

O termo Big Data: quebra de paradigma dos n-V's

Marcos de Souza^a, Fernanda Gomes Almeida^b e Renato Rocha Souza^c

Resumo: Os dados são considerados o novo petróleo da era digital. O crescimento e integração de grandes volumes de dados digitais – Big Data – têm sido utilizados para tomada de decisão em diferentes áreas. Entretanto, percebe-se que não existe uma universalização do conceito de Big Data, sobretudo, no meio acadêmico por se tratar de um termo relativamente novo. A pesquisa tem como objetivo identificar os conceitos relacionados ao termo Big Data e; identificar as características que compõem o conceito. A pesquisa foi elaborada com base em artigos científicos e livros digitais e impressos publicados entre os anos de 2011 e 2018. Como resultado percebe-se que a literatura apresenta tradicionalmente 5V's como características de um Big Data, sendo volume, variedade, velocidade, veracidade e viabilidade. Contudo, a pesquisa permitiu identificar mais 2V's e diversos conceitos para o termo, o que não permite assim uma consolidação do termo. Com o avanço tecnológico, novas possibilidades têm surgido e colocado as características em cheque. Sugere-se como pesquisas futuras uma caracterização real do termo de forma que o mesmo não esteja atrelado a determinada letra alfabética.

Palavras-chave: Big Data. Conceito. Viabilidade. Venalidade. n-V's.

The term Big Data: paradigm break of n-V's

Abstract: Data is considered the new oil of the digital era. The growth and integration of large volumes of digital data – Big Data – have been used for decision making in several areas. However, this is not an international model of big data. A research aims to identify the data related to the term Big Data and; identify as characteristics that make up the concept. The research was based on scientific and printed data between the years 2011 and 2018. As a result it realizes that a literature is traditionally produced as a large volume of data, being volume, variety, speed, veracity and feasibility. In this article, you may have more information about the concepts of a term that is not mandatory. With technological technology, new features have emerged and placed as features in check. It is suggested as a new form of characterization of the term so that it is not tied to an alphabetic letter.

Keyword: Keywords: Big data. Concept. Viability. Venality. N-V's.

-
- a Universidade Federal de Minas Gerais (UFMG). E-mail: marcosdesouza82@gmail.com. ORCID: <https://orcid.org/0000-0002-9829-7249>. Currículo: <http://lattes.cnpq.br/3958131052236839>
- b Universidade Federal de Minas Gerais (UFMG). E-mail: usernanda@gmail.com. ORCID: <https://orcid.org/0000-0001-7913-827X>. Currículo: <http://lattes.cnpq.br/5601300780102290>
- c Fundação Getúlio Vargas (FGV). E-mail: rsouza.fgv@gmail.com. ORCID: <https://orcid.org/0000-0002-1895-3905>. Currículo: <http://lattes.cnpq.br/4726949697973381>

1 Introdução

Os dados são considerados o novo petróleo na era digital, possuindo um valor altamente estratégico. Tão importante quanto o petróleo, os dados são a arma mais poderosa da sociedade. Grandes empresas tecnológicas como Amazon, Apple, Facebook, Google e Microsoft apresentaram um faturamento maior que 25 bilhões de dólares apenas no primeiro semestre de 2017, demonstrando que as organizações que detêm os dados também detêm o poder (ECONOMIST, 2017).

A produção e armazenamento de dados têm passado por um processo evolutivo desde a popularização dos computadores e a difusão da internet, dessa forma, tem ganhando destaque a partir de uma produção massiva de dados realizada por meio de sites, redes sociais digitais, smartphones, sistemas informatizados, satélites, sensores de carros, aviões e trens, câmeras de segurança entre outros tipos de serviços como streaming (AMARAL, 2016). Além da evolução referente a capacidade de armazenamento de dados, faz necessário destacar também o quantitativo de vezes que os dados são compartilhados pelos usuários (MARQUESONE, 2016).

