

BIG DATA E GESTÃO DA INFORMAÇÃO: MODELAGEM DO CONTEXTO DECISIONAL APOIADO PELA SISTEMOGRAFIA

BIG DATA Y GESTIÓN DE LA INFORMACIÓN: MODELAJE EN CONTEXTO DE LA TOMA DE DECISIONES APOYADO POR LA SISTEMOGRAFIA

William Barbosa Vianna*
Moisés Lima Dutra**
Enzo Morosini Frazzon***

RESUMO

Introdução: O estudo justifica-se pela escassez e utilidade de estudos no campo da informação que abordem o fenômeno do *big data* sob a ótica da gestão da informação.

Objetivo: O objetivo do estudo é identificar e representar os elementos gerais do processo de tomada de decisão no contexto do *big data*.

Metodologia: Trata-se de um estudo de caráter exploratório e natureza teórica e dedutiva.

Resultados: Resultou na identificação dos principais elementos envolvidos no processo decisório no ambiente do *big data* e na sua representação sistemográfica.

Conclusões: Foi possível desenvolver uma representação que permitirá ulteriores desenvolvimentos de simulação computacional.

Palavras-chave: *Big Data*. Apoio à decisão. Sistemografia. Modelagem. Gestão da informação.

* Doutor em Engenharia de Produção pela Universidade Federal de Santa Catarina (UFSC). Professor do Departamento de Ciência da Informação da UFSC. E-mail: wwilliam@hotmail.com

** Doutor em Ciências da Computação (Doctorat En Informatique) pelo Université Claude Bernarde Lyon 1, França. Professor do Departamento de Ciência da Informação da Universidade Federal de Santa Catarina (UFSC). E-mail: m.oisés.dutra@ufsc.br

*** Doutorado em Engenharia de Produção pelo Universität Bremen, Alemanha. Professor do Departamento de Engenharia de Produção e Sistemas da Universidade Federal de Santa Catarina (UFSC). E-mail: enzo.frazzon@gmail.com

1 INTRODUÇÃO

Um dos desafios para Gestão da Informação é apoiar de forma estruturada os processos de tomada de decisão no contexto de *Big Data*, considerando que a informação é um dos ativos principais para os processos decisórios ou de gestão.

Segundo estudo de Gantz e Reinsel (2012), a quantidade de informações existente no universo digital crescerá em até 50 vezes entre os anos de 2010 e 2020 e totalizará um crescimento de até 300 vezes se considerado os anos de 2005 a 2020, partindo de 130 exabytes para aproximadamente 40 zetabytes.

Na prática, essa explosão das redes móveis, a computação em nuvem e as novas tecnologias tem dado origem a um número incompreensivelmente grande de informações (RAPPORTEUR, 2010).

Segundo a revista CIO (2012) dados úteis podem vir de qualquer lugar e estar em toda parte em diversos formatos, pois há diversos dispositivos que capturam dados de diferentes fontes, os quais podem proporcionar ganhos de eficiência se bem trabalhados.

Entretanto, segundo as revistas CIO (2012) e *Computerworld* (2012) a maior parte dos líderes não sabe lidar com essa grande variedade e quantidade de informações, e não tem conhecimento dos benefícios que uma análise bem feita desses dados poderia trazer ao seu negócio ainda que, se não forem trabalhados de forma adequada e estratégica, esses dados podem prejudicar a organização e levar a uma “armadilha de mercado” (CIO, 2012; COMPUTERWORLD, 2012).

Davenport (1998) identifica que não é possível negar que decisões baseadas em dados inúteis têm custado bilhões de dólares em produtos encalhados, em aquisições que não funcionam, em investimentos em instalações ou equipamentos que não produzem o que se espera deles.

Verifica-se a necessidade de que os líderes definam o que fazer com tantos dados, o que é um problema que não envolve só mudança

de tecnologia, mas adaptação de processos e, sobretudo, mudanças na análise e gestão dos dados (COMPUTERWORLD, 2012; MERITALK BIG DATA EXCHANGE, 2013).

E ainda segundo a CIO (2012), as organizações que estiverem atentas ao novo contexto de informações e que estão se preparando para a tomada de decisão em tempo real a partir das informações coletadas dentro de um grande volume de dados gerados, terão vantagem competitiva de 20% sobre os seus concorrentes.

Para Le Coadic (2004), a ciência da informação encontra um ponto de inflexão na explosão da informação e implosão do tempo, conjunção de dois fenômenos que dão origem aos elevados fluxos de informação que impactam nos processos de tomada de decisão.

Segundo a Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação, o contínuo desenvolvimento das tecnologias da informação e evolução constante do ambiente Web têm proporcionado novas formas de acessar, recuperar, armazenar e gerir e interagir com a informação e os objetos tradicionais para representar a informação já não são suficientes, os ambientes de interação com a informação têm sido transformados, bem como o comportamento das pessoas. E isso têm exigido novas abordagens para os estudos em Ciência da Informação (ENANCIB, 2014).

Segundo Pinto, Elias e Vianna (2014) as decisões no campo da Ciência da Informação vêm se tornando cada vez mais complexas por envolver um número cada vez maior de variáveis, algumas de difícil controle, outras intangíveis.

Nesse sentido, considera-se que estudos que associem o contexto de informações do “*big data*” com a tomada de decisão são úteis e válidos para Gestão da Informação e conseqüentemente para o desenvolvimento da ciência da informação.

Para tanto, pergunta-se: **que elementos são relevantes para organizar a informação e apoiar a tomada de decisão no contexto do *big data*?**

2 OBJETIVO E METODOLOGIA

O objetivo desse estudo é identificar e representar os elementos gerais do processo de tomada de decisão no contexto do *big data*. Trata-se de um estudo de caráter exploratório e natureza teórica e dedutiva, pois pretende investigar de forma preliminar o fenômeno do *big data* associado a processos decisórios, tendo em vista a gestão da informação.

Triviños (1987) descreve o estudo de tipo exploratório como aquele que permite ao investigador aumentar sua experiência em torno de determinado problema. O pesquisador parte de uma hipótese e aprofunda seu estudo nos limites de uma realidade específica, buscando antecedentes, maior conhecimentos para, em seguida, planejar uma pesquisa descritiva ou de tipo experimental.

Sampieri, Collado e Lucio (1994), justificam o uso do estudo exploratório quando o objetivo é examinar um tema que tenha sido pouco ou nada estudado anteriormente, permitindo que se obtenha um maior grau e familiaridade com os fenômenos envolvidos.

No caso desse estudo, tendo em vista a escassez de produção relacionada aos termos *big data* e decisão no campo da ciência da informação, caracteriza-se como pesquisa básica, uma vez que algumas investigações são fundamentais para sustentar futuras aplicações.

