

BIG DATA E SEGURANÇA*

Antonio Wilson Nogueira Filgueiras – UEMG – Unidade Carangola
Daniela Martins Sampaio – UEMG – Unidade Carangola
Harlen Pereira Ferreira – UEMG – Unidade Carangola
Nilson Ned Corrêa – UEMG – Unidade Carangola
Jádison Leite Brandão – UEMG – Unidade Carangola

RESUMO: Esse artigo tem como objetivo realizar um estudo sobre o Big Data, fenômeno premente do mundo moderno causado pelo excesso de dados da sociedade da informação. Sabe-se que na sociedade atual e com os grandes avanços da tecnologia houve uma crescente quantidade de dados digitais, segundo a IBM, em 2008, foram produzidos cerca de 2,5 quintilhões de bytes todos os dias, surpreendentemente 90% dos dados do mundo foram criados nos últimos dois anos decorrentes as adesões das grandes empresas à internet, como exemplo, as redes sociais, dados de GPS, dispositivos embutidos e móveis. Uma das barreiras encontradas na expansão de dados é a falta de armazenamento. Sendo assim necessária se faz a criação de Data Center específico para o armazenamento em massa de dados, uma vez que o Big Data vem para suprir, de certa forma, essa necessidade de armazenamento com o conceito de oferecer, agregar e tirar conclusões de dados até então espaços.

PALAVRAS-CHAVE: Big Data. Armazenamento. Dados. Tecnologia. Privacidade.

INTRODUÇÃO

Postagens no twitter, o envio de um e-mail, o preenchimento de um formulário eletrônico, uso de cartões de crédito em transações comerciais, transações bancárias e ligações telefônicas geram dados que vão sendo guardados, acumulando e um enorme volume de dados diversificados. Segundo a Revista Ciênciahoje (Agosto/2013), esse volume de dados não é mais gerado apenas por grandes centros tecnológicos, governamentais e empresas como antes acontecia. Nos dias atuais, esse volume de informações criado, conta com a participação dos usuários domésticos no mundo todo que utilizam de equipamentos eletrônicos presentes desde a camada mais privilegiada até a camada popular, produzindo informações, como por exemplo, fotos tiradas a partir de câmeras digitais, o envio de SMS, arquivos multimídia produzidos por celulares e dispositivos móveis entre outros equipamentos eletrônicos.

Segundo o IDC (2011, 2012), foram gerados 1,7 Zettabytes e 2,7 Zettabytes de dados digitais respectivamente e ainda prevê que em 2015 o volume produzido anualmente alcance 8 Zettabytes. Para entender melhor as proporções acerca desse crescimento, temos que 90% dos dados disponíveis no mundo foram gerados apenas durante 2010 e 2011. Segundo a divulgação da IBM (International Business Machines) no final de 2012.

O grande problema em meio a essa massa de dados crescente é como armazená-los e tratá-los, de maneira que eles possam ser úteis e o que isso implicará para a sociedade. Portanto,

¹*XI EVIDOSOL e VIII CILTEC-Online - junho/2014 –<http://evidosol.textolivre.org>

estudos e pesquisas tem se voltado para essa área e dado origem ao termo Big Data. O objetivo desse artigo é realizar um estudo sobre o tema, para que possamos conhecer melhor esse fenômeno premente do mundo moderno. Para tanto, a metodologia utilizada foi a da pesquisa bibliográfica e na internet. Averiguar as implicações que o uso do Big Data tem em relação à segurança de dados e, nesse caso, a privacidade da informação é um fator a ser discutido. Embora o uso do Big Data possua seu aspecto positivo, também se faz necessária a discussão dos contras que ele pode vir a oferecer.