O conceito de Big Data foi cunhado a partir do crescimento e integração de grandes volumes de dados digitais produzidos por diversos meios e tecnologias, bem como pelo uso de ferramentas que perpassam as etapas de produção, compartilhamento, curadoria e análise de maneira eficiente dessa gigantesca quantidade de dados, resultando em informações de valia e permitindo descobrir novos padrões significativos de conhecimentos (SAKR, 2016; ELSHAWI; SAKR, 2018).

Buscas realizadas utilizando o termo Big Data nas ferramentas de pesquisa acadêmica Google Scholar e o Portal de Periódicos da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) apresentou como resultado, respectivamente, 4.820.000 e 1.739.128 documentos que abordam de alguma forma a temática Big Data. Já a pesquisa realizada no buscador Google apresentou aproximadamente 4.170.000.000 resultados. Assim, trata-se de um quantitativo bastante representativo, uma vez que se refere a um termo que foi cunhado há poucos anos.

A definição de Big Data é considerada pobre (BOYD; CRAWFORD, 2012), equivocada e genérica, pois refere-se apenas ao tamanho (DAVENPORT, 2017). Contudo, este trabalho parte do princípio que o Big Data é um termo em fase de amadurecimento e que ainda está em processo de consolidação de conceitos.

Justifica-se a importância desse trabalho pela notoriedade do termo Big Data nos últimos anos, tanto no meio acadêmico quanto empresarial e pela discrepância entre os conceitos encontrados na literatura. Como hipótese, acredita-se que, por se tratar de um termo relativamente novo e principalmente por estar atrelado aos meios e avanços tecnológicos, possam surgir cotidianamente novas características que contribuem para essa diversidade de conceitos.

2 Objetivos

O trabalho tem como objetivo identificar os conceitos relacionados ao termo Big Data e, identificar as características que compõem o conceito.

3 Procedimentos Metodológicos

Esta é uma pesquisa de natureza básica, abordando o problema de forma qualitativa. Do ponto de vista dos seus objetivos, é classificada como exploratória e descritiva. Da perspectiva dos procedimentos técnicos, identifica-se como pesquisa bibliográfica (GIL, 2010; Gil, 2016).

A pesquisa consiste em um levantamento, análise e seleção de trabalhos que trazem uma conceituação do termo Big Data, bem como suas características, publicados entre os anos de 2011 e 2018.

Os trabalhos foram identificados a partir de pesquisas realizadas no Portal de Periódicos da CAPES e no Google Scholar. Além desses, foram utilizados artigos científicos, livros digitais e impressos da área de domínio.

As publicações selecionadas serviram como base para a identificação dos conceitos relacionados ao termo Big Data, bem como a identificação das características que compõem o seu conceito.

4 Resultados

Um número expressivo de documentos acadêmicos apresenta uma série de conceitos sobre o termo Big Data. Esses conceitos são apresentados no Apêndice A. Verifica-se dessa forma, uma diversidade de conceitos o que pode contribuir para a não consolidação do termo.

A consultoria Gartner desenvolveu um estudo no ano de 2001 no qual apresentou o volume, a variedade e a velocidade – conhecidos como 3V's – como um dos maiores desafios para a manipulação e a gestão de dados. Esse viés tem sido discutido com maior relevância no meio acadêmico (Laney, 2001). Posteriormente, a veracidade (qualidade) valor (relevância) foram incorporadas aos 3V's e constituíram assim o 5V's. Essas características têm ganhado destaque nas investigações relacionadas ao meio empresarial (Assunção et al., 2015). Tais características são utilizadas para definir o termo Big Data.

Contudo, além dos 5V's, outras características podem surgir, formando assim os n-V's, uma vez que o conceito de Big Data está atrelado aos avanços tecnológicos.