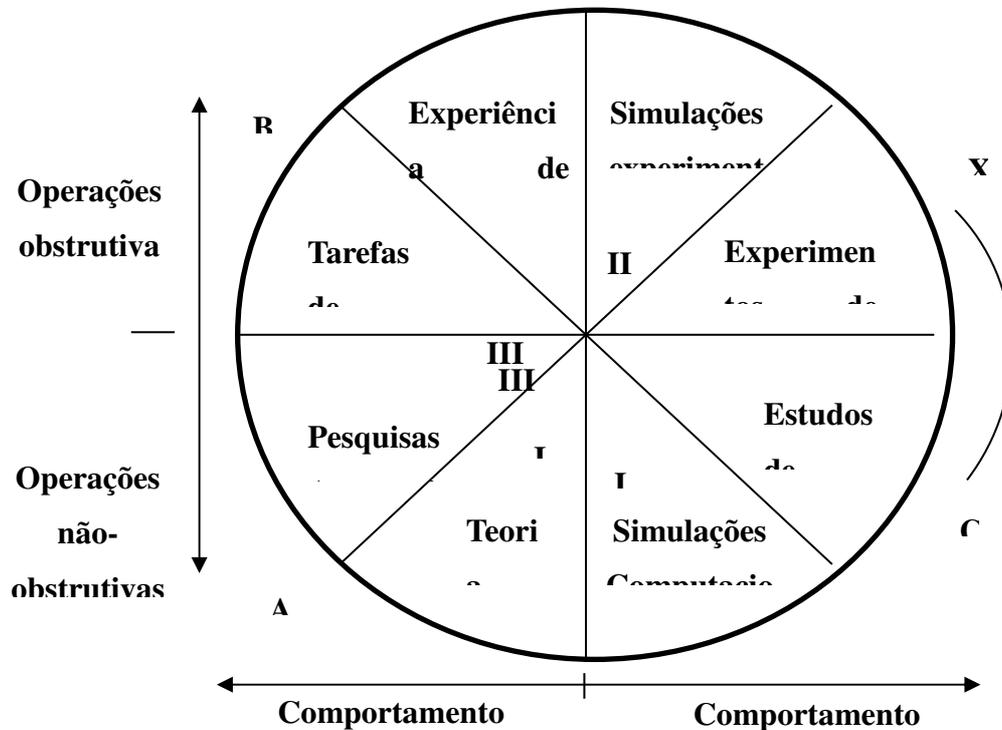
Segundo Schwartzman (1979), a pesquisa básica acumula informações e conhecimentos que podem eventualmente levar a resultados acadêmicos ou aplicados importantes, mas sem fazê-lo diretamente.

Macgrath (1982) propõe um modelo circunplexo para representar as etapas de pesquisa e facilita a identificação quanto a diversos aspectos, conforme se pode verificar na figura 1.

O circunplexo das estratégias de pesquisa permite que o pesquisador percorrer circularmente as diferentes estratégias

orientando-o na escolha da melhor opção de acordo com os objetivos que procura atingir em cada etapa do processo investigativo.

Figura 1 - Circumplexo das etapas de pesquisa de Macgrath



Fonte: Macgrath (1982).

Segundo o modelo, as etapas A, B e C da Figura 1, representam o ponto circumplexo em que existe uma atenção máxima sobre um dado objetivo de pesquisa, em detrimento dos outros dois. Percorrendo-se circularmente o circumflexo é possível fazer a escolha de uma estratégia, onde sempre será uma negociação compensatória de um objetivo em relação a outros dois. Sendo assim, os pontos A, B e C são máximos do respectivo objetivo, e mínimo dos outros dois.

No caso do tipo de experimento têm-se: I. Ocorre no ambiente natural; II. Ocorre em ambiente controlado e criado artificialmente; III. Comportamento do sistema não depende do ambiente; IV. Não requer observação do comportamento do sistema.

Nesse sentido, de acordo com o modelo proposto (Figura 1) esse estudo situa-se no campo IV do circumplexo de Macgrath (1982), ou

seja, no campo da teoria formal, de caráter dedutivo, modelo orientado em direção ao comportamento universal dos sistemas, etapa essa que será preliminar para as simulações computacionais e outros passos de validação empírica, de acordo com o objetivo do pesquisador.

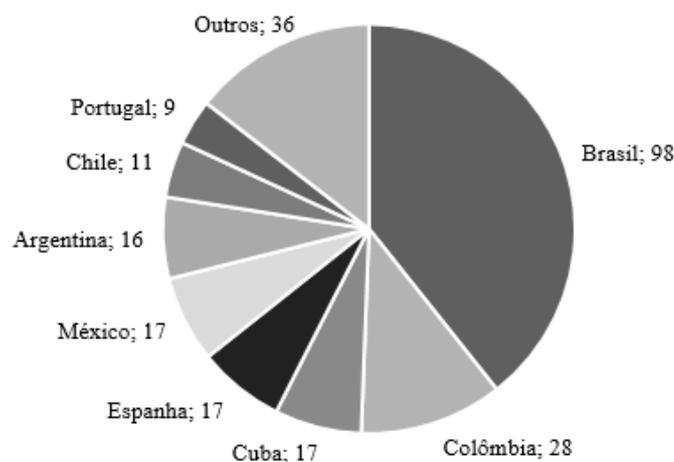
No que se refere a fonte de dados, esse estudo é bibliográfico com seleção intencional de referencial teórico baseado no conhecimento do pesquisador e análise geral da literatura.

Para Oliveira (2007) a pesquisa bibliográfica é uma modalidade de estudo e análise de documentos de domínio científico tais como livros, periódicos, enciclopédias, ensaios críticos, dicionários e artigos científicos. Como característica diferenciadora ela pontua que é um tipo de estudo direto em fontes científicas, sem precisar recorrer diretamente aos fatos/fenômenos da realidade empírica.

3 BIG DATA E DECISÃO NO CAMPO DA CIÊNCIA DA INFORMAÇÃO NO BRASIL

Na base de dados SCIELO, uma busca em 10/07/2015 com o termo “*big data*” retornou 249 artigos sendo 98 publicados no Brasil, seguido pela Colômbia (28), Cuba e Espanha (17), representados no gráfico 1.

