DOI: 10.5748/9788599693100-11CONTECSI/PS-694

BIG DATA: PUBLICATION EVOLUTION AND RESEARCH OPPORTUNITIES

Simone S. Luvizan (Fundação Getulio Vargas, São Paulo, Brasil) - sluvizan@hotmail.com
Fernando S. Meirelles (Fundação Getulio Vargas, São Paulo, Brasil) - fernando.meirelles@fgv.br
Eduardo H. Diniz (Fundação Getulio Vargas, São Paulo, Brasil) - eduardo.diniz@fgv.br

As supply and demand of data and tools to manage them increase, creating the Big Data (BD) paradigm, the academic engagement in this discussion becomes key. Based in bibliometric data and Baskerville and Myers (2009) lens, this paper identify an academic and non-academic BD concernment wave raised in 2011 and continuously growing since then. Data about publications in portuguese were not conclusive, letting doubt about Brazilian academy enthusiasm for BD or his adoption of the expression BD, question to be explored by future studies. To a deeper investigation of BD in IS field, some ICT publications were analyzed to capture the level of BD participation in their agenda and to present their contents briefing and word cloud representation. Finally, we propose 5 topics for future researches, highlighting the importance of being attentive to emergent phenomena's to build a solid IS research tradition that could effectively collaborate with the progress of theory and practice.

Keywords: big data; analytics; IS research agenda; IS fashion waves.

BIG DATA: EVOLUÇÃO DAS PUBLICAÇÕES E OPORTUNIDADES DE PESQUISA

O aumento da oferta e demanda de dados e de ferramentas para manejá-los, constituindo o paradigma do Big Data (BD), torna cada vez mais relevante seu debate no meio acadêmico. A partir de dados bibliométricos e das lentes de Baskerville e Myers (2009), este artigo identifica a onda de interesse acadêmico e não-acadêmico sobre BD que se projeta em 2011 e cresce continuamente desde então. Dados das publicações em português não são conclusivos e põem em dúvida o entusiasmo da academia brasileira ou sua adesão ao termo BD, o que poderia ser estudado por outros trabalhos. Para explorar o campo de IS, alguns periódicos foram analisados para captar o nível de participação do tema e relatar brevemente seus conteúdos e a nuvem de palavras que os representa. Por fim, apresentamos 5 temáticas para futuras pesquisas, destacando a importância de manter-nos atentos aos fenômenos emergentes para construir uma tradição de pesquisa de IS que possa de fato contribuir com a ciência e com a prática.

Palavras-chave: big data; inteligência analítica; agenda de pesquisa SI; ondas de interesse SI.

INTRODUÇÃO

Novas tecnologias têm surgido nos últimos anos para endereçar as limitações técnicas das ferramentas convencionais em lidar com as demandas de processamento cada vez mais sofisticadas, tempos de resposta cada vez menores e crescentes volumes de dados (LETOUZÉ, 2012; GOLDMAN et. al, 2012). A chamada WEB 2.0 com suas redes sociais e ferramentas de colaboração, a "inteligência" embarcada nos objetos para captar, armazenar e trocar informações possibilitada pela "Internet das coisas" e as plataformas móveis, são apenas alguns exemplos de novos geradores de dados. Tais ferramentas encontram-se em franco desenvolvimento, porém o patamar alcançado até aqui já permitiu a criação de diversas aplicações capazes de processar bases de dados gigantescas em tempo real ou muito próximo disso (GOLDMAN et. al, 2012). O termo Big Data (BD) surgiu para denominar o fenômeno destes grandes volumes de informações (LETOUZÉ, 2012; GOLDMAN et. al, 2012).

O fenômeno BD tem sido amplamente discutido no meio executivo de IS. Projeções confirmam o interesse dos profissionais pelo tema e sua intensão de adotá-lo em suas operações. Segundo pesquisa global do Gartner, 42% dos líderes de TIC estimam investir em projetos de BD dentro de um ano (INFORMATIONWEEK, 2013). Esta expectativa também está presente nos governos e outras organizações políticas e sociais. Gestão de cidades baseadas em colaboração com cidadãos, apoio ao serviço social, aplicações para órgãos de inteligência da segurança pública e até ferramentas de monitoramento e estratégia eleitoral são alguns exemplos das possibilidades no cenário público (LETOUZÉ, 2012; PSFK, 2011; SMOLAN e ERWITT, 2012). Iniciativas não-governamentais também têm alertado para o potencial do BD para o desenvolvimento (BD4D), visando estimular o debate sobre suas possibilidades de contribuir para solução de grandes problemas da humanidade (LETOUZÉ, 2012; PSFK, 2011; SMOLAN e ERWITT, 2012).

Pesquisadores, no entanto, alertam para a escassez de estudos refletindo as implicações da utilização destas ferramentas em questões organizacionais, culturais e sociais mais abrangentes (BOYD e CRAWFORD, 2011). Especialmente no Brasil, o volume de pesquisas acadêmicas em BD ainda é pequeno. Para contribuir com esta discussão, o objetivo deste artigo é identificar a onda de interesse pelo BD entre acadêmicos e não-acadêmicos, refletindo especialmente sobre a participação e abordagens no campo de IS. Partindo do conceito de Baskerville e Myers (2009) da moda nos estudos de IS, este artigo utiliza-se de dados bibliométricos para identificar a curva de interesse acadêmico e não-acadêmico sobre BD na base de dados analisada, fazendo especiais considerações sobre o cenário brasileiro. Na sequência, alguns periódicos da área de tecnologia são analisados para identificar o nível de participação do tema em suas publicações e um breve relato de seus conteúdos é apresentado. Por fim, coloca-se as conclusões dos resultados identificados juntamente com a proposta de 5 grupos temáticos que poderiam inspirar novos trabalhos.

BIG DATA – CONCEITOS E UM BREVE HISTÓRICO

Pesquisas sobre grandes volumes de dados não são uma novidade. Alguns autores sugerem que elas iniciaram ainda na década de 70, investigando métodos de processamento de dados e chegaram aos anos 90 estudando, por exemplo, a modelagem e desenvolvimento de software para grandes volumes de dados (PARK e LEYDESDORFF, 2013). Os anos 2000, no entanto, são marcados por um salto não apenas nas possibilidades técnicas de processamento, armazenagem e transmissão de dados, mas também pela explosão de fenômenos de geração de dados que nos levaram a volumes sem precedentes

na história da humanidade. Neste contexto, o termo BD não indica um fenômeno composto por elementos totalmente novos, mas um conjunto de questões, novas e clássicas, que combinadas em novo cenário tecnológico, social e econômico, deram origem a um novo paradigma.

A definição de BD adotada neste trabalho foi a encontrada com mais frequência na literatura acadêmica e não-acadêmica, sendo também a que nos parece mais coerente. Ela propõe que BD é o fenômeno do processamento de grandes volumes de dados, com os quais as ferramentas tradicionais não são capazes de lidar na velocidade requerida (GOLDMAN et al, 2012). Não é, portanto, um volume específico que classifica o fenômeno, que também é marcado por outras características, como a complexidade e velocidade de processamento necessárias (DEMCHENKO et al, 2013; PARK e LEYDESDORFF, 2013). Logo, a definição de "big" deve ser analisada no contexto individualizado, já que o volume considerado grande em uma determinada situação pode não ser considerado grande em outra. Esta classificação também deve variar ao longo do tempo para a mesma demanda, devido aos rápidos avanços da capacidade das ferramentas envolvidas, de forma que o grande de hoje pode ser o médio de amanhã (PARK e LEYDESDORFF, 2013). Para facilitar a classificação alguns autores sugerem que estamos diante de um fenômeno de BD quando o tamanho dos dados faz parte do problema de pesquisa (PARK e LEYDESDORFF, 2013).

Os muitos desafios enfrentados pelo BD foram inicialmente sumarizados em 3 V's: Volume (basicamente tamanho e quantidade de dados), Velocidade (dinâmica de crescimento e processamento dos dados) e Variedade (diversidade de origens, formas e formatos dos dados) (DEMCHENKO et al., 2013). Posteriormente, foram agregados os elementos Valor (significados que podem ser atribuídos aos dados, valor agregado oferecido por tais significados) e a Veracidade (autenticidade, reputação da origem, confiabilidade dos dados), constituindo-se nos 5 V`s do BD (DEMCHENKO et al. 2013), conforme figura 1.