4.1 Big Data e suas características

O Big Data é um fenômeno tecnológico e humano no qual são produzidos dados em diversos formatos e que são armazenados em grandes quantidades de dispositivos de armazenamento. Refere-se às características inerentes de um Big Data os n-V's, sendo eles:

- Volume – relaciona-se ao quantitativo de informações armazenadas em dispositivos físicos ou virtualizados. No atual cenário, os armazenamentos de dados perpassam por um crescimento exponencial mediante a tecnologia de computação nas nuvens. Os dados têm sido criados em escala de zetabytes por meio de diversas fontes e aplicações por atividades cotidianas (VASCONCELOS; BARÃO, 2017). Considerando que bilhões de pessoas geram informações diariamente por meio da internet, estima-se que o volume de informações produzidas possa dobrar a cada 18 meses. O atributo volume é considerado a característica mais importante no conceito de Big Data e faz dimensão sem precedentes do volume de dados (MACHADO, 2018);

- Variedade – define a diversidade de informações e dados produzidos e que possuem formatos heterogênicos, podendo ser dados estruturados, não estruturados ou semiestruturados que são criados, compartilhados e consumidos em diferentes formatos e uma variedade de fontes de dados, como, por exemplo, upload de fotos, transmissão de vídeos em tempo real, textos publicados em redes sociais (MACHADO, 2018). A variedade também indica a variação das áreas de aplicações de um Big Data, tais como as áreas governamentais, financeiras, transportes e automação, varejistas, seguros e outras possibilidades (VASCONCELOS; BARÃO, 2017);
- Velocidade – além dos desafios referentes ao volume e variedade de dados, o Big Data faz referência a agilidade em que os dados são coletados, processados, analisados e utilizados (MARQUESONE, 2016; VASCONCELOS; BARÃO, 2017; MACHADO, 2018). Um fator que contribuiu positivamente para a característica da velocidade está na evolução das telecomunicações, tais como largura de banda e infraestrutura física, possibilitando assim que os dados sejam transportados com uma maior velocidade por meio da internet (VASCONCELOS; BARÃO, 2017);
- Veracidade – diz respeito a qualidade dos dados e a confiabilidade sobre as fontes dos dados. O Big Data está inserido em um contexto de variedade e volume de dados, sendo comum a existência de dados inconsistentes. Dessa forma, torna-se necessário garantir a autenticidade das fontes e dos dados coletados de forma que sejam confiáveis para uma determinada solução (MARQUESONE, 2016; VASCONCELOS; BARÃO, 2017; MACHADO, 2018);
- Valor – caracteriza-se pelo quão valioso e significativo um dado pode ser para uma determinada solução em um Big Data, obtendo assim valor qualitativo a partir do volume, variedade e velocidade de dados (MARQUESONE, 2016). Torna-se importante ressaltar que se trata da relevância dos dados após o processamento. Os dados brutos perpassam por análise e processamento onde poderão ser acrescidos de valor, transformando-se em vantagem competitiva para negócios. (VASCONCELOS; BARÃO, 2017; MACHADO, 2018).
- Viabilidade – destina-se a identificar o relacionamento entre variáveis e padrões latentes de dados (BIEHN, 2013). Esse conceito pode ser discutido, por exemplo, em modelagem de tópicos que faz uso de machine learning e processamento de linguagem natural em grandes corpora textuais para identificar a emergência dos tópicos latentes.
- Venalidade – projetado futuramente em formatos que possam ser vendidos (DAVENPORT, 2017). A questão da venda de dados já ocorre por meio de empresas prestadoras de serviços online, entretanto, o autor discute a questão de um formato específico para o conceito de grandes volumes de dados.

5 Considerações Finais

O termo Big Data, novo se comparado a outros conceitos científicos consolidados na literatura, apresenta uma diversidade de conceitos, contribuindo para a não consolidação do termo. Esse fato pode ser observado no Apêndice A que mostra uma não uniformidade entre as

definições elaboradas pelos diversos autores, inclusive com críticas, disparidades, simplicidade e homogeneidade.