Gráfico 1 - "big data", por país

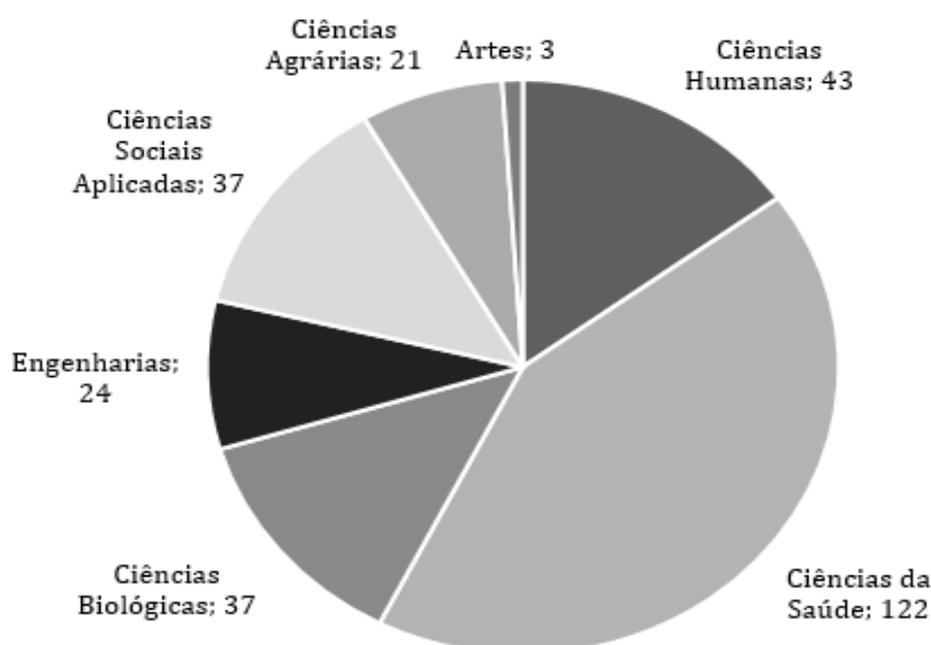


Fonte: SCIELO (2015).

O Brasil destaca-se na produção científica sobre *big data*, seguido pela Colômbia, Espanha, México, Cuba e Argentina.

As principais áreas de conhecimento das ocorrências identificadas foram: Ciências da Saúde (122); Ciências Humanas (43); Ciências Sociais Aplicadas e Biológicas (37) e Engenharias (24), representados no gráfico 2.

Gráfico 2 - "big data", por área de conhecimento



Fonte: SCIELO (10/07/2015)

Destaca-se a produção no campo da saúde, seguida das ciências sociais e humanas. Da produção em Ciências Sociais Aplicadas, 15 artigos são brasileiros.

A associação do termo *big data* AND decisão, não retornou resultados assim como *big data* AND ciência da informação ou *big data* AND gestão da informação. A busca pelos termos nos títulos também não obteve qualquer resultado.

A busca na Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação – BRAPCI, restou sem resultados tanto para

o termo individual como para associações citadas tanto em palavras-chave como em título. O mesmo se deu na busca nos Anais do Grupo de Trabalho 4 (GT-4) do Encontro Nacional de Ciência da Informação de 2014.

Das 37 publicações em Ciências Sociais Aplicadas obtidas na base SCIELO, somente com o termo “*big data*”, identificou-se que 15 artigos foram publicados no Brasil, sendo 1 no campo da Ciência da Informação, ou seja: BARBOSA, Eduardo Batista de Moraes; SENA, Galeno de. *Scientific data dissemination a data catalogue to assist research organizations*. Ci. Inf., Brasília, v. 37, n. 1, p. 19-25, abr. 2008.

Dessa forma, pelo indicativo acima é possível verificar a escassez de produção científica tanto no campo das ciências sociais aplicadas no Brasil, quanto no da Ciência da Informação e subárea de Gestão da Informação; indexada nas palavras-chave com o termo “*big data*”, bem como na associação com ciência e gestão da informação.

A abordagem da temática do *big data*, de maneira geral, encontra-se em estágio inicial na literatura em geral e, particularmente no campo da ciência da informação apresenta-se como um desafio a investigar. E ainda mais quando associada aos processos de tomada de decisão, ciência da informação e gestão da informação.

4 ASPECTOS RELEVANTES PARA ABORDAGEM DO FENÔMENO *BIG DATA* PARA TOMADA DE DECISÃO

A operacionalização científica do termo *Big Data* demanda alguns aprofundamentos necessários não somente à sua compreensão como fenômeno, mas também ao seu sentido, dimensões e características a serem abordadas, para além de uma expressão mercadológica e genérica por vezes mal definida ou utilizada inadequadamente.

Para fins desse estudo, considera-se que *Big Data* se refira à terceira época da era da informação relacionada diretamente a enorme quantidade de dados estruturados e desestruturados, cujo tamanho, natureza, origem ou formato estão além da habilidade de ferramentas de

software comuns de capturar, armazenar, gerenciar e analisar (MANYIKA et al., 2011; MINELI; CHAMBERS; DHIRAJ, 2013).

Na prática, *Big Data* resume de maneira geral um conceito de explosão de dados de forma incontrolável e a necessidade de transformar esses dados em informações que possam ser utilizadas para direcionar os negócios e as estratégias das organizações, minimizar riscos, e apoiar o processo de tomada de decisões.

Para Zikopoulos et al. (2012) *Big Data* se caracteriza por quatro aspectos: volume, velocidade, variedade e veracidade de dados. Volume refere-se a quantidade de dados disponível que vem crescendo de forma exponencial. A Velocidade diz respeito a rapidez que os dados podem ser capturados, processados e utilizados. Variedade tem a ver com as diversas fontes em que os dados podem ser encontrados e Veracidade se relaciona à capacidade de gerar informações úteis e oportunas.

Segundo a IBM (2013), o *Big Data* se caracteriza por cinco aspectos: a) Volume: refere-se à maior quantidade de dados sendo gerado a partir de uma variedade de fontes; b) Variedade : refere-se o uso de vários tipos de dados para analisar uma situação ou evento; c) Velocidade: o tempo para resolução de algum problema, como tratamento de dados, sua obtenção, atualização e gravação, devendo ser feito em tempo hábil, muitas vezes feito em tempo real; d) Veracidade: confiabilidade dos dados, ela garante o máximo possível de consistência nos dados e; e) Valor: benefício do investimento, o uso de Big Data exige um investimento que deve gerar retorno para as empresas como na melhoria da qualidade dos serviços e aumento da receita.

Para Davenport (2012) as organizações estão diante de uma expansão de dados que são muito volumosos e desestruturados para serem gerenciados e analisados através de meios tradicionais, levando a necessidade de se pensar em formas de analisar e processar tais dados a fim de gerar informações pertinentes e oportunas.

McAfee e Brynjolfsson (2012) evidenciam outro aspecto relevante de Big Data: a velocidade em que dados podem ser capturados e processados, quase em *real time*, podendo dar a uma organização vantagem competitiva.

Na intersecção entre os elementos do contexto do *big data* com a tomada de decisão verificam-se ao menos três problemáticas críticas que merecem ser aprofundadas.

Primeiramente, há o problema do tamanho, desestrutura e variedade de fontes de dados.

Há dificuldade quando o tamanho do conjunto de dados vai além da capacidade para capturar, armazenar, gerenciar e analisar das ferramentas de software de banco de dados típicos (MANYIKA et al., 2011).

Os analistas deparam-se com a existência de grandes volumes de dados, inúteis em si mesmos e estão obrigados a extrair os subconjuntos relevantes para poder desenvolver algum tipo de análise (INFORMATION BUILDERS, 2013).

As organizações estão lidando com uma expansão de dados que são muito volumosos e desestruturados para serem gerenciados e analisados por meios tradicionais, levando à necessidade de se pensar em formas de analisar e processar tais dados a fim de gerar informações pertinentes e oportunas (DAVENPORT, 2012).