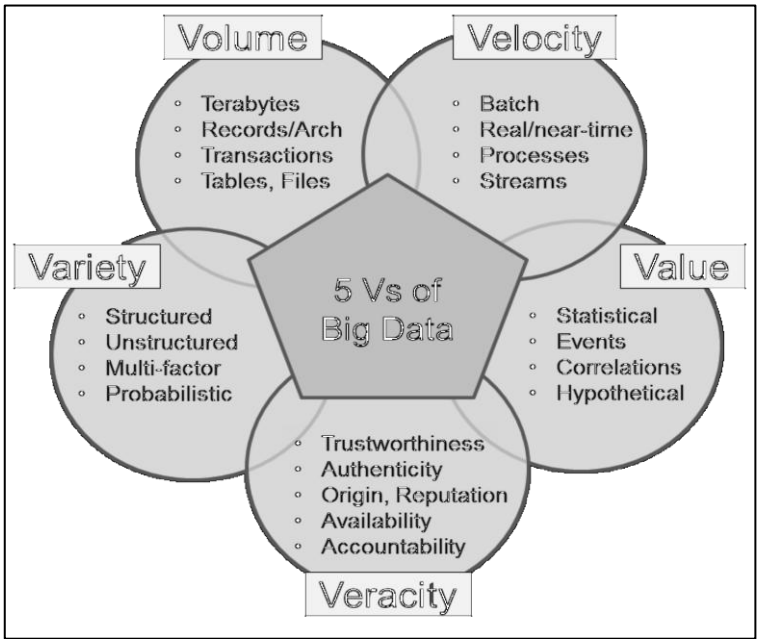


Figura 1: Os 5 V's do Big Data
Fonte: DEMCHENKO et al., 2013

Embora as aplicações e objetivos específicos das ferramentas de BD sejam muito variadas, pode-se dizer que seus usuários compartilhem de uma tríade de expectativas, expressa pelo acrônimo MAD - Magnetism, Agility, Depth - (COHEN et al., 2009). O Magnetismo é a capacidade de atrair dados sobre um determinado tema de diferentes fontes, sejam eles de qualquer formato, estrutura ou origem. A Agilidade indica a grande capacidade de adaptação do sistema à evolução dos dados. A Profundidade se refere ao nível de detalhe possibilitado pelas análises produzidas a partir do grande conjunto de dados e à complexidade do processamento realizado, podendo envolver conceitos estatísticos sofisticados e aprendizado de máquina (COHEN et al., 2009).

A importância e amplitude deste tema vêm atraindo interesses entre praticantes e acadêmicos em diversas áreas. Institutos de pesquisa renomados como Mackinsey e Gartner, além de organismos internacionais como a ONU, já incorporaram o tema em suas agendas de pesquisa e atuação há alguns anos (GARTNER, 2013; MANYIKA, 2011; PSFK,2011). Na iniciativa privada também é crescente o interesse pelo tema, seja entre as empresas que apostam no potencial de valor de projetos de BD para seus negócios, ou entre aquelas que atuam ou planejam atuar oferecendo produtos/serviços nesta área para abocanhar os investimentos esperados das primeiras. As expectativas também são grandes no setor público, onde as aplicações nas diversas esferas de atuação prometem elevar a gestão pública a níveis sem precedentes de eficiência, controle e transparência (LETOUZÉ, 2012; PSFK, 2011; SMOLAN e ERWITT, 2012). Há também grandes expectativas sobre o impacto nas relações políticas, seja pelo uso de ferramentas de BD pelos políticos ou pelo novo modelo de organização e engajamento da sociedade através das redes sociais.

Apesar de já participar das discussões precursoras desta temática, a academia ainda tem muito a avançar nos estudos sobre este novo paradigma. Diversos pesquisadores alertam para a falta de estudos acadêmicos que abordem BD sob perspectivas de análise mais amplas e integrativas (BORGMAN, 2012; BOYD e CRAWFORD, 2011; DEMCHENKO, et al., 2013). Neste sentido, conhecer o nível do interesse acadêmico sobre a questão e os assuntos sobre os quais tais interesses repousam pode ser de grande valia para a evolução do conhecimento nesta área. E este é o propósito deste artigo.

TEORIAS E MÉTODOS

Nesta sessão, faremos uma descrição dos procedimentos metodológicos adotados e a base conceitual que os sustentam para reforçar a clareza e a transparência deste trabalho. Isto é recomendado por diversos autores, que alertam sobre a falta de atenção dada por muitas pesquisas na descrição dos procedimentos e fundamentação dos métodos utilizados, colocando em questão seu rigor científico, base para legitimidade de resultados e construção de teoria em qualquer campo de estudos (BAGOZZI, 2011; BURRELL e TOYAMA, 2009; RINGLE et. al., 2012; STRAUB, 1989; WHETTEN, 1989). Primeiramente, apresentamos a base conceitual e uma visão geral dos procedimentos da pesquisa, seguida da descrição das principais etapas do processo.

Base Conceitual

A abordagem conceitual deste trabalho será baseada na proposta de "IS fashion waves" de Baskerville e Meyers (2009) construída a partir da teoria da Moda na Gestão (Management Fashion) de Abrahamson (1996). Inspirado nas teorias neo-institucional e da inovação, Abrahamson (1996) propõe que uma comunidade formadora influencia os gestores seguidores, conduzindo-os a uma crença coletiva e transitória sobre os benefícios de determinadas técnicas para o avanço da gestão. Com base neste conceito, Baskerville e Meyers (2009) definem o IS Fashion como: "...uma crença relativamente transitória e

coletiva na pesquisa e prática de IS, disseminada pelos formadores da moda, que uma técnica ou tecnologia leva à inovação em IS." (BASKERVILLE e MEYERS, 2009, p. 649).

Formadores e seguidores da moda interagem num ambiente, sob a influência de normas, padrões de racionalidade e forças sócio-psicológicas e tecno-econômicas, que participam da dinâmica da "moda" e sustentam seu fluxo cíclico (BASKERVILLE e MEYERS, 2009). Diversos críticos afirmam que a academia não deve se influenciar pela "moda", mantendo as agendas de pesquisa imunes aos apelos do mercado ou dos praticantes, mais vulneráveis à ela (BASKERVILLE e MEYERS, 2009). Assim como Abrahamson (1996), Baskerville e Meyers (2009) também se colocam em uma posição neutra, já que seus estudos não se propõe a discutir o mérito da "moda", mas sim utilizar-se deste conceito para identificar ondas de interesse e verificar a conexão entre as agendas de pesquisa da academia e os focos de atenção da prática, examinando as influências mútuas entre elas. Analogamente, neste trabalho não pretendemos criticar ou apoiar a adoção de modismos nos estudos acadêmicos, mas sim utilizar-nos desta base conceitual para estabelecer uma medida que indique o nível de interesse manifestado na literatura acadêmica e não-acadêmica sobre BD e, consequentemente, discutir o engajamento acadêmico atual nesta questão. Vale também lembrar que, por se tratar de um tema emergente e não de um ciclo concluído, não será possível analisar a dinâmica do processo de formação e solvência da onda, mas apenas identificar o início de sua curva, informação suficiente para os propósitos deste artigo.

Procedimentos de Pesquisa

Visão Geral. Para identificar o interesse temático na literatura acadêmica e praticante, adotamos uma estratégia semelhante àquela utilizada por Baskerville e Meyers (2009), baseada na pesquisa bibliográfica para dimensionar a presença do tema nas publicações do campo e calcular o índice de participação deste ao longo do tempo. A progressão deste índice anual define a curva de interesse, evidenciando ondas concentradas em períodos específicos, as chamadas "fashion waves" (BASKERVILLE e MEYERS, 2009). Os dados de três periódicos utilizados para o cálculo desta curva foram estratificados, para uma análise de seus conteúdos realizada em duas etapas. Na primeira, os dados (título, resumo e termos do assunto) foram utilizados para geração de nuvem de palavras, que evidenciou a predominância de termos, indicando questões mencionadas com mais frequência. Na segunda, a leitura completa deste extrato que, juntamente com a análise das nuvens, deu origem a um breve sumário do conteúdo geral das publicações analisadas. Exercícios alternativos de busca por outros termos de pesquisa foram realizados para explorar o comportamento da produção acadêmica no Brasil, uma vez que a busca inicial não identificou sua onda de interesse por BD. Seus resultados, embora não permitam comparações com a onda de interesse internacional, trouxeram elementos que permitem levantar questões sobre a atuação acadêmica brasileira neste assunto.