Outro fator que contribui para falta de consolidação do termo Big Data está nas suas características. Inicialmente conhecidas por 3V's – volume, variedade e velocidade; mais tarde com o acréscimo de veracidade e viabilidade tornando-se 5V's; porém, as características que compõem o conceito de Big Data ultrapassam esse quantitativo, podendo ser acrescentado também viabilidade e venalidade, tornando-se assim 7V's. Essa evolução de características de palavras iniciadas com a letra “V” tem acontecido mediante ao surgimento de novas necessidades que aparecem ao se realizar algum tipo de processamento com Big Data e também por conta do avanço tecnológico. Entretanto, acredita-se que as tecnologias estejam longe de serem estabilizadas. Com isso, novas possibilidades ou novos “V's” poderão surgir nos próximos anos, como por exemplo, a Variabilidade – capacidade de submeter-se a variações de mudanças; e Visualização de dados associados ao Big Data. Percebe-se que as características associadas ao conceito de um Big Data no que diz respeito aos 5V's tem se apresentado de forma sólida na literatura, entretanto, com o avanço tecnológico, novas possibilidades têm surgido e colocado as características em cheque. Dessa maneira, surge o que podemos chamar de n-V's, um paradigma ainda em construção, aberto para as novas características ou representações do termo Big Data.

Ressalta-se que, em busca realizada no Portal de Periódicos da CAPES utilizando os termos “n-V's Big Data”, “nV's Big Data”, “n- Vs Big Data” e “nVs Big Data”, apesar da quantidade elevada de resultados que tratam do Big Data, não foram encontrados resultados que abordem os n-V's da forma que foi proposta neste trabalho – o paradigma dos n- V's.

Sugere-se como pesquisas futuras uma quebra de paradigmas com relação à caracterização real do termo de forma que o mesmo não esteja atrelado à determinada letra alfabética, e sim ao que de fato representa um Big Data. Além disso, perspectivas futuras sobre o Big Data e tecnologias podem contribuir para a formação de um conceito consolidado, embora acredita-se que possa levar anos para realização de uma definição universalizada do conceito.

Referências

AMARAL, F. **Introdução à ciência de dados: mineração de dados e big data**. Rio de Janeiro: Alta Books, 2016.

ASSUNÇÃO, M. D. et al. Big Data computing and clouds: trends and future directions. **Journal of Parallel and Distributed Computing**, v. 79, p. 3-15, 2015.

BIEHN, N. The missing V's in big data: viability and value. **Wired**, New York, 2013. Disponível em: <<https://www.wired.com/insights/2013/05/the-missing-vs-in-big-data-viability-and-value>>. Acesso em: 10 set. 2018.

BOYD, D.; CRAWFORD, K. Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon. **Information, Communication & Society**, v. 15, n. 5, p. 663, 2012.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. **Portal de Periódicos**. Disponível em: <<http://www.periodicos.capes.gov.br/>>. Acesso em: 05 set. 2018.

CRAWFORD, K.; SCHULTZ, J. Big Data and Due Process: toward a Framework to Redress Predictive Privacy Harms. **Boston College Law Review**, v. 55, n. 1, p. 93-128, 2014. Disponível em: <<http://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4>>. Acesso em: 11 set. 2018.

DAVENPORT, T. H. **Big Data no trabalho: derrubando mitos e descobrindo oportunidades**. Rio de Janeiro: Alta Books, 2017.

ECONOMIST. **The world's most valuable resource: data and the new rules of competition**. London, v. 423, n. 9039, 2017.

ELSHAWI, R.; SAKR, S. Big Data Systems Meet Machine Learning Challenges: towards Big Data Science as a Service. **Big Data Research**, v. 5, n. 10, p. 7, 2018. Disponível em: <<https://arxiv.org/pdf/1709.07493.pdf>>. Acesso em: 5 set. 2018.

FRICKÉ, M. Big data and its epistemology. **Journal of the Association for Information Science and Technology**, v. 66, n. 4, p. 651-661, 2015.

GANTZ, J.; REINSEL, D. Extracting value from chaos. **IDC Review**, v. 1142, n. 2011, p. 1-12, 2011. Disponível em: <<https://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>>. Acesso em: 09 set. 2018.

BIG DATA. In: IT GLOSSARY. Stamford: Gartner Group, 2012. Disponível em: <<https://www.gartner.com/it-glossary/big-data/>>. Acesso em: 09 set. 2018.

GEORGE, G.; HAAS, M. R.; PENTLAND, A. Big Data and Management. **Academy of Management Journal**, 2014, Vol. 57, No. 2, 321–326. Disponível em: <<http://dx.doi.org/10.5465/amj.2014.4002>>. Acesso em: 07 set. 2018.