1 SURGE UM NOVO TERMO NO SÉCULO XXI: BIG DATA

Enfrenta-se um grande problema em relação ao armazenamento de dados, desde um documento importante, fotos pessoais, arquivos multimídia até diversidades de dados irrelevantes que, mesmo assim, ocupam espaço. Segundo a revista InfoEscola,

O conceito de *Big Data* é focado no gigantesco armazenamento de dados, com enorme velocidade. O Big Data é baseado no conceito de 5V: Valor, Veracidade, Variedade, Volume e Velocidade. O Big Data é tida por muitos, como a solução de eventuais situações problemáticas da economia. E devido ao modelo econômico adotado pela globalização, o termo *Just in time*, totalmente dependente da necessidade de uma expansão virtual, se tornou a palavra de ordem das negociações e forçou a ampliação de estrutura para armazenamento de dados.(ADAMI, 2014, p.1)

Devido à grande quantidade de dados pessoais fornecidos voluntariamente na rede, como os dados fornecidos em redes sociais, e-mails, lojas eletrônicas, o termo Big Data surge como um conjunto de estudos com o objetivo de suprir a necessidade de armazenamento e gerenciamento desse gigantesco volume de dados. O termo Big Data se popularizou a partir de 2013 e não há um conceito específico para ele, mas para muitos cientistas a principal idéia de Big Data é unir uma infinidade de informações não estruturadas que ao serem usadas de maneira consciente se tornam poderosas na tomada de decisões cada vez mais rápidas e precisas. Diante disso, há o consenso de que três dessas características, também chamadas de três ‘V’s’ são os principais objetivos que se espera alcançar com o Big Data: Volume, Velocidade e Variedade (Ciênciahoje, p. 19). A proposta é tratar uma quantidade massiva de dados diversificados em uma grande velocidade de modo que se possa utilizá-los instantaneamente.

Através da análise de dados não estruturados, como, posts no facebook, consegue-se descobrir para qual time determinada pessoa torce e a partir desse ponto, os anunciantes podem enviar propagandas de produtos relacionados ao time favorito do usuário, diminuindo os riscos de se cometer uma gafe com o consumidor. Através da análise de dados é possível chegar ao conhecimento de interesses pessoais baseando-se em informações fornecidas pelos próprios usuários e assim, anúncios são direcionados diretamente a esses mesmos usuários que, conseqüentemente, se tornam o público alvo de campanhas publicitárias.

Porém, o Big Data não se aplica somente ao marketing de vendas. Um bom exemplo é o “Flu Trends” do Google, que é baseado nos dados de seu buscador. A empresa desenvolveu um projeto que consegue identificar e analisar a informações de modo que identifique tendências, repetições, perfis de usuários como é o caso de um usuário pesquisando sobre gripe, o programa é capaz de prever a propagação antes que os números oficiais reflitam a situação real e com esse mesmo programa é possível identificar e analisar a inflação, a taxa de desemprego, as variações da

Bolsa, entre outros.

Outro exemplo a ser citado para mostrar o uso dessas grandes massas de dados é a eleição presidencial nos EUA no ano de 2013. Segundo o portal de notícias G1, os funcionários contratados para auxiliar na campanha de Obama utilizaram de grandes massas de informações como telefonemas, mensagens nas redes sociais, dados dos perfis dos eleitores para que pudessem prever e reverter o resultado ao favor do candidato. O resultado final foi certo - Obama foi reeleito.

2 PRIVACIDADE E BIG DATA

Sabe-se que com o crescimento desordenado do número de dados digitais, a privacidade, sem dúvida, será uma questão muito importante a se discutir. Um dos fatores que influenciam muito no crescimento do número de dados são as redes sociais, pessoas postando fotos, vídeos, mensagens textuais, expondo sua vida pessoal na internet, facilitando o acesso para aqueles que se interessam em descobrir assuntos pessoais.

Segundo o portal de notícias G1, a cada cinco segundos, no mundo inteiro, 17 mil fotos e 144 mil posts são publicados no Facebook e por dia, são 247 milhões de e-mails. Dados de usuário são armazenados cada vez mais na internet e com a tecnologia de armazenamento nas nuvens, o processo tem ficado ainda mais fácil e tem sido tentador para os usuários suas informações na rede. Fotos e arquivos particulares, dados pessoais estão na internet, facilitando uma análise total da vida de uma pessoa. Existem casos em que as