Além disso, há aspectos específicos dos contextos em que o volume de dados disponível é relativamente pequeno, esparso, circunscrito culturalmente, de difícil obtenção e/ou não disponível em formato digital.

Em segundo lugar, há o problema das formas comumente utilizadas para tomada de decisão.

Do ponto de vista estratégico, as organizações sentem cada vez mais a necessidade encontrar instrumentos que facilitem a aquisição, o processamento e a análise de grandes quantidades de dados, provenientes de fontes diferentes e dispersas pela organização, e que

servam como uma base sólida para descobrir novo conhecimento (OLSZAK; ZIEMBA. 2007)

É usual que as pessoas situadas em pontos elevados da hierarquia organizacional tomem decisões com base em sua experiência, o que fazem geralmente levando em conta padrões de relacionamentos que internalizaram ao longo de suas carreiras (MCAFEE; BRYNJOLFSSON, 2012).

Segundo McAfee e Brynjolfsson (2012), na ausência de tecnologias e ferramentas capazes de dar suporte adequado à decisão no contexto do *Big Data*, prevalece a tendência dos tomadores de decisão em se basear na intuição nos processos decisórios.

Além disso, destaca-se a necessidade de lidar com as diversas acepções polissêmicas do termo '*Big Data*', a inserção de seu uso em contextos diversificados das organizações e representação da informação cujas fronteiras, objetivos e interesses não estão bem definidos, mudam ao longo do processo de tomada de decisão e demandam resultados o máximo possível em tempo real em ambientes de incerteza e complexidade.

De forma geral, do ponto de vista de abordagens teóricas encontradas de forma recorrente na literatura, identifica-se decisão no contexto do *big data* tem se valido predominantemente de abordagens e ferramentas decorrentes de *Business Intelligence* e *Big Data Analytics*.

5 ASPECTOS BÁSICOS DOS CONCEITOS DE BUSINESS INTELLIGENCE E BIG DATA ANALYTICS E A TOMADA DE DECISÃO

Os conceitos de *Business Intelligence* e *Big Data Analytics* são os mais recorrentes na literatura que trata de *big data* no sentido de gestão da informação e seus processos de decisão.

No que se refere a *Business Intelligence*, o termo surgiu em 1996 a partir do *Gartner Research Group* sendo concretizado através de diferentes sistemas.

Turban, Sharda e Delen (2010) definem *Business Intelligence* como um termo “guarda-chuvas” que combina arquiteturas, ferramentas, bases de dados, ferramentas analíticas, aplicações e metodologias.

Para Santos e Ramos (2006) os sistemas de *Business Intelligence* combinam o levantamento de dados operacionais que permitam o seu posterior armazenamento em repositórios adequados para análise, exploração e apresentação de informação essencial para a tomada de decisão.

Nas últimas décadas, tem-se verificado um grande crescimento no número de produtos e serviços de *Business Intelligence* oferecidos, assim como na adoção destes por parte das organizações (CHAUDHURI; NARASAYYA, 2011).

Em geral, o *Business Intelligence* utiliza *Data Warehouses* ou *Data Marts* como repositórios de informação organizacional e incluem sistemas de processamento analítico de dados e mecanismos de *Data Mining*, sendo os primeiros para uma análise multidimensional dos dados e os segundos para a extração de conhecimento e identificação de padrões e modelos preditivos nos dados.

No que se refere ao conceito de *Big Data Analytics*, segundo a INTEL (2012), tratam-se de diversas estratégias de tecnologia para deixar a percepção mais rica, profunda e precisa no que se refere a clientes, analisando padrões e correlações, ganhando por fim vantagem competitiva. Objetiva processar um fluxo constante de dados em tempo real, de forma que as organizações possam tomar decisões com mais rapidez, acompanhar as tendências emergentes, corrigir cursos rapidamente, e investir em novas oportunidades de negócios.

No desdobramento do *Big Data Analytics* têm-se ainda diversas tecnologias e sistemas cujo objetivo é melhorar a qualidade da gestão estratégica nas organizações por meio de novas tecnologias e diversas técnicas para extração, transformação, processamento e apresentação de dados e fornecimento de informação em tempo real.

A identificação das principais abordagens utilizadas no contexto do *big data* aponta para uma grande diversificação de seu potencial de solução, usos e ferramentas desenvolvidas no ambiente corporativo e competitivo com base em métodos computacionais.

Não obstante o paradigma tradicional quantitativo apresente reconhecida solidez do ponto de vista lógico e técnico, do ponto de vista empírico apresenta limitações principalmente porque no contexto do *Big Data* as decisões normalmente não podem ser tomadas com base num só critério, como no paradigma tradicional, mas a partir de objetivos, interesses ou critérios múltiplos de um ou mais decisores, além da interação de diversificados elementos exigem modelagem para aplicações em diversos ambientes.

Segundo Bourne et al. (2002) dentre as causas para prevalência de abordagens fortemente focadas nas dimensões quantitativas estariam: a) demasiado tempo e esforço requerido para mudança; b) consequências pessoais oriundas da implementação de novas medidas; c) decisão de não alterar o conjunto de medidas para não mostrar performance adversa; d) medo por parte daqueles que serão avaliados; e) visão das medidas como ataques pessoais; f) falta de continuidade no comprometimento da alta gerência; g) benefícios incertos; h) cada um ter sua relação de prioridades.

Uma abordagem estritamente quantitativa da decisão comporta uma parte de ilusão, pois hoje é difícil modelar uma realidade complexa e cambiante, visto, ela não ser unicamente composta de elementos materiais ou técnicos facilmente quantificáveis, mas possui dimensão social importante, onde as interações informais entre as entidades individuais não podem ser esquecidas (SHIMIZU, 2001).

Outra questão a considerar é quanto às limitações de modelos genéricos. Corre-se o risco de, diante da complexidade e da incerteza reduzir as questões a busca de “soluções ótimas” baseadas em modelos matemáticos e simulações ausentes de clareza quanto ao problema, sua natureza, os objetivos da a seleção de critérios aleatórios sob os quais o

problema será avaliado, ou seja, soluções para problemas que não se sabe exatamente quais são; soluções para problemas engessadamente dados em contextos mutantes ou ainda soluções descontextualizadas. (ENSSLIN et al, 2000; VIANNA; ENSSLIN; ENSSLIN, 2007).

Tangen (2003), Denton (2005), Neely e Powell (2004) consideram que a adoção de medidas que se dizem aplicáveis a todas as organizações, ou a todas de um mesmo setor, uma vez que resolver os problemas por meio do uso de medidas genéricas pré-existentes, por serem mais fáceis de medir, é um erro estratégico.

Em decorrência da limitação da adoção unilateral da objetividade nos processos decisórios há necessidade de interação entre os elementos objetivos e subjetivos num processo de tomada de decisão. Desta forma fica impossível negar a importância dos fatores subjetivos e deixá-los à parte na tentativa de utilizar uma abordagem inteiramente objetiva (ROY; VANDERPOOTEN, 1996).