GIL, A. C. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

_____. **Métodos e técnicas de Pesquisa Social**. 6. ed. São Paulo: Atlas, 2016.

GOMES, R. D. P. **Big Data: desafios à tutela da pessoa humana na sociedade da informação**. Rio de Janeiro: Lumen Juris, 2017.

GOOGLE. Disponível em: <<http://www.google.com/>>. Acesso em: 05 set. 2018.

GOOGLE SCHOLAR. Disponível em: <<https://scholar.google.com.br/>>. Acesso em: 05 set. 2018.

HIJMANS, H. **The european union as guardian of internet privacy**. Brussels: Springer International Publishing, 2016.

INTERNATIONAL BUSINESS MACHINES (IBM). O que é Big Data, 2014. Disponível em: <<https://i.pinimg.com/originals/f6/5c/fd/f65cfdfe4813b82155e9a35eacf590f4.jpg>>. Acesso em: 10 set. 2018.

INSTITUTO DE TECNOLOGIA & SOCIEDADE DO RIO (ITSRio). Big Data no projeto Sul Global: relatório sobre estudos de caso. Rio de Janeiro: p.26, 2016. Disponível em: <https://itsrio.org/wp-content/uploads/2016/03/ITS_Relatorio_Big-Data_PT-BR_v2.pdf>. Acesso em: 9 set. 2018.

LANEY, D. 3D data management: controlling data volume, velocity and variety. **META Delta, Stamford**, 6 fev. 2001. Application Delivery Strategies. Disponível em: <<https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>>. Acesso em 9 set. 2018.

MACHADO, F. N. R. **Big Data**: o futuro dos dados e aplicações. São Paulo: Érica, 2018.

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

MARQUESONE, R. **Big Data**: técnicas e tecnologias para extração de valor dos dados. São Paulo: Casa do Código, 2016.

PROVOST, F.; FAWCETT, T. **Data Science para negócios**: o que você precisa saber sobre mineração de dados e pensamento analítico de dados. Rio de Janeiro: Alta Books, 2016.

SAKR, S. **Big data 2.0 processing systems**: a survey. Springer International Publishing, 2016.

TAURION, C. **Big data**. Brasport, 2013.

VASCONCELOS, J. B.; BARÃO, A. **Ciência dos dados nas organizações**: aplicações em python. Lisboa: FCA, 2017.

VOLPATO, T; RUFINO, R. R.; DIAS, J. W. **Big data – transformando dados em decisões**. 2014. Disponível em: <http://web.unipar.br/~seinpar/2014/artigos/graduacao/Tiago_Volpato.pdf>; Acesso em: 9 set. 2018.

