALAVI, M.; LEIDNER, D.E. Review: knowledge management and knowledge management systems: conceptual foundations and research issues. **MIS Quarterly**, v. 25, n. 1, p. 107-136, 2001.

- AL-AMMARY, J. The Strategic Alignment between Knowledge Management and Information Systems Strategy: The Impact of Contextual and Cultural Factors. **Journal of Information & Knowledge Management**. 13, 1, 1450006-1-1450006-12, Mar. 2014.
- BARATA, Germana. América Latina e o impacto de suas publicações científicas. **Cienc. Cult.** São Paulo, v. 62, n. 2, 2010. Disponível em <[http://cienciaecultura.bvs.br/scielo.php?script=sci\\_arttext&pid=S0009-67252010000200007&lng=en&nrm=iso](http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252010000200007&lng=en&nrm=iso)>. Acesso em 15 Mar. 2015.
- BATRA, S. Big Data Analytics and its Reflections on DIKW Hierarchy. **Review of Management**. 4, 1/2, 5-17, June 2014
- CHEN, Min; MAO, Shiwen; LIU, Yunhao. Big data: A survey. **Mobile Networks and Applications**, v. 19, n. 2, p. 171-209, 2014.
- CHYI LEE, Ching; YANG, Jie. Knowledge value chain. **Journal of management development**, v. 19, n. 9, p. 783-794, 2000.
- CHOO, C. W. The knowing organization: How organizations use information to construct meaning, create knowledge, and make decisions. **Oxford University Press**, USA, 2006.
- COLLIS, J.; HUSSEY, R. **Pesquisa em administração: um guia prático para alunos de graduação e pós-graduação**. 2. ed. Porto Alegre: Bookman, 2005.
- DAVENPORT, T. H. Enterprise analytics: Optimize performance, process, and decisions through big data. **Upper Saddle River**, New Jersey: FT Press OperationsManagement, 2012.
- DAVENPORT, T. H; BART, P.; BEAN, R. How Big Data is Different. **MIT Sloan Management Review**, July 30, pp. 43-6, 2012.
- DAVENPORT, T. H.; PRUSAK, L. Conhecimento empresarial. São Paulo: Publifolha, 1999.
- DI MARTINO, B. et al. Big data (lost) in the cloud. **International Journal of Big Data Intelligence.**, Vol.1, No.1/2, pp.3 – 17, 2014
- ERICKSON, S.; ROTHBERG, H. Big Data and Knowledge Management: Establishing a Conceptual Foundation. **Electronic Journal Of Knowledge Management**, 12(2), 101-109, 2014.
- FREITAS JUNIOR, J. C. S.; MAÇADA, A. C. G. **Processo decisório no contexto de big data: estudo de caso em uma empresa do varejo**. 2014. In: XVII Seminários em Administração da FEA/USP (SEMEAD). São Paulo, 2014.
- FREITAS JUNIOR, J. C. S.; MAÇADA, A. C. G.; BRINKHUES, R. A.; DOLCI, P. C. **Dimensões de Big Data e o Processo Decisório: Estudos de Casos Múltiplos no Varejo**. In: V Encontro de Administração da Informação, 2015, Brasília. EnADI 2015. Rio de Janeiro: ANPAD, 2015. v. 1. p. 1-17.
- GAGGIOLI, A. CyberSightings. **CyberPsychology, Behavior & Social Networking**. 16, 11, 850, Nov. 2013.
- GARTNER IT GLOSSARY. Disponível em: <<<http://www.gartner.com/it-glossary/big-data/>>> Acesso em 15 nov. 2012.
- GHOCHANI, M; et al. Simulation of customer behavior using artificial neural network techniques. *International Journal of Information, Business & Management*. 5, 2, 59-68, May 2013. ISSN: 20769202.
- HANSEN, M. T. et al. The search-transfer problem: The Role of weak ties in sharing knowledge across organization subunits. **Administrative Science Quarterly**, n. 44, p. 82 – 111, 1999.
- JOHNSON, JE. BIG DATA + BIG ANALYTICS = BIG OPPORTUNITY. **Financial Executive**. 28, 6, 50-53, Jul. 2012.
- KABIR, N.; CARAYANNIS, E. Big Data, Tacit Knowledge and Organizational Competitiveness. **Proceedings Of The International Conference On Intellectual Capital, Knowledge Management & Organizational Learning**, 220-227, 2013.