Para Roy e Vanderpootem (1996) é impossível negar a importância destes fatores subjetivos e colocá-los de lado em favor do uso de uma abordagem totalmente objetiva.

Esse aspecto demonstra a relevância no desenvolvimento e aperfeiçoamento de associações dos instrumentos comumente utilizados no contexto do *big data* com aqueles que têm sido utilizados no apoio à decisão.

Tal propósito e desafio tem sido identificado como um dos desafios de abordagens relevantes do *big data*.

Segundo a *Information Builders* (2013), fabricante especializado em *Business Intelligence* (BI), o cargo de “*Data Scientist*” (cientista de dados), dará mais ênfase à gestão e supervisão da informação e não apenas à parte científica, aspecto que se encontra em seu estágio inicial de desenvolvimento.

Nesse contexto, para tomada de decisão é um desafio entender como esses dados evoluem e como utilizá-los, sendo que essa tarefa criará, em 2015, segundo a grande empresa norte-americana de

consultoria Gartner Group, cerca de 4,4 milhões de empregos – muitos deles para profissionais capacitados a analisar esses dados – em um mercado mundial avaliado, já para este ano, em US\$ 34 bilhões (cerca de R\$ 70 bilhões).

As competências em “*data steward*” (gestor de dados) serão cada vez mais necessárias aos profissionais da informação e demandam entender os dados e saber como qualificá-los e usá-los no decurso do processo decisório. Em outras palavras, tão importante quanto a tecnologia é ter pessoas capazes de saber o que fazer com essas informações e que estejam qualificadas na tomada de decisão.

6 TOMADA DE DECISÃO E MODELAGEM PRELIMINAR

A tomada de decisão exige aplicação hábil de métodos e processos científicos e não científicos pelos quais os indivíduos interpretam dados ou informações para produzir resultados de inteligência perspicazes e recomendações viáveis (FLEISHER; BENSOUSSAN, 2007).

Segundo Mandel, Simon e Delyra (1997) em todas as etapas da manipulação da informação é necessária a presença do ser humano e suas limitações influenciam na capacidade de aquisição de dados e processamento de grande volume de informações e este aspecto constituem o principal gargalo do processo.

Entretanto, de maneira geral, as abordagens que envolvem processos de tomada de decisão tradicionalmente consideram que a decisão é condicionada pelo objetivo de maximização do lucro, ou seja, atinge-se uma solução ótima, selecionada dentre diversas alternativas, geralmente através de um problema de programação linear, com um objetivo sujeito a um conjunto x de restrições.

Segundo Bana e Costa (1993) é preciso aceitar definitivamente que a subjetividade está onipresente nos processos de tomada de decisão.

Os novos contextos que envolvem *big data* as decisões necessitam considerar cada vez mais uma variedade de fontes tais como as redes sociais que necessitam, por vezes, o uso de técnicas de análise de sentimento, isto é, a análise de quantidades imensas de dados desestruturados criados pelos clientes em mídias sociais, *blogs* e outras fontes na internet (BROWN; MANYIKA; CHUI, 2011).

Para Simon (1947) os tomadores de decisão possuem habilidades limitadas para avaliar todas as possíveis alternativas de uma decisão, bem como lidar com as consequências incertas da decisão tomada.

Simon (1979) observa ainda que a teoria clássica leva sempre as mesmas conclusões e que os princípios da racionalidade perfeita são contrários aos fatos que realmente ocorrem no processo real de tomada de decisão.

É ainda interessante considerar que Simon (1990) utilizou ensaios com recursos computacionais (*hardware e software*) e sistemas de Inteligência Artificial, tanto para entender como para aumentar a capacidade de pensamento e ampliar a fronteira da racionalidade limitada, ou seja, considerar as habilidades que determinam a qualidade da tomada de decisão e solução de problemas, além de serem armazenadas na mente das pessoas, podem ser armazenadas em computadores (SIMON, 1987).

Segundo Balestrin (2002), Simon ainda usou a tecnologia computacional para simular técnicas de pesquisa operacional na modelagem de sistemas complexos uma vez que considera que a modelagem é a principal – talvez a primeira – técnica para estudar o comportamento de sistemas complexos, sendo usados para prever, para analisar e para prescrever algo sobre esses sistemas.

Dessa forma, verifica-se a necessidade de processos de modelagem do contexto decisional, tendo em vista aumentar o conhecimento sobre o processo de tomada de decisão em contextos reais. Tal modelagem ao evidenciar o problema e os interesses e

objetivos dos decisores pode melhor qualificar e legitimar a seleção e o uso dos recursos computacionais e/ou abordagens e ferramentas decorrentes de *Business Intelligence* e *Big Data Analytics*. Tal já ocorre de diversas formas com os instrumentos de Pesquisa Operacional.

7 A ESTRUTURAÇÃO DO CONTEXTO DECISIONAL

A definição do que fazer com tantos dados, como selecionar dados úteis, estruturá-los e analisá-los, ou seja, o processo de tomada de decisão no contexto do *big data* pode ser facilitado pela modelagem preliminar do contexto decisional, onde será possível evidenciar o problema e os objetivos e interesses aos quais se destinam a obtenção da informação, ou seja a consideração adequada do tomador de decisão.

Para Roy e Vanderpooten (1996), o jogo destas confrontações e interações, sob os vários efeitos de compensação e amplificações do sistema, é o que chamamos de processo decisório. Assevera o autor que ao privilegiar o caminho realista, admite-se que existe um problema real e que a formulação consiste em representar o mesmo (descrevê-lo) e só então, modelá-lo ou seja, representá-lo matematicamente ou computacionalmente.

Segundo Landry (1995), a definição de problema é de importância primordial no processo de encontrar as formas para solucioná-lo, na literatura isto é amplamente reconhecido, mas na prática é ignorado.

Eden et al. (1983) define problema como uma situação onde alguém deseja que alguma coisa seja diferente de como ela é, e não se esta seguro de como obtê-la ou modificá-la.

Nesta perspectiva, segundo os dizeres Landry (1995) os problemas são entendidos por meio de quatro aspectos: a) necessitam ter um “dono”, b) é necessário evidenciar a situação percebida como insatisfatória, c) ser significativo e d) possível de ser resolvido.

Ensslin, Montibeller Neto e Noronha (2001) sustenta que um problema pertence a uma pessoa ou grupo, isto é, ele é um constructo que o indivíduo faz sobre os eventos associados ao contexto decisório, não é uma entidade física e nenhuma situação é inerentemente ou “objetivamente” um problema.

Para Bana e Costa (1993) é preciso aumentar os conhecimentos dos atores sobre o problema, clarificar vários elementos, torná-los operacionais, encontrar suas inter-relações, ou seja, proceder à estruturação que servirá de base das operacionalizações.