Apêndice A – Conceitos de Big Data

| Referência | Conceitos |
|-----------------------------------|--|
| (GANTZ; REINSEL, 2011) | “[...] as tecnologias de Big Data descrevem uma nova geração de tecnologias e arquiteturas projetadas para extrair economicamente o valor de volumes muito grandes e uma variedade de dados, permitindo alta velocidade de captura, descoberta, e/ou análise.” |
| (BOYD; CRAWFORD, 2012) | “Big Data é, em vários sentidos, um termo pobre.” “Definimos big data como um fenômeno cultural, tecnológico e acadêmico.” “Big Data é menos sobre dados qu |
| (GARTNER GROUP, 2012) | “Big Data são ativos de informações de grande volume, alta velocidade e / ou de alta variedade que exigem formas inovadoras e econômicas de processamento de informações que permitem uma melhor percepção, tomada de decisões e automação de processos.” |
| (EUROPEAN COMMISSION, 2013) | “Big Data se refere ao crescimento exponencial tanto da disponibilidade quanto no uso automatizado de informação: refere-se a conjuntos de dados digitais gigantescos detidos por empresas, governos e outras organizações de grande porte, que são amplamente analisados (daí o nome: analytics) usando algoritmos de computador. Big Data pode ser usado para identificar tendências mais gerais e correlações, mas também pode ser processado, de modo a afetar diretamente os indivíduos.” |
| (MAYER-SCHÖNBERGER; CUKIER, 2013) | “Big Data refere-se a coisas que se podem fazer em grande escala, que não podem ser feitas em escala menor, de forma a extrair novas ideias ou criar novas formas de valor, de maneira que acabam mudando mercados, organizações, a relação entre cidadãos e os governos, dentre outros.” |
| (TAURION, 2013) | “Big Data = volume + variedade + velocidade.” |
| (CRAWFORD; SCHULTZ, 2014) | “Big Data é um termo generalizado e impreciso, que se refere ao uso de grandes conjuntos de dados na ciência de dados e análise preditiva.” |
| (GEORGE et. al.2014) | “O Big data é formado por uma crescente pluralidade de fontes de informação, entre eles cliques na web, transações em dispositivos móveis, conteúdo gerado por usuários, mídias sociais, bem como conteúdo gerado intencionalmente através de redes de sensores ou transações comerciais, tais como consultas de vendas e transações de compra.” |
| (IBM, 2014) | “Big Data é um termo utilizado para descrever grandes volumes de informações de dados e que ganha cada vez mais relevância à medida que a sociedade se depara com o aumento sem precedentes no número de informações.” |
| (FRICKÉ (2015) | “Big Data não é incompatível com a experimentação. Mas é o amigo da observação passiva: encoraja a observação passiva. Realizar pesquisas e fazer observações é uma observação passiva.” “Big Data deve ser capaz de nos fornecer melhores evidências para nossas teorias.” |
| (AMARAL, 2016) | “Big Data é o fenômeno em que dados são produzidos em vários formatos e armazenados por uma grande quantidade de dispositivos e equipamentos.” |
| (HIJMANS, 2016) | “A evolução da era big data implica, por sua própria natureza, em falta de controle, uma vez que o volume de dados é sem precedentes, diverso em variedade e movendo-se com uma velocidade que se aproxima cada vez mais do tempo real.” |
| (ITSRio, 2016) | “[...] o conjunto de dados cuja existência só é possível em consequência da coleta massiva de dados que se tornou possível nos últimos anos, graças à onipresença de aparelhos e sensores na vida cotidiana e do número crescente de pessoas conectadas a tais tecnologias por meio de redes digitais e também de sensores.” |
| (PROVOST; FAWCETT, 2016) | “[...] o termo big data significa conjunto de dados que são grandes demais para os sistemas tradicionais de processamento e, portanto, exigem novas tecnologias para processá-los” |
| (DAVENPORT, 2017) | “O big data é inegavelmente grande, mas o termo é um tanto ou quanto equivocado.” |
| (MACHADO, 2018) | “Este grande volume de dados, esta explosão de dados gerados universalmente a cada instante, é chamado de Big Data, que está expondo uma nova onda de tecnologia e arquitetura destinada a extrair valor de uma imensa variedade de dados, o que permite alta velocidade com o objetivo de capturar, descobrir e analisar estas informações importantes e valiosas no âmbito de gestão de negócios.” |

Fonte: elaborado pelo autor com base em: (VOLPATO; RUFINO; DIAS, 2014; GOMES, 2017).

Vídeo da apresentação

Título: O termo Big Data: quebra de paradigma dos n-V's.



Disponível em: http://dadosabertos.info/enhanced_publications/idt/video.php?id=6

Transcrição da apresentação

Olá essa é uma apresentação referente ao trabalho aprovado junto ao 2º Workshop de Informação Dados e Tecnologia – WIDAT, que nessa edição de 2018 acontece na Universidade Federal da Paraíba. É uma pesquisa realizada com contribuição dos pesquisadores Fernando Gomes Almeida da Universidade Federal de Minas Gerais, Renato Rocha Souza também da UFMG e da Fundação Getúlio Vargas do Rio de Janeiro e por mim Marcos de Souza da Universidade Federal de Minas Gerais. O título da pesquisa é “O TERMO BIG DATA: quebra de paradigma dos n-V’s”.