Fonte de dados. Os dados foram extraídos da base de dados Business Source Complete através da interface EBSCOhost, um sistema de busca e referência através do qual pode-se localizar e acessar dados bibliográficos e textos completos em diversas bases de dados. A base selecionada abriga conteúdo acadêmico, comercial e profissional produzido desde 1886, sendo uma alternativa interessante para pesquisas em diversas áreas ligadas à gestão de negócios, incluindo MIS (EBSCO, 2013). Os acessos foram feitos através de usuário conectado à rede FGV, utilizando-se portanto da assinatura desta junto ao EBSCO.

Coleta de Dados. Os dados foram coletados entre os dias 12/11/2013 e 24/11/2013, selecionando períodos anuais, de 2000 a 2013. Em resumo, as buscas foram sem termo de busca na coleta geral e com o termo "Big Data" (entre aspas para captar toda a expressão) em texto completo na coleta dos trabalhos que mencionam BD. A partir deste resultado inicial, selecionamos os três periódicos acadêmicos da área de tecnologia com maior quantidade de artigos mencionando "Big Data" e que fossem classificados no ISI Web of Knowledge - 2012 JCR Social Science Edition. Para estes periódicos foram captadas as quantidades de artigos mencionando BD e também a quantidade total de artigos publicados anualmente. As quantidades identificadas a cada busca foram registradas em planilha eletrônica (MS-Excel), constituindo-se no banco de dados da pesquisa. Para a análise de conteúdo dos periódicos selecionados, os dados de resumo, título e termos do assunto de cada artigo contendo a expressão "Big Data" foi exportada, gerando-se um arquivo de texto consolidado por periódico. Estes dados foram utilizados para a geração da nuvem de palavras e para a análise de conteúdo. Durante este processo, identificamos resultados que faziam parte do conteúdo de publicações acadêmicas, mas não correspondiam à produção acadêmica propriamente dita, como notas do editor ou chamadas de trabalhos. Estes itens foram excluídos das quantidades iniciais, alterando o ranking de periódicos acadêmicos com mais trabalhos mencionando BD e, consequentemente, alterando a seleção dos 3 periódicos para análise de conteúdo. Os periódicos apresentados neste artigo correspondem aos 3 primeiros do ranking final.

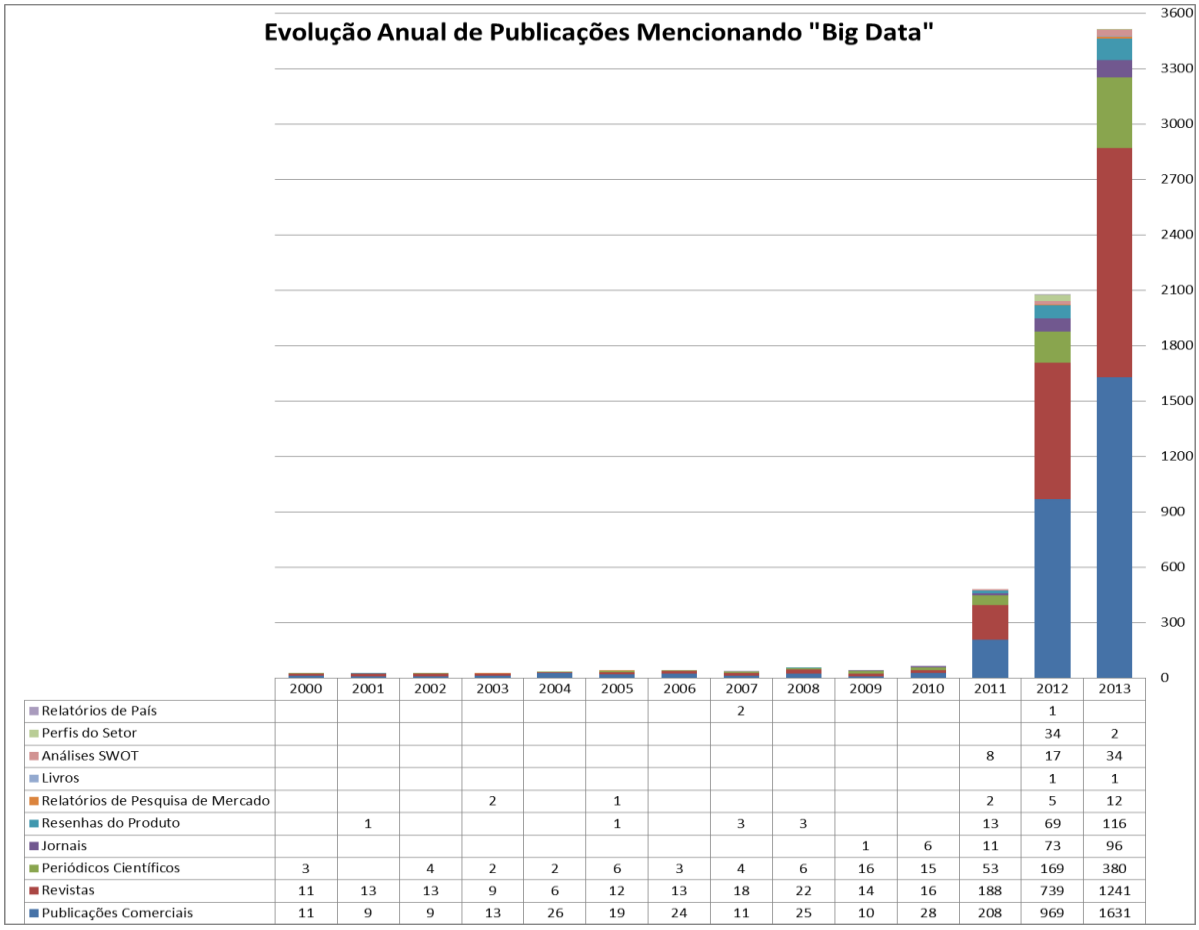
Tratamento e análise dos dados. Para a análise das quantidades registradas nas buscas foram gerados quadros e gráficos apresentados abaixo com seus respectivos esclarecimentos. Para a análise de conteúdo, os arquivos de texto com título, resumo e termos do assunto dos periódicos selecionados foram tratados para exclusão de outras informações que tenham sido agregadas durante a exportação dos dados (marcações, endereços, etc), da identificação dos campos e de palavras sem representatividade para o assunto do artigo, mas que apareciam com frequência, como *article, editor, author, discuss, present*, etc. Este texto final foi utilizado para a geração da nuvem de palavras de cada periódico, através da ferramenta on-line Tagxedo (www.tagxedo.com). A análise dos resultados foi baseada nos dados das buscas captados em planilha, textos extraídos, gráficos e índices gerados, conforme será apresentado a seguir. Embora os dados tenham sido coletados até o ano 2013, nas comparações entre períodos iniciais e finais, considerou-se o ano 2012, já que o ano de 2013 ainda não estava concluído e não poderia ser comparado com outros anos completos.

RESULTADOS

Diferentemente da experiência de Baskerville e Myers (2009), que observaram temáticas cujos ciclos de ascendência e declínio já haviam sido concluídos, este trabalho focaliza um fenômeno emergente. Aqui não se pretende analisar o ciclo de vida completo do interesse pelo tema ou sua comparação com outros, uma vez que sua dinâmica ainda está em curso. Esta análise se deterá à emergência e crescimento do interesse, buscando discutir suas implicações no meio acadêmico e praticante.