KIM, G-H; TRIMI, S.A.;JI-HYONG, C. Big Data Applications in the Government Sector. **communications of the ACM**, vol. 57, no. 3, 2014.

KOSCIEJEW, M. The Era of Big Data. *Felicitier*. 59, 4, 52-55, Aug. 2013.

KWON, O. et al. Data quality management, data usage experience and acquisition intention of big data analytics. **International Journal of Information Management**. 2014.

MAHRT, M.; SCHARKOW, M. The Value of Big Data in Digital Media Research. **Journal of Broadcasting & Electronic Media**, 57(1), 20-33, 2013.

MANYIKA, J.; CHUI, M.; BROWN, B.; BUGHIN, J.; DOBBS, R.; ROXBURGH, C.; BYERS, A. H. **Big Data: The Next Frontier for Innovation, Competition, and Productivity**. McKinsey Global Institute, May 2011. Disponível em [http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation). Acesso em 08 de novembro de 2014.

LAMONT, J. Big data has big implications for knowledge management. **KM World**. 21, 4, 8-11, Abr. 2012.

LARA-NAVARRA, P; SERRADELL-LÓPEZ, E; MANIEGA-LEGARDA, D. EVOLUCIÓN DE LOS REPOSITARIOS DOCUMENTALES. EL CASO SOCIALNET. (Spanish). : Evolution of documents repositories. The SocialNet case. (English). **El Profesional de la Información**. 22, 5, 432-439, Set. 2013.

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. Big Data: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana. Rio de Janeiro: Elsevier, 2013.

NONAKA, I.; TAKEUCHI, H. Criação de conhecimento na empresa. Rio de Janeiro: Elsevier, 1997. 13º Reimpressão.

ORACLE. **From Overload to Impact: An Industry Scorecard on Big Data Business Challenges**, 2012. Disponível em < <http://www.oracle.com/us/industries/oracle-industries-scorecard-1692968.pdf>>. Acesso em 12 de julho de 2015.

POLANYI, M. **The tacit dimension**. Garden City: Anchor Books, 1967

RAJESH, k,V,N. Big Data Analytics: Applications and Benefits. **The IUP Journal of Information Technology**, Vol. IX, No. 4, 2013

SAMBAMURTHY, V.; SUBRAMANI, M. Special issue on information technologies and knowledge management. **MIS Quarterly**, v.29, pp. 193-195, 2005.

SPLENDIANI, A; et al. Knowledge sharing and collaboration in translational research, and the DC-THERA Directory. **Briefings in Bioinformatics**. 12, 6, 562-575, Nov. 2011.

SUKUMAR, SR; FERRELL, RK. 'Big Data' collaboration: Exploring, recording and sharing enterprise knowledge. **Information Services & Use**. 33, 3, 257-270, Jul. 2013.

WEBSTER, J.; WATSOM, R.T. Analyzing the past to prepare for the future: Writing a literature review. **MIS Quarterly**, v.26, n.2, p.13-23, 2002.

WOLFSWINKEL, J. F.; FURTMUELLER, E.; WILDEROM, C.P. M. Using grounded theory as a method for rigorously reviewing literature. **European Journal of Information Systems**. 22, 45-55. 2013

ZHANG, D. Granularities and Inconsistencies in Big Data Analysis. **International Journal of Software Engineering & Knowledge Engineering**. 23, 6, 887-893, Ago. 2013.

ZIKOPOULOS, P.C. et al. Understanding Big Data: Analytics for Enterprise-Class Hadoop and Streaming Data. **McGraw-Hill**, New York, 2012.