Trata-se de uma pesquisa de referencial teórico, mas como muita contribuição no que diz respeito especificamente dos conceitos e das características de um Big Data. Pois bem para a gente poder falar de Big Data a gente tem que falar de dados, uma matéria que saiu no ano de 2017 no “The Economist” diz que os dados são considerados um novo petróleo na era digital, possuindo um valor altamente estratégico, nessa matéria destacam por exemplo as organizações que detém os dados também detém todo o poder. Ainda mais falando na questão econômica, alguns exemplos, Apple, facebook, Google e a Microsoft tiveram um faturamento de 25 bilhões de dólares apenas no primeiro semestre de 2017. Quer dizer, no ano passado é um faturamento consideravelmente muito alto, empresas que trabalham com dados.

Toda essa questão de produção e armazenamento de dados tem passado por um processo evolutivo desde a popularização dos computadores e da difusão da internet, desta forma pegar destaque a partir de uma produção massiva de dados realizada, por exemplo, por sites, por redes sociais digitais, por smartphones, sistemas informatizados, satélites, sensores de carros, aviões e trens, câmeras de segurança entre outros tipos de serviços como, por exemplo, stream que é esse agora que vocês estão vendo neste momento.

Além da evolução referente à capacidade de armazenamento de dados a autora Rosângela Marquesone, ela destaca que faz necessário ressaltar a quantidade de vezes que esses mesmos dados são compartilhados pelos próprios usuários. Partindo deste princípio de dados a gente consegue chegar ao conceito de Big Data cunhado a partir do crescimento e integração do grande volume de dados digitais como esses excitados agora há pouco, por exemplos diversos meios de tecnologias bem como o uso de ferramentas que perpassa pela por algumas etapas sendo elas produção, compartilhamento, curadoria e análise de uma maneira eficiente nessa gigantesca quantidade de dados.

Como resultado nessas informações elas geram, por exemplo, padrões significativo de novos conhecimentos. A Gartner desenvolveu um estudo no ano de 2001 no qual apresentou três características de um Big Data, sendo elas volume, variedade e velocidade, conhecido na literatura como 3-v’s, como um dos maiores desafios da manipulação e gestão de dados, esse viés tem sido discutido massivamente com muita relevância no meio acadêmico. Posteriormente surgiram mais dois v’s que é a “veracidade” relacionada à qualidade e o “valor” relacionado à relevância, formando assim ou incorporando assim os 5 v’s também na literatura, só que com a veracidade e com valor está voltado mais para um viés empresarial. Então essas cinco características, elas contribuem para uma definição de Big Data.

Justifica-se a importância desse trabalho pela notoriedade do termo de Big Data nos últimos anos, tanto no meio acadêmico quanto no meio empresarial e a discrepância no qual você encontra os conceitos na literatura, não existe uma homogeneidade de um conceito de Big Data e como hipótese acredita-se que por se tratar de um termo relativamente novo, principalmente por estar atrelado aos meios e avanços tecnológicos, pode surgir forma cotidianas novas características que contribuem com a diversidade desse conceito.

O objetivo dessa pesquisa está em identificar os conceitos relacionados ao termo Big Data e as características que compunham este conceito, então a definição de Big Data por exemplo por alguns autores é muito criticado, então por exemplo Boyd e Crawford falam que o tema Big Data é considerado muito pop, é equivocado e genérico por Davenport em um livro de 2017 recém-lançado. O conteúdo desse trabalho ele parte do princípio que Big Data é um termo em fase de amadurecimento e que ainda está em processo de consolidação do seu conceito.

Trata-se de uma pesquisa de natureza básica abordando o problema de forma qualitativa do ponto de vista dos seus objetivos a pesquisa se classifica como exploratória e descritiva, da perspectiva dos procedimentos técnicos identifica-se como pesquisa bibliográfica, lógico que estamos nos apoiando em nomes de autores como Gil por exemplo. A pesquisa consiste em um levantamento de análise e seleção de trabalhos que faz a conceituação de Big Data, bem como suas características publicadas entre o período de 2011 e 2018, esse período temporal nós encontramos diversos conceitos que de fato não existe uma homogeneidade do conceito em cima e sim uma discrepância, hoje apresentado no apêndice desta pesquisa.