O gráfico da evolução anual das publicações mencionando "Big Data" desde o ano 2000 (figura 2), evidencia que, embora estivesse presente anteriormente e apresentasse um aumento gradativo, houve um grande salto neste número em 2011, dando início a uma curva crescente acentuada. De 65 itens listados em 2010, passamos a 483 em 2011, 2077 em 2012 e 3513 em 2013, indicando uma tendência de crescimento e alcance, já que há

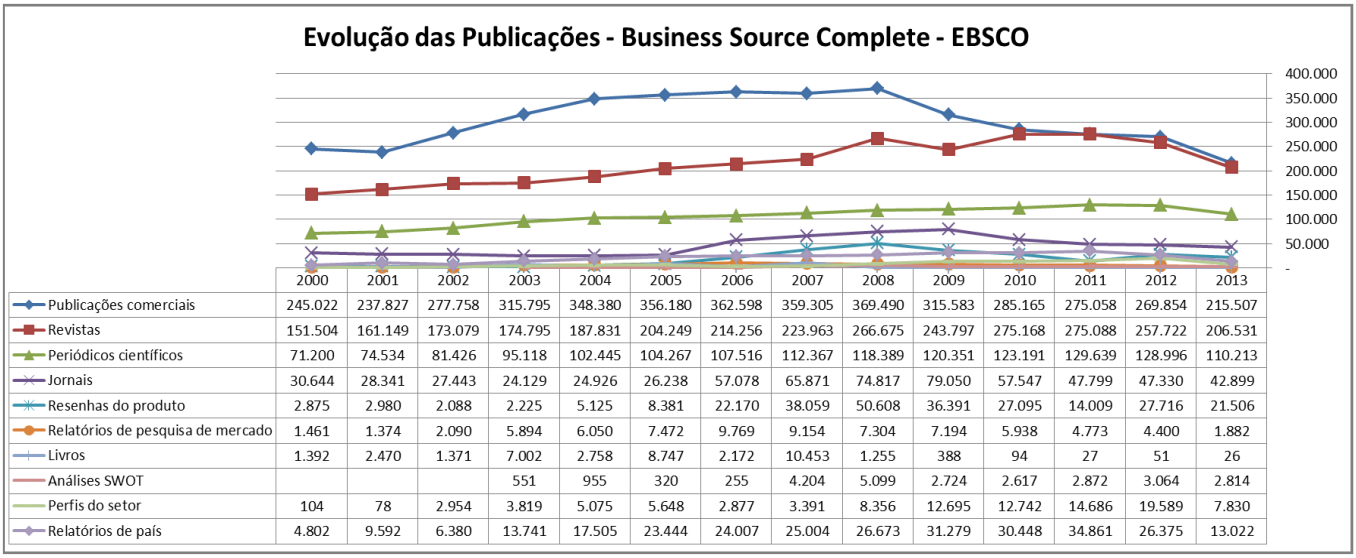
aumento de quantidade e disseminação entre as publicações de todas as categorias (acadêmicas e não-acadêmicas).



	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Relatórios de País								2					1	
Perfis do Setor													34	2
Análises SWOT												8	17	34
Livros													1	1
Relatórios de Pesquisa de Mercado				2		1						2	5	12
Resenhas do Produto		1				1		3	3			13	69	116
Jornais										1	6	11	73	96
Periódicos Científicos	3		4	2	2	6	3	4	6	16	15	53	169	380
Revistas	11	13	13	9	6	12	13	18	22	14	16	188	739	1241
Publicações Comerciais	11	9	9	13	26	19	24	11	25	10	28	208	969	1631

Figura 2: Gráfico da Evolução Anual de Publicações Mencionando “Big Data”
Fonte: Elaborado pelos Autores
Nota: Ano de 2013 incompleto, pois os dados foram coletados em novembro/2013

Outras questões, como a variação no volume total de publicações no período, por exemplo, também podem interferir na elevação do volume de publicações sobre um tema específico. De fato, a figura 3 demonstra uma oscilação não apenas no volume total de publicações da base, mas também uma variação interessante na sua dispersão entre as categorias. Destaca-se a queda brusca na quantidade de livros, além da redução nas publicações comerciais. Em contrapartida, há um aumento em outras categorias importantes, como revistas e jornais. Em periódicos científicos, observa-se crescimento gradativo e consistente, elevando-os de 71.200 em 2000 ao patamar de 128.996 em 2012, o que significa um crescimento de 81%. Vale lembrar que este retrato reflete a situação do universo de publicações considerada nesta base de dados, não podendo ser tomada como a totalidade da produção científica da área. Um estudo mais abrangente, incluindo outras fontes de dados poderia verificar se esta variação corresponde ao comportamento geral do campo, explorando as implicações deste quadro.



	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Publicações comerciais	245.022	237.827	277.758	315.795	348.380	356.180	362.598	359.305	369.490	315.583	285.165	275.058	269.854	215.507
Revistas	151.504	161.149	173.079	174.795	187.831	204.249	214.256	223.963	266.675	243.797	275.168	275.088	257.722	206.531
Periódicos científicos	71.200	74.534	81.426	95.118	102.445	104.267	107.516	112.367	118.389	120.351	123.191	129.639	128.996	110.213
Jornais	30.644	28.341	27.443	24.129	24.926	26.238	57.078	65.871	74.817	79.050	57.547	47.799	47.330	42.899
Resenhas do produto	2.875	2.980	2.088	2.225	5.125	8.381	22.170	38.059	50.608	36.391	27.095	14.009	27.716	21.506
Relatórios de pesquisa de mercado	1.461	1.374	2.090	5.894	6.050	7.472	9.769	9.154	7.304	7.194	5.938	4.773	4.400	1.882
Livros	1.392	2.470	1.371	7.002	2.758	8.747	2.172	10.453	1.255	388	94	27	51	26
Análises SWOT				551	955	320	255	4.204	5.099	2.724	2.617	2.872	3.064	2.814
Perfis do setor	104	78	2.954	3.819	5.075	5.648	2.877	3.391	8.356	12.695	12.742	14.686	19.589	7.830
Relatórios de país	4.802	9.592	6.380	13.741	17.505	23.444	24.007	25.004	26.673	31.279	30.448	34.861	26.375	13.022

Figura 3: Gráfico da Evolução Anual das Publicações Em Geral no Business Source Complete - EBSCO
Fonte: Elaborado pelos Autores
Nota: Ano de 2013 incompleto, pois os dados foram coletados em novembro/2013

Para uma análise mais contextualizada do nível de interesse por um determinado assunto, é preciso relativizar sua presença no universo considerado (BASKERVILLE e MYERS, 2009; ABRAHAMSON e FAIRCHILD, 1999). Para isso adotamos o mesmo procedimento de Baskerville e Myers (2009), que consiste em calcular um índice de participação simplesmente dividindo a quantidade dos estudos que mencionam BD pela quantidade total de trabalhos publicados. A evolução anual deste índice na figura 4 revela a curva de interesse pelo tema no período em cada categoria. Este quadro confirma o início de um interesse mais significativo e generalizado a partir de 2011 e que segue crescente desde então, caracterizando o surgimento do que estamos denominando neste trabalho de onda de interesse.

Por ser um processo em andamento, não é possível analisá-lo de forma ampla, já que muitas coisas podem acontecer ao longo de seu ciclo de vida. Não se pode precisar, por exemplo, se realmente trata-se de uma onda de interesse pontual, de rápida ascensão e declínio ou uma questão que permanecerá representativa durante um longo período. De qualquer forma, estes dados são suficientes para a questão desta pesquisa cujo objetivo não é fazer uma análise retrospectiva de um fenômeno passado, mas identificar o status atual de um fenômeno em curso.

A proximidade entre o início da onda das categorias acadêmica (periódicos científicos) e não-acadêmicas (todos os demais) sugere que a academia está atenta às temáticas atuais, como também foi observado no estudo de Baskerville e Myers (2009). Como indicado naquele trabalho, aqui também nota-se uma pequena antecedência da onda entre os trabalhos não-acadêmicos, sugerindo uma postura seguidora da academia, que pode ser reflexo de maior agilidade na adesão dos não-acadêmicos aos tema emergentes ou da diferença dos ciclos de produção e publicação dos trabalhos acadêmicos, mais longos do que os não-acadêmicos.