Ao final do processo de modelagem preliminar espera-se que: a) o que é importante se sobressaia mais claramente do que é menos importante; b) seja possível separar as conclusões robustas das frágeis; c) estejam dissipados maus entendimentos, redundância e ambiguidades que possam ocorrer na comunicação dos decisores; d) seja possível evitar as armadilhas ilusórias do bom senso; e) haja ênfase, uma vez entendidos, nos resultados não controversos. (ROY; VANDERPOOTEN, 1996).

7.1 A Modelagem Preliminar do Contexto Decisional e a Contribuição da Sistemografia

No que se refere a uma proposta de modelagem preliminar do contexto decisional com incorporação da subjetividade, considera-se que é fundamental assumir que, inicialmente resolver um problema é um processo, com alguns dados novos sendo incorporados e outros descartados ao longo do tempo, e que novas questões podem aparecer e substituir as originalmente propostas e os fenômenos podem ser caracterizados por um alto grau de complexidade interna, deslocando as fronteiras de interação entre objeto, sujeito e ambiente, o que gera novos contextos (VIANNA; ENSSLIN, 2008).

Dentre as convicções de cientificidade que fundamentam a proposta se encontram: a) limitação de soluções, ferramentas e técnicas aplicáveis a qualquer contexto decisional; b) necessidade de

modelagem preliminar do contexto decisional com incorporação da subjetividade.

Para tal, é fundamental considerar e representar na modelagem os elementos objetivos e subjetivos que interagem no contexto decisório, num sistema aberto do qual são componentes os atores, com seus valores e objetivos, e as ações (alternativas e possibilidades) com suas características (BANA E COSTA, 1993).

Nesse sentido, o modelo ao ser operacionalizado não será mais visto como a aproximação de uma realidade externa aos atores. Ao contrário, é considerada uma representação construída das suas percepções em relação aos elementos da realidade e tem como objetivo funcionar como uma ferramenta útil.

Para Le Moigne (1977), modelar é conceber, para um objeto, um modelo que permita conhecê-lo, compreendê-lo, interpretá-lo e antecipar seu comportamento. O observador é um sujeito ativo que procede a uma descrição comunicável do que percebe e do que concebe.

No caso, a sistemografia é considerada uma técnica de modelagem gráfica de sistemas complexos para a concepção, análise e simulação dos sistemas de um modo geral, visando busca da racionalidade, agilidade e flexibilidade desses sistemas (BRESCIANI FILHO, 2001; LE MOIGNE, 1990).

No que se refere à “sistemografia” é uma designação de Le Moigne (1977) para demonstrar a capacidade de o sistema de agir como um instrumento para modelar objetos por meio da concepção do modelo e sua representação por meio de signos. Há um vínculo profundo entre forma e função que precisa ser considerado na representação do objeto.

Segundo Neto e Leite (2010) a sistemografia consiste basicamente em cinco etapas: 1) identificar o fenômeno, definir o contexto de pesquisa, identificar o sistema, os sistemas relacionados, o ambiente imediato e geral, as fronteiras de cada sistema; 2) desenvolver o modelo geral de referência por agregação de todos os conhecimentos acessíveis a ele e que possam ser utilizados para observar o fenômeno

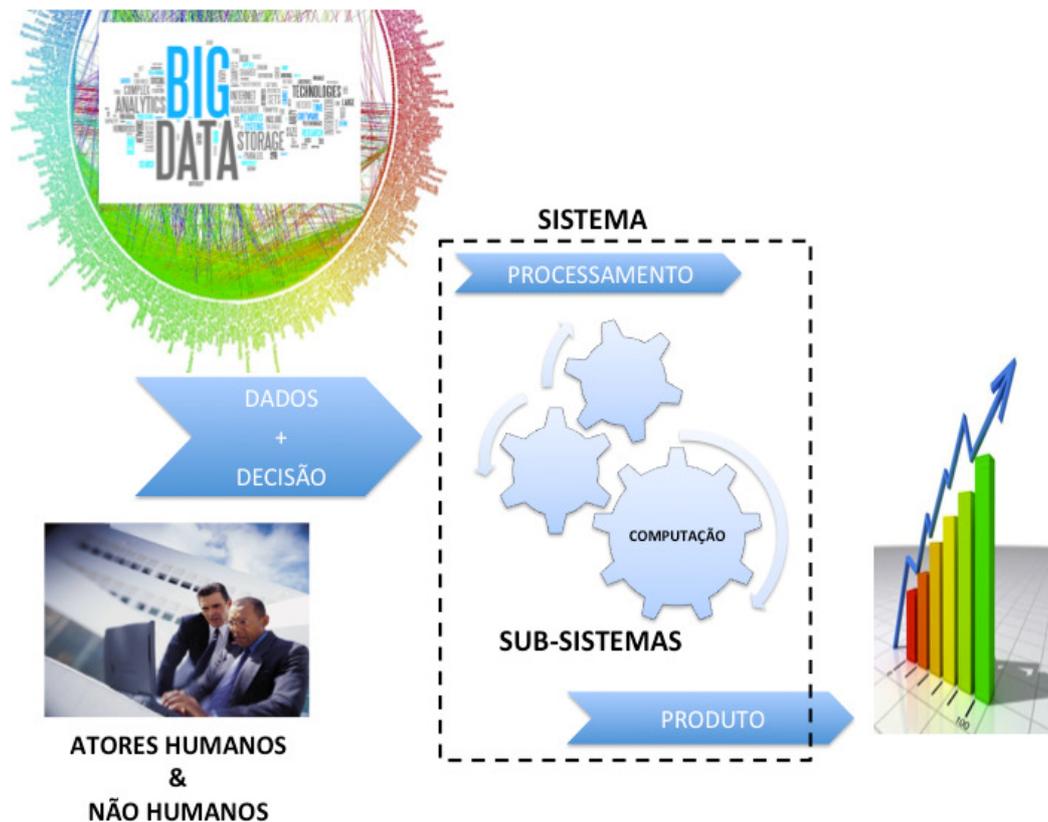
identificado; 3) observar a realidade e o fenômeno real por isomorfismos com base no modelo conceitual de base; 4) desenvolver modelos da realidade mais inteligíveis aos olhos do pesquisador/observador e atores; 5) agir sobre a realidade.

A partir das etapas propostas por Neto e Leite (2010), será apresentada uma representação sistemográfica geral da etapa 1 (identificação do fenômeno) para visualização gráfica dos elementos até aqui levantados, tendo em vista possibilitar a geração de protocolos novos e úteis de pesquisa, a identificação clara de variáveis e dimensões para o avanço do conhecimento no campo da gestão da informação.

7.2 Representação Sistemográfica do Contexto Decisional

Tendo em vista a modelagem preliminar do contexto decisório para o uso de dados obtidos no contexto do *Big Data*, propõe-se a representação dos elementos que compõe o contexto decisional na figura 2, onde se esboça uma representação sistemográfica do contexto geral.

Figura 2 – Representação sistemográfica do contexto decisional de *Big Data*



Fonte: Os autores

A representação sistemográfica apresenta como *inputs* dados e decisores. A explicitação desses dois elementos de entrada é preliminar ao uso das ferramentas de processamento que resultarão nos produtos desse processo.