Os trabalhos foram identificados através de pesquisas realizadas no portal de periódicos da CAPES, no Google Acadêmico além de livros que foram utilizados e artigos científicos, livros digitais, impressos na área de domínio de Big Data. As publicações selecionadas serviram como base para a identificação dos conceitos relacionados ao termo Big Data e com a identificação das características que compõem este conceito.

As características já encontrada massivamente na literatura, nós temos o volume, variedade, velocidade nos três v's, focado na área acadêmica a veracidade e valor voltado para o meio empresarial, entretanto essa pesquisa trouxe alguns resultados muito interessantes, por exemplo mais dois v's que a viabilidade e a venalidade. Além disso outro diferencial dessa pesquisa é que o termo que nós utilizamos para o título "n-v's" não foi encontrada na base de periódicos da CAPES, algumas variações mesmo assim foram zero resultados encontrados.

Do conceito de Big Data, a gente pode falar que é um fenômeno tecnológico e humano no qual são formados dados em diversos formatos que são armazenadas em grande quantidade de dados em dispositivos de armazenamento, refere-se assim as características inerentes de Big Data os n-v's, em que a literatura hoje fala em 5-v's, nós já estamos identificando sete e possam surgir novos v's, novas características que serão apresentados.

Desses novos 2-v's nós vamos destacar aqui a "viabilidade" que destina-se a investigar um relacionamento entre as variáveis e padrões latente de dados, é uma citação de Biehn de 2013, o que seria isso é padrões latentes de dados na viabilidade pode ser utilizado, por exemplo, há um conceito de modelagem de tópicos, faz uso de machine learning, por exemplo, em cima de processamento de linguagem natural, por exemplo eu tenho grandes corpora de dados e aí eu

quero identificar a viabilidade por exemplo da imersão desses tópicos, qual é a frequência desses tópicos, a relação desses tópicos, em cima de um grande Big Data? Então eu trato com relação a isso essa parte latente que é de viabilidade.

Davenport em 2017 a questão da legalidade, projeção futuramente de formatos de Big Data que possam ser vendidos. A gente pode falar assim: ah mas hoje dados são vendidos, por exemplo se eu adquirir um número novo de uma linha telefônica no outro dia já tem alguém oferecendo algum determinado tipo de produto, a venda de fato realmente acontece, mas o que Davenport trabalha, que ele está pleiteando futuramente ou almejando é a questão de um formato de venda de dados.

Considerações finais o termo Big Data comparado a outros conceitos científicos já consolidados na literatura apresenta uma diversidade de conceitos e contribui para uma não consolidação do termo, este fato pode ser observado junto da pesquisa certamente terão acesso aos anais, no apêndice dela nós construímos uma linha temporal de 2011 até 2018 onde vocês poderão observar, por exemplo, uma não uniformidade entre as definições elaboradas pelos autores, inclusive com críticas, com disparidade de conceito e a não homogeneidade das características do conceito.

Essa evolução de características de palavras iniciadas com a letra “v” tem acontecido mediante o surgimento de novas necessidades que aparece para realizar o processamento do Big Data, também por conta dos avanços tecnológicos, isso que as tecnologias estejam longe de serem estabilizadas e com isso novas possibilidades de novas “v’s” poderão surgir nos próximos anos, como por exemplo a variabilidade que diz respeito à capacidade de submeter-se a variações mudança e a própria visualização de dados que já existem mas que pode ser conceituada, associada a um Big Data, então a gente está falando e 9-v’s, n-v’s de acordo com as necessidades de processamento e avanços tecnológicos vão surgir “n” características para tentar formar um conceito sólido de Big Data.

Sugere-se como pesquisas futuras, quebra de paradigmas com a caracterização real do termo. Por fim só mais 2-v’s que são apresentados potencializando mais 2-v’s chegando assim a 9-v’s, para poder caracterizar Big Data, são características que acabam sendo descaracterizados, não sei, seria uma questão mais de discussão. Fica a mesma dica para pesquisas futuras, tenhamos de fato um conceito real, quanto tempo isso vai levar eu não sei responder, mas de toda forma nós agradecemos a contribuição realizada para o evento e todas as críticas, sugestões vocês podem entrar em contato conosco.

Obrigado.