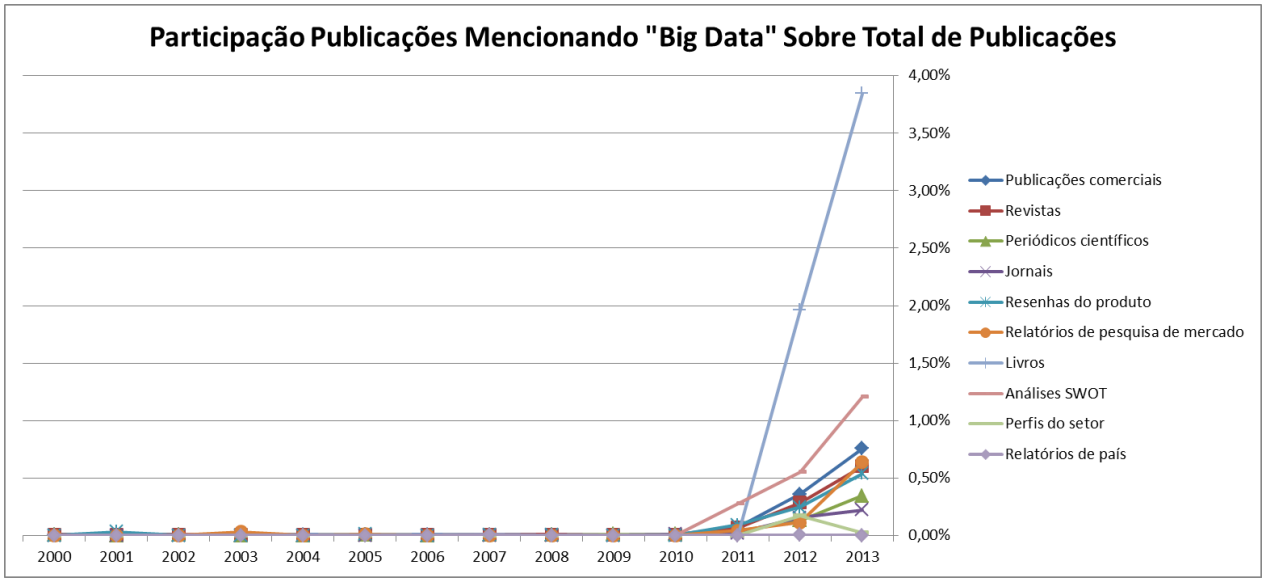
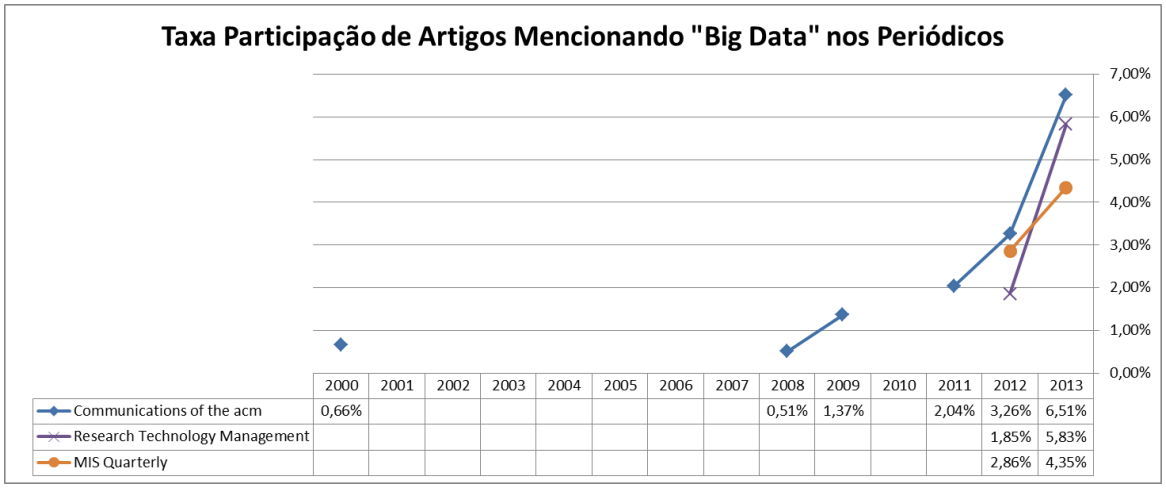


Figura 4: Gráfico da Evolução Anual da Participação das Publicações Mencionando "Big Data" Sobre Total
Fonte: Elaborado pelos Autores
Nota: Ano de 2013 incompleto, pois os dados foram coletados em novembro/2013

Por ser uma base geral da área de gestão e negócios, a onda identificada refere-se a este campo como um todo. Para identificar o comportamento nos estudos de IS, faz-se necessário uma análise mais detalhada sobre a produção do campo. Assim, foram selecionados alguns periódicos científicos que atuam no campo de IS para uma análise mais específica de sua abordagem sobre BD. De acordo com os critérios descritos anteriormente, os três periódicos selecionados foram: Communication of the ACM, Research Technology Management e MIS Quarterly.

Como há diferenças no volume de publicações entre os periódicos, suas quantidades absolutas não são comparáveis e também foi necessário calcular o índice de participação dos estudos que mencionam BD em cada periódico sobre a quantidade total de suas publicações. Além de isolar as diferenças de volume entre os periódicos, isto resolve também a distorção que pode haver na análise do próprio periódico ao longo do tempo, já que seu volume de publicação também pode oscilar.

A figura 5 mostrando a evolução do índice de participação nos três periódicos analisados indica que eles também fazem parte da onda de interesse iniciada em 2011 e acompanham seu crescimento consistente nos anos seguintes. Nota-se ainda que o periódico Communication of the ACM já abordava pontualmente o tema antes deste período, corroborando a informação da figura 2 sobre a produção geral, que também indica haver alguma presença acadêmica nesta questão mesmo antes do interesse mais generalizado.



	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Communications of the acm	0,66%								0,51%	1,37%		2,04%	3,26%	6,51%
Research Technology Management													1,85%	5,83%
MIS Quarterly													2,86%	4,35%

Figura 5: Gráfico da Evolução Anual da Participação dos Artigos Mencionando "Big Data" por Periódico
Fonte: Elaborado pelos Autores
Nota: Ano de 2013 incompleto, pois os dados foram coletados em novembro/2013

Por ser um fenômeno emergente, a identificação dos conteúdos abordados também pode trazer contribuições para o avanço do conhecimento sobre o tema. Abaixo o sumário do conteúdo destas publicações produzido a partir dos dados referentes ao título, termos do assunto e resumo dos artigos dos periódicos selecionados.

Communications of the ACM

Com a maior quantidade de artigos selecionados entre os periódicos analisados, esta é a publicação de temática mais variada entre os três. Um primeiro grupo de artigos aborda questões técnicas relacionadas a algorítmos, ferramentas de gestão de dados, linguagens de programação/desenvolvimento, plataformas e infraestrutura em geral. Neste item enquadram-se temas como *cloud-computing, MapReduce, Hadoop, NoSQL, Hazi*, programação modular e inteligência artificial. Um segundo grupo de artigos discute a aplicação destas ferramentas, trazendo assuntos como BI, *data mining, process mining*, *marketing* direcionado pelo comportamento do cliente, *crowdsourcing,* aplicações *mobile*, representação gráfica de grandes volumes de dados e interfaces gestuais. Um terceiro grupo de artigos propõe discussões sobre segurança e privacidade, lembrando que esta temática também aparece em diversos artigos dos grupo anteriores, mesmo que tangencialmente. Por fim, alguns artigos falam sobre educação, destacando os desafios atuais da formação na área de tecnologia.

Figura 6: Nuvem de Palavras dos Artigos Mencionando "Big Data" no *Communications of the ACM*
Fonte: Elaborado pelos Autores

MIS Quarterly

Alguns artigos discutem questões ligadas à temática de BI (Business Intelligence) ou BA (Business Analytics), como a abordagem dos CIOs para melhorar a orientação à informação em suas empresas, a jornada para o desenvolvimento das capacidades de BI na empresa e recomendações para maximizar o valor de suas iniciativas de BA. Outros discutem temas mais abrangentes, como privacidade da informação do cliente e aquisição, entendimento e uso dos dados gerados em tempo real, refletindo sobre as oportunidades e as demandas de novas capacidades para sua implementação.