Tanto dados como decisores são elementos centrais do processo decisório no contexto do *big data*, tal como se explicitou nas seções anteriores. Entretanto, para que faça sentido, a participação dos decisores é que legitima o processo possibilitando maior incorporação dos aspectos subjetivos ao processo.

Os *inputs* representados de forma geral na figura 2, necessitarão de explicitação posterior dos elementos (dados + decisão) e suas inter-relações que serão considerados na fase seguinte, ou seja, quais dados e informações serão explorados do universo do *big data*.

A proposta altera a lógica tradicional dos chamados sistemas de apoio à decisão que são fundamentados em softwares que através de informações e modelos especializados se propõe apoiar a tomada de decisão.

As engrenagens não definidas que se associam à computação surgirão do processo de modelagem podendo ser, por exemplo: comunicação, tratamento de dados, ontologias, tratamento técnico ou outras.

Os decisores precisarão explicitar seus objetivos e interesses e, a partir deles, os dados ou as coleções de informações que se relacionam, geralmente desestruturados, poderão ser melhor processados pelos instrumentos de *Business Intelligence*, *Big Data Analytics* ou outros.

Tal procedimento, embora útil, não se encontra evidenciado na literatura, lacuna essa já discutida nas seções anteriores.

Há uma escassez de pessoas com as habilidades necessárias para aproveitar as ideias que grandes conjuntos de dados podem gerar, sendo essa uma das restrições mais importantes quanto a capacidade de uma organização capturar o potencial de *Big Data*. (MANYIKA et al., 2011).

Dessa forma, considera-se que esse estudo seja uma contribuição para o aumento de conhecimento e de possibilidades de gestão da informação apoiada pela sistemografia e sua consequente potencialidade de apoiar a modelagem do contexto decisional.

8 CONSIDERAÇÕES FINAIS

O estudo objetivou identificar e representar os elementos gerais de um contexto de tomada de decisão que envolva *big data*.

Na seção 3 verificou-se que a literatura sobre o tema é recente e escasso no Brasil no campo da ciência da informação há uma identificação de sua relevância, entretanto, a área carece de estudos sobre o assunto.

Na seção 4 identificou-se que entre os principais problemas relacionados ao *big data* estão a questão do tamanho, desestrutura e variedade de fontes de dados, bem como as formas comumente utilizadas para tomada de decisão baseadas na intuição ou em modelos quantitativos e computacionais.

A seção 5 apresentou aspectos básicos dos conceitos de *Business Intelligence* e *Big Data Analytics* que são ferramentas que tem emergido como uma área de sistemas de apoio à decisão.

Na seção 6 analisou-se a importância da modelagem preliminar do contexto decisional, tendo em vista a racionalidade limitada do tomador de decisão.

A seção 7 resgatou alguns conceitos fundamentais da estruturação do contexto para o processo de decisão, particularmente a necessidade de consideração dos elementos objetivos e subjetivos; a importância do decisor e seus objetivos e interesses; e o papel da sistemografia na modelagem preliminar do contexto decisional.

No item 7.2 é apresentado um modelo geral de representação sistemográfica que evidencia, sintetiza e explicita os elementos anteriores da discussão, bem como sua utilidade e possibilidades de desenvolvimento.

Futuros estudos deverão ampliar e detalhar a partir de problemas reais, 'dados e decisão' e seus respectivos *inputs*, bem como selecionar ferramentas computacionais para simulação, conforme o modelo Mcgrath (1982), para prever o comportamento particular do sistema.

REFERÊNCIAS

BANA E COSTA, Carlos A. Três convicções fundamentais na prática do apoio à decisão. **Revista Pesquisa Operacional**, Rio de Janeiro, v. 13, n. 1, jun. 1993.

BALESTRIN, Alsones. Uma análise da contribuição de Herbert Simon para as teorias organizacionais. **REAd - Revista Eletrônica de Administração**, Porto Alegre, v. 8, n. 4, jul./ago. 2002.

BOURNE, Mike et al. The success and failure of performance measurement initiatives: perceptions of participating managers. **International Journal of Operations & Production Management**, United Kingdom, v. 22, n. 11, p. 1288-1310, 2002.

BRESCIANI FILHO, Ettore. Métodos de estudo de sistema: sistemografia. **Revista Eletrônica de Informática**, Campinas, v. 1, n. 1, 2001. Disponível em:
<http://www.puccampinas.edu.br/centros/ceatec/revista_eletronica/primeira_edicao.html>. Acesso em: 28 dez. 2014.

BROWN, Brad; CHUI, Michael; MANYIKA, James. Are you ready for the era of 'big data'? **McKinsey Quarterly**, New York, out. 2011.

CHAUDHURI, Surajit; NARASAYYA Vivek. New frontiers in business intelligence. In: INTERNATIONAL CONFERENCE ON VERY LARGE, 37., 2011, Seattle, Washington. **Proceedings...** Seattle, Washington, 2011. p. 1052-1053.

DAVENPORT, Thomas H. **Ecologia da informação**: porque só a tecnologia não basta para o sucesso na era da informação. São Paulo: Futura, 1998.

DAVENPORT, Thomas H. **Enterprise analytics**: optimize performance, process, and decisions through big data. New Jersey: FT Press Operations Management, 2012.

DENTON, Keith. Measuring relevant things. **International Journal of Productivity and Performance Management**, United Kingdom, v. 54, n. 4, p. 278-287, 2005.

EDEN, Colin; JONES, Sue; SIMS, David. **Messing about in problems**. Oxford: Pergamon, 1983.

ENANCIB - ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 4., 2014, Belo Horizonte. Disponível em:
<<http://enancib2014.eci.ufmg.br/sobre>> Acesso em: 17 jul. 2015.

ENSSLIN, Leonardo; DUTRA, Ademar; ENSSLIN, Sandra Rolim. MCDA: a constructivist approach to the management of human resources at a governmental agency. **International Transactions in Operational Research**, Oxford, v. 7, n. 1, p. 79-100, 2000.

ENSSLIN, Leonardo; MONTIBELLER NETO, Gilberto, NORONHA, Sandro MacDonald. **Apoio à decisão**: metodologias para estruturação de problemas e avaliação multicritério de alternativas. Florianópolis: Insular, 2001.

FLEISHER, Craig S.; BENSOUSSAN, Babette. **Business and competitive analysis**: effective application of new and classic methods. New Jersey: Pearson Education, 2007.

GANTZ, John; REINSEL, David. The digital universe in 2020: big data, nigger digital shadows, and biggest growth in the far east. **IDC**, New York, v. 3, Dec. 2012.

IAROZINSKI NETO, Alfredo; LEITE, Maria Silene. A abordagem sistêmica na pesquisa em engenharia de produção. **Produção**, São Paulo, v. 20, n. 1, Mar. 2010.