Figura 7: Nuvem de Palavras dos Artigos Mencionando "Big Data" no *MIS Quaterly*
Fonte: Elaborado pelos Autores

Research Tehnology Management

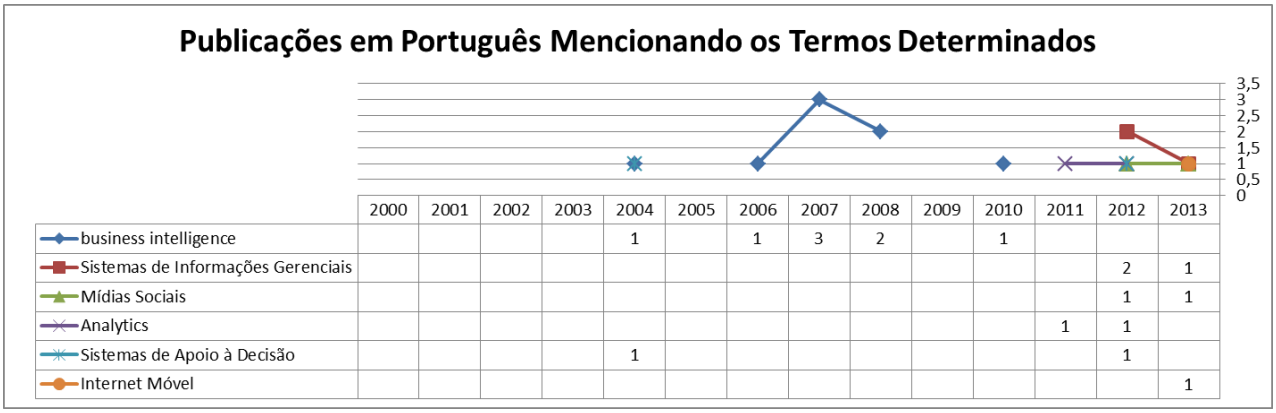
Enquanto alguns trabalhos falam sobre tendências futuras e políticas governamentais (EUA), outros artigos discutem o paradoxo de tecnologias disruptivas como as mídias sociais, a conexão do fenômeno BD com a emergência da WEB 3.0 (Internet das Coisas) e a utilização de BD na Gestão do Conhecimento e da Inovação. Alguns abordam a combinação de estratégias analíticas e intuitivas na gestão inovadora e as oportunidades do mundo VUCA (Volatile, Unstructured, Complex, Ambiguous) em vários contextos e as novas habilidades necessárias para inovação e pesquisa & desenvolvimento neste cenário.

Figura 7: Nuvem de Palavras dos Artigos Mencionando "Big Data" no *Research Technology Management*
Fonte: Elaborado pelos Autores

Os Resultados Sobre os Estudos no Brasil

Na tentativa de identificar o panorama dos estudos de BD no Brasil, as mesmas buscas demonstradas acima foram realizadas, agregando-se o critério idioma português, considerando que grande parte da produção científica brasileira incluída nesta base está em português. Os resultados obtidos nestas buscas foram insignificantes, corroborando o estudo de Park e Leydesdorff (2013) que também não captou presença significativa de estudos sobre BD na área de pesquisa do Brasil, deixando-o de fora da lista dos 20 países com maior nível de participação na rede de relacionamento da área.

Tal situação despertou a dúvida se realmente há uma desatenção da academia brasileira ao tema ou se ela não aderiu ao termo BD. Esta dúvida é reforçada pelo alerta de Meirelles (2013), apontando problemas de taxonomia em Business Intelligence (BI), que na literatura brasileira foi chamado de Business Intelligence, Business Analytics, Analytics, Sistema de Apoio à Decisão, Sistema de Informações Gerenciais, Sistema de Suporte ao Executivo, Inteligência Analítica e Inteligência de Negócios, entre outros. Embora BD e BI sejam fenômenos diferentes, diversos estudos de BD no âmbito das organizações convergem com assuntos relacionados a BI (CHEN et al, 2012). Assim, para tentar identificar a existência de uma onda de interesse similar à observada no cenário internacional, associada a termos alternativos, as buscas foram refeitas utilizando as expressões mencionadas acima nas denominações de BI no Brasil e também para "mídias sociais", "internet das coisas" e "internet móvel", assuntos também frequentes na temática de BD. A figura 9 demonstra o resultado destas seleções, indicando que estes temas apareciam mais dispersos em anos anteriores e parecem um pouco mais concentrados em 2012 e 2013, sugerindo maior presença dos mesmos nas publicações brasileiras. Vale lembrar, no entanto, que este panorama em números absolutos não é suficiente para determinar se há uma onda de interesse e, em caso positivo, qual sua intensidade. Não se pode afirmar também que os estudos identificados abordem de fato os novos desafios discutidos em BD. Estes números apenas nos dão indícios de que a academia brasileira de negócios e IS não tenha aderido com o mesmo entusiasmo à taxonomia BD, embora o termo seja bastante presente nas publicações não-acadêmicas do Brasil. Sugere ainda indícios de que, embora não compartilhe da nomenclatura, haja discussões crescentes sobre a problemática envolvida neste tema. Estudos mais detalhados sobre a produção brasileira poderiam identificar mais precisamente o posicionamento da academia local.



	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
business intelligence					1		1	3	2		1			
Sistemas de Informações Gerenciais													2	1
Mídias Sociais													1	1
Analytics												1	1	
Sistemas de Apoio à Decisão					1								1	
Internet Móvel														1

Figura 9: Gráfico das Publicações em Português Mencionando Termos Alternativos
Fonte: Elaborado pelos Autores
Nota: Ano de 2013 incompleto, pois os dados foram coletados em novembro/2013

CONCLUSÕES

Este artigo analisou a evolução das publicações mencionando BD, colocando-a na perspectiva do total publicado neste período e traçando o que chamou de curva de interesse sobre o tema para a base de dados Business Source Complete, disponível através do EBSCO. A curva de interesse detectada evidenciou um crescimento notável e progressivo sobre o tema a partir de 2011, tanto nas categorias acadêmicas como nas não-acadêmicas. A curva das publicações não-acadêmicas parece projetar-se levemente antes da acadêmica, sugerindo que a academia tenha um comportamento seguidor e não formador deste interesse, como também observado nos estudos de Baskerville e Myers (2009) sobre outras temáticas. Assim como naquela pesquisa, este trabalho não tem elementos para precisar se este deslocamento realmente reflete a relação formador/seguidor destes atores ou é consequência da diferença de seu processo de publicação, que é mais longo nos veículos acadêmicos, podendo ocasionar um lapso temporal no início e no final destes ciclos de interesse concentrado, dando a impressão que a academia demore mais para entrar e para sair deles (BASKERVILLE e MYERS, 2009).

Este mesmo nível de interesse não foi observado ao se verificar estudos em português, levantando a questão sobre sua abordagem na academia brasileira. Os dados desta pesquisa não são suficientes para se fazer afirmações a este respeito, mas sugerem a necessidade de futuros estudos que examinem este cenário para identificar se há desatenção ao tema e, neste caso, quais as causas deste desinteresse, no mínimo intrigante num país gerador de dados como o Brasil. Tal estudo poderia também investigar a hipótese deste resultado ser reflexo da não adesão da academia brasileira ao termo "Big Data", explorando também suas causas e a existência de uma taxonomia alternativa. Vale acrescentar que, se a adoção de uma taxonomia diferenciada for confirmada, deve-se refletir sobre seus impactos no campo, já que ela pode dificultar o diálogo dos estudos locais com os estudos internacionais, não apenas influenciando o resultado das buscas dos pesquisadores brasileiros, mas especialmente, tornando os trabalhos brasileiros "invisíveis" aos pesquisadores internacionais que buscarão suas referências através de termos que não os captarão. Esta discussão pode ser ainda mais pertinente na medida em que, a despeito do crescente interesse da academia sobre as temáticas de TIC no Brasil, este campo de pesquisa brasileiro ainda não se consolidou no cenário mundial e sua participação em publicações internacionais ainda é inexpressiva (POZZEBON e DINIZ, 2012).