INFORMATION BUILDERS. **Conceito de Big Data vai perder relevância**. 2013. Disponível em: <<http://www.computerworld.com.pt/2013/12/24/conceito-de-big-data-vai-perder-relevancia>>. Acesso em: 17 jul. 2014.

LANDRY, Maurice. A note on the concept of problem. **Organization Studies**, Berlin, n. 16, p. 315-343, 1995.

LE COADIC, Yves-François. Princípios científicos que direcionam a ciência e a tecnologia da informação digital. **Transinformação**, Campinas, v. 16, n. 3, p. 205-213, set./dez. 2004.

LE MOIGNE, Jean-Louis. **La théorie du système général, théorie de la modélisation**. Paris, 1977.

LE MOIGNE, Jean-Louis. **La modélisation des systemes complexes**. Afcet systemes. Paris: Dunod, 1990.

MCAFEE, Andrew, BRYNJOLFSSON, Erik. Big Data: the management revolution. **Harvard Business Review**, Boston, v. 90, n. 10, p. 60-66, out. 2012.

MANDEL, Arnaldo, SIMON, Imre; DeLYRA, Jorge L. **Informação**: computação e comunicação. São Paulo: Universidade de São Paulo, 1997. Disponível em: <<http://www.ime.usp.br/~is/abc/abc/node1.html>>. Acesso em: 9 jan. 2015.

MANYIKA, James et al. **Big data**: the next frontier for innovation, competition, and productivity. McKinsey Global Institute Report, 2011. Disponível em: <http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation>. Acesso em: 11 jan. 2015.

MCGRATH, Joseph E. Dilemmas: the study of research choices and dilemmas. In: McGrath, J.E.; Martin, J.; Hulka, R. A. (Ed.). **Judgement calls in research**. Beverly Hills: Sage, 1982. p. 69-102

MINELI, Michael; CHAMBERS, Michele; DHIRAJ Ambiga. **Big data, big analytics**: emerging business intelligence and analytic trends for today's businesses. New Jersey: John Wiley & Sons, 2013.

NEELY, Andy; POWELL, S. The challenges of performance measurement. **Management Decision**, York, v. 42, n. 8, p. 1017-1023, 2004.

OLSZAK, Celina M.; ZIEMBA, Ewa. Approach to Building and Implementing Business. **Interdisciplinary Journal of Information**, Santa Rosa, Califórnia, n. 2, p. 135-148, 2007.

PINTO, Adilson Luiz; ELIAS, Ezmir Dippe; VIANNA, William Barbosa. Requisitos para métricas em arquivos: critérios específicos para arquivometria. **Perspectivas em Ciência da Informação**, v. 19, n. 3, p. 134-148, jul./set. 2014.

OLIVEIRA, Maria Marly de. **Como fazer pesquisa qualitativa**. Petropolis: Vozes, 2007.

RAPPORTEUR, David Bollier. **The promise and peril of big data**. Washington: The Aspen Institute. 2010.

ROY, Bernard; VANDERPOOTEN, Daniel. The european school of MCDA: emergence, basic features and current works. **Journal of Multi-Criteria Decision Analysis** Oxford, v. 5, n.1, p. 22-38, 1996.

SAMPIERI, Roberto Hernández; COLLADO, Carlos Fernández; LUCIO, Pilar Baptista. **Metodología de la investigación**. México: McGraw Hill, 1994.

SANTOS, Maribel Yasmina; RAMOS, Isabel. **Business intelligence: tecnologias da informação na gestão do conhecimento**. Lisboa: FCA, 2006.

SCHWARTZMAN, Simon. **Formação da comunidade científica no Brasil**. São Paulo: Nacional, 1979.

SIMON, Herbert. A. **Administrative Behavior**. New York: Macmillan, 1947.

SIMON, Herbert. A. Rational decision making in business organizations. **American Economic Review**, Nashville, v. 69, p. 493-513, 1979.

SIMON, Herbert A. et al. Decision making and problem solving. **Management Science**, Providence, v. 17, n. 5, p. 11-21, 1987.

SIMON, Herbert A. Prediction and prescription in systems modeling. **Operations Research**, Baltimore, v. 38, p. 7-14, 1990.

SHIMIZU, Tamio. Decisão nas organizações. São Paulo: Atlas, 2001.

TANGEN, Stefan. An overview of frequently used performance measures. **Work Study, Local**, v. 52, n. 7, p. 347-354, 2003.

TRIVIÑOS, Augusto Nivaldo Silva. **Introdução à pesquisa em ciências sociais: a pesquisa qualitativa em educação**. São Paulo: Atlas, 1987.

TURBAN, Efraim; SHARDA, Ramesh E.; DELEN, Dursun. Decision support and business intelligence systems. 9. ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2010.

VIANNA, William Barbosa; ENSSLIN, Sandra Rolim; ENSSLIN, Leonardo. O design na pesquisa quali-quantitativa em engenharia de produção questões a considerar. **Revista Gestão Industrial**, Ponta Grossa, v. 3, p. 172-185, 2007.

VIANNA, William Barbosa; ENSSLIN, Leonardo. O design na pesquisa quali-quantitativa em engenharia de produção: questões epistemológicas. **Revista Produção On-line**. Florianópolis, v. 8, n. 1, mar. 2008.

ZIKOPOULOS, Paul *et al.* Understanding big data: analytics for enterprise class hadoop and streaming data. New York: McGraw-Hill, 2012.

Title

Big data and information management: modeling the context decisional supported by sistemography.

Abstract

Introduction: The study justified by the scarcity of studies in the field of information science that addressing the phenomenon of big data from the perspective of information management. that will allow further development of computer simulation.

Objective: The objective is to identify and represent the general elements of the decision-making process in the context of big data.

Methodology: It is an exploratory study and theoretical and deductive nature

Results: It resulted in the identification of the main elements involved in decision-making on big data environment and its sistemografic representation

Conclusions: It was possible to develop a representation which will allow further development of computer simulation.

Keywords: Big Data. Decision Support. Sistemography. Modeling. Information Management.

Titulo

Big data y gestión de la información: modelaje en contexto de la toma de decisiones apoyado por la sistemografia

Resumen

Introducción: El estudio Justificada por los estudios de escasez y de servicios públicos en el ámbito de la información que abordan el fenómeno de grandes volúmenes de datos desde la perspectiva de gestión de la información.

Objetivos: El objetivo del estudio es identificar y representar los elementos generales del proceso de toma de decisiones en el contexto de grandes volúmenes de datos.

Metodología: Es un estudio exploratorio y la naturaleza teórica y deductiva.

Resultados: El resultado fue la identificación de los principales elementos que intervienen en la toma de decisiones sobre el medio ambiente de grandes datos y su representación sistemográfica.

Conclusiones: Fue posible desarrollar una representación que permitirá un mayor desarrollo de la simulación por ordenador.

Palabras clave: Big Data. Ayuda a la decisión. Sistemography. Modelaje. Gestión de la información.

Recebido: 18.09.2015

Aceito: 09.04.2016