Uma análise mais focada sobre três periódicos internacionais da área de tecnologia sugere que esta temática também venha ocupando crescente espaço na agenda de pesquisas da academia de CS e IS. Seus dados demonstram a semelhança entre sua curva de interesse e a curva de interesse geral, evidenciando ainda a presença de trabalhos sobre o tema mesmo antes do interesse mais generalizado. A revisão dos temas abordados nestas publicações indica uma vasta amplitude temática. O MIS Quarterly, único periódico entre os três especializado em IS, aprofunda-se na questão pela perspectiva do BI, embora um de seus artigos revisando a evolução do BI reafirme seu entendimento de que BD e BI sejam fenômenos diferentes, ainda que apresentem uma área de convergência importante (CHEN et al, 2012). Mesmo que já apresente alguns trabalhos expandindo-se para outras questões do BD, o cenário retratado neste periódico sugere a necessidade de se alargar os horizontes das discussões sobre BD no campo de IS. Como uma última contribuição deste artigo para esta empreitada, abaixo apresentamos algumas ideias para futuros estudos.

OPORTUNIDADES PARA FUTUROS ESTUDOS DE BD

Este trabalho soma-se a outros que clamam por maior participação de BD na academia mundial, destacando as lacunas no campo de IS e a ausência de uma atuação consistente

das pesquisas do Brasil. Para colaborar com seu avanço, apresentamos 5 grupos de questões que poderiam inspirar novos trabalhos:

Estudos multidisciplinares, discutindo contextos, aplicações e questões específicas das áreas, sob óticas mais abrangentes. A pesquisa do McKinsey Global Institute demonstra que, mesmo havendo boas perspectivas para a utilização de BD em todos os setores, a percepção do potencial de ganhos em alguns deles é maior de que em outros (MANYIKA et al, 2011). Segundo a pesquisa, isso se dá porque alguns setores já visualizam mais claramente aplicações promissoras e/ou porque em alguns há maiores barreiras para a identificação de oportunidades e/ou implantação das mesmas (MANYIKA et al, 2011). Estudos multidisciplinares, mesclando o conhecimento dos pesquisadores de IS e das áreas de aplicação poderiam acelerar o avanço das implementações para realizar o potencial já percebido, assim como contribuir para dissipar as barreiras que diminuem o potencial dos setores vistos atualmente como poucos promissores. Vale também ressaltar que a multidisciplinaridade proposta não se limita ao trabalho em conjunto com áreas de aplicação, mas também com outras disciplinas que expandam o horizonte das discussões emergentes. Por exemplo, WACHS et al. (2011) discute as implicações dos padrões culturais na eficiência dos sistemas de reconhecimento gestual, questão com a qual pesquisadores mais especializados em disciplinas ligadas à cultura poderiam contribuir. Outro exemplo, os trabalhos conjuntos com especialistas organizacionais, que poderiam contribuir para a compreensão das implicações e oportunidades do uso mais intenso das ferramentas de busca no âmbito dos dados corporativos, cuja adesão ainda é pequena se comparada com o uso intenso que fazemos dos buscadores na Internet (MEIRELLES, 2013). Por fim, estudos que adotem perspectivas sociomateriais também poderiam contribuir, já que "nós temos que considerar como as ferramentas participam da modelagem do mundo conosco enquanto as utilizamos" (BOYD e CRAWFORD, 2012, p. 675). Estudos qualitativos utilizando-se de modelos como o Multilevel Framework, que focaliza as dinâmicas sociais e materiais como fenômenos embricados (POZZEBON e DINIZ, 2012), poderiam ser um bom caminho para esta corrente.

Padrões metodológicos. Diversos autores alertam sobre a importância da participação acadêmica nas discussões mais amplas sobre a aplicação de BD para definir ou, pelo menos, influenciar os padrões metodológicos a serem adotados no desenvolvimento das ferramentas de BD (BOYD e CRAWFORD, 2012; CHOW-WHITE e GREEN, 2013). Além disso, esta atuação acadêmica poderia também contribuir para disseminar a informação sobre o impacto que tais padrões exercem sobre o funcionamento e resultados das ferramentas, conscientizando seus usuários para a possibilidade de vieses implícitos em sua utilização. Um exemplo desta questão se refere ao uso de BD na própria área de pesquisa no âmbito acadêmico ou não. A disponibilidade de volumes gigantescos de dados e de ferramentas para processá-los reacende a discussão sobre o conceito de conhecimento, a relevância das teorias e de métodos de pesquisa baseados em hipóteses, por exemplo (BOYD e CRAWFORD, 2012). Diversos autores alertam para o risco de pesquisas que, baseadas em grandes volumes de informação, produzirão respostas corretas para perguntas erradas revestidas convincentemente (BOYD e CRAWFORD, 2012; CHOW-WHITE e GREEN, 2013).

Ética, segurança e acesso. A evolução das possibilidades tecnológicas viabiliza uma série de transações de captura e acesso a dados inimagináveis há alguns anos. No entanto, o fato de ser possível não significa que seja ético e é necessário discutir os sistemas que estão

direcionando e regulando estas práticas (BOYD e CRAWFORD, 2012). É preciso refletir, por exemplo, sobre as possibilidades da "Internet das Coisas" lembrando que estamos falando da "Internet das Nossas Coisas" e, portanto, é preciso discutir os níveis de controle e de consciência que o dono das coisas deve ter sobre o que elas estão fazendo na Internet. No centro deste debate estão questões como privacidade, direitos individuais versus interesses coletivos, segurança da informação, propriedade e direitos de acesso, sustentabilidade e governança das estruturas, entre outros (BORGMAN, 2012; BOYD e CRAWFORD, 2012; DEMCHENKO, et al., 2013; MANYIKA et al, 2011;). Estas questões permeiam as discussões sobre BD em todas as áreas de aplicação, mas algumas são mais sensíveis ou estão mais expostas a elas. O setor de governo e política, por exemplo, é fortemente implicado por tais assuntos. Primeiramente, por seu papel fundamental neste debate para construção de políticas públicas capazes de alavancar o desenvolvimento deste campo e, ao mesmo tempo, proteger os direitos individuais e coletivos de empresas e cidadãos (BOYD e CRAWFORD, 2012; MANYIKA et al, 2011). Os episódios do vazamento de documentos secretos e do acesso a e-mails de terceiros pelos serviços de inteligência americana ilustram algumas destas questões e a urgência de seu debate. Além disso, grande foco se dá nesta área também devido ao potencial de mudanças que o uso de tais ferramentas pode trazer para a dinâmica da política, a organização e serviços dos governos, os seus controles e mecanismos participatórios (Chen et al, 2012; LETOUZÉ, 2012; MANYIKA et al, 2011; PSFK, 2011; SMOLAN e ERWITT, 2012). O "Big Brother" da gestão de cidades, o papel decisivo das mídias sociais no processo eleitoral de Obama, o fenômeno recente das manifestações públicas organizadas pelas redes sociais que invadiram as ruas no Brasil em 2013, as iniciativas de eGoverno brasileiras prometendo melhores condições de fiscalização para os órgãos públicos e maior nível de transparência para os cidadãos, são apenas alguns exemplos entre os inúmeros existentes. Apesar desta vasta agenda no âmbito dos governos, a discussão sobre ética, segurança e acesso não se restringe a ele e deve ser abraçada por todos os segmentos da sociedade e todos os campos acadêmicos (BOYD e CRAWFORD, 2012; DEMCHENKO et al, 2013; LETOUZÉ, 2012; MANYIKA et al, 2011). Cabe destacar a profusão de temas a serem explorados pelas empresas e estudos organizacionais em vários campos. Alguns autores alertam, por exemplo, para as novas demandas que os sistemas das bases tecnológicas emergentes impõem aos processos de auditoria (VASARHELYI, 2013). Para eles, a nova geração de ferramentas de auditoria deverá trazer seus componentes integrados aos processos de negócio, uma vez que os métodos tradicionais de auditoria não são capazes de tratar estes novos sistemas, caracterizados por elementos como o grande volume e diversidade dos dados, complexidade e velocidade do processamento e estruturas baseadas em nuvem (VASARHELYI, 2013). Este é apenas um exemplo ilustrativo da diversidade de temas a serem explorados, não apenas focando as ferramentas em si, mas os impactos de sua adoção nos processos de segurança e governança organizacional.

Estrutura. Embora possa soar como um tema mais técnico a ser coberto completamente pela agenda dos pesquisadores de outras áreas, como Computer Science, há aqui diversas lacunas a serem exploradas em IS. Por exemplo, em seu estudo sobre eScience, Demchenko et al. (2013) propõe que a construção de uma comunidade científica efetivamente colaborativa, demanda um modelo de infraestrutura que a suporte e que deve ser concebido a partir de um entendimento amplo de todo o ciclo de vida dos dados de eScience (DEMCHENKO et al, 2013). O concepção destes modelos e a definição destes fluxos de informação e processo se darão sob quais influências e pressões? Quem participará delas? De forma mais geral, qual o impacto que a adoção de ferramentas de BD

exercem sobre a gestão da estrutura em todas as áreas? Quais os impactos nos custos da empresa, por exemplo? Quais os impactos dos novos modelos de fornecimento de estrutura ofertados na era do BD, como o Resource-as-a-Service por exemplo, sobre a gestão e o custo da infraestrutura? A temática abordada pelo Communications of the ACM no período analisado demonstra que há um engajamento dos pesquisadores técnicos neste tema. Advogamos que a participação de IS nesta reflexão pode trazer uma outra dimensão de discussões imprescindíveis para a evolução do tema no âmbito acadêmico e com grandes potenciais de contribuições aos praticantes.

Formação do capital humano. É recorrente na literatura acadêmica e não-acadêmica, que uma das alavancas para a realização do potencial de BD é a capacidade dos profissionais e indivíduos para bem utilizá-lo. Diversos autores alertam para a atual falta destes profissionais no mercado e seu iminente agravamento devido ao aumento da adoção pelas empresas sem contrapartida equivalente na formação de profissionais (CHEN et al., 2012; LETOUZÉ, 2012; MANYIKA et al, 2011). Entendemos que esta lacuna possa ser atacada pela academia em 3 frentes de trabalho. A primeira, debruçando-se sobre a identificação do perfil que se deve buscar desenvolver, suas competências e habilidades. Gartner (2013) sugere que o perfil deste profissional pode ser definido como uma combinação de engenheiro e artista, alertando para a pluralidade de conhecimentos necessário e o antagonismo entre eles, sob a ótica dos perfis clássicos. A segunda frente poderia se dedicar à revisão dos conteúdos e modelos acadêmicos utilizados pelas instituições de ensino, assim como à sua implementação. Aqui vale lembrar que não estamos falando apenas de cursos ligados à tecnologia, mas de todos os cursos, observando que os desafios entre eles serão distintos, tanto pelas diferentes demandas de cada um, quanto pela diferente situação atual em que se encontram. Por outro lado, há que se atentar também para a revisão curricular dos níveis anteriores, já que alguns autores defendem a necessidade de se incluir disciplinas da ciências da computação desde a educação fundamental, como acontece com a física, química, biologia, etc (CERF, 2013). Já a terceira frente discutiria a formação das pessoas fora do ambiente educacional. A demanda por estes profissionais já existe e o mercado não poderá aguardar pelo ciclo de formação dos profissionais que recém ingressem em cursos revisitados nas universidades, para atender sua demanda de pessoal (MANYIKA et al, 2011). Por isso, é necessário que as empresas também assumam seu papel formador. Destaca-se também que esta formação não se limita ao nível dos profissionais. À medida que os indivíduos aderem cada vez mais ao uso de ferramentas de BD, mesmo que não o façam em seu exercício profissional e que não estejam se dando conta deste uso, eles também são parte da demanda por formação ou conscientização.

Por fim, estes são apenas alguns exemplos com os quais não pretendemos esgotar as lacunas temáticas a serem percorridas pelos estudos de BD. Nossa intenção é apenas compartilhar informações e ideias que possam inspirar outros pesquisadores da área a se engajarem em esforços de pesquisa consistentes capazes de realizar o potencial de valor deste fenômeno para a ciência e para a prática.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRAHAMSON, E. (1996). Management Fashion. *The Academy of Management Review*, *21* (1), 254-285.

ABRAHAMSON, E & FAIRCHILD, G. (1999). Management Fashion:Lifecycles, Triggers, and Collective Learning Processes. *Administrative Science Quarterly*, *44* (4), 708-740.

BAGOZZI, R. (2011). Measurement and meaning in information systems and organizational research: methodological and philosophical foundations. *MIS Quarterly*, *35* (2), 261-292.

BASKERVILLE, R.L. & MYERS, M.D. (2009). Fashion Waves in IS Research & Practice. *MIS Quarterly*, *33* (4), 647-662.

BORGMAN, C. (2012). The conundrum of sharing research data. Journal of the American Society for *Information Science and Technology*, *63* (6), 1059-1078.

BOYD, D. & CRAWFORD, K. (2012). Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information, Communication & Society*, *15* (5), 662-679.

BURRELL, J. & TOYAMA, K. (2009). What Constitutes a Good ICTD Research? *Information Technologies and International Development*, *5* (3), 82-94.

CERF, V. (2013). Computer Science Education-Revisited. *Communications of the ACM*, *56* (8).

CHEN, H.; CHIANG, R. & STOREY, V. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, *36* (4), 1165-1188.

CHOW-WHITE, P & GREEN, S. (2013). Data Mining Difference in the Age of Big Data: Communication and the Social Shaping of Genome Technologies from 1998 and 2007. *International Journal of Communications*, *7*, 556-583.

COHEN, J., et al. (2009). MAD Skills: New Analysis Practices for Big Data. PVLDB, 2(2).

DEMCHENKO, Y., et al. (2013, maio). Addressing Big Data Issues in Scientific Data Infrastructure. *First International Symposium on Big Data and Data Analytics in Collaboration (BDDAC 2013)*. Part of The 2013 International Conference on Collaboration Technologies and Systems (CTS 2013), San Diego, USA.

EBSCO (2013). Recuperado em 23 novembro 2013 de http://web.ebscohost.com/ehost/search/selectdb?sid=bcc145fb-aef4-47b0-a9de-e9b6a2132d55%40sessionmgr112&vid=1&hid=127

Gartner (2013, outubro). *Gartner Symposium/ITxpo 2013*. Orlando, FL, USA.

GOLDMAN, A., et al. (2012). Apache Hadoop: Conceitos Teóricos e Práticos, Evolução e Novas Possibilidades. XXXI JORNADAS DE ATUALIZAÇÕES EM INFORMÁTICA.

INFORMATIONWEEK (2013). Big Data: 42% Dos Líderes De TI Planejam Investir No Próximo Ano. InformationWeek Brasil, 26 março 2013. Recuperado em 07 abril 2013 de

http://informationweek.itweb.com.br/13454/big-data-42-dos-lideres-de-ti-planejam-investir-no-proximo-anno/

LETOUZÉ, E. (2012). Big Data for Development:Challenges&Opportunities. UN Global Pulse.

MANYIKA, J et al. (2011). Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute.*

MEIRELLES, F. (2013). Pesquisa Annual – Administração de Recursos de Informática. *GVcia – Centro de Tecnologia de Informação Aplicada da FGV-EAESP*, São Paulo, 2013.

PARK, H. & LEYDESDORFF, L. (2013). Decomposing Social And Semantic Networks In Emerging "Big Data" Research. *Journal of Informetrics*, *7* (3), 756-765.

POZZEBON, M. & DINIZ, E. (2012). Theorizing ICT and Society in the Brazilian Context: a Multilevel, Pluralistic and Remixable Framework. *BAR – Brazilian Administration Review*, Rio de Janeiro, *9* (3), 287-307, Jul/Set.

PSFK (2011). Future Of Real-Time Information. *UN Global Pulse*.

RINGLE, C., SARSTEDT, M. & STRAUB, D. (2012). A Critical Look at the Use of PLS-SEM in MIS Quarterly. MIS Quarterly, *36* (1).

SMOLAN, R. & ERWITT, J. (2012). *The Human Face of Big Data* – IPAD APP. California: Against All Odds.

STRAUB, D. (1989). Validating instruments in MIS research. *MIS Quarterly*, *13* (2), 147-169.

VASARHELYI, M. (2013). Formalization of Standards, Automation, Robots, and IT Governance. *Journal of Information Systems*, *27* (1), 1-11.

WACHS, J. et al. (2011). Vision-based hand-gesture applications. *Communications of the ACM*, *54* (2), 60-71.

WHETTEN, D.A. (1989). What Constitutes a Theoretical Contribution? *The Academy of Management Review*, *14*, 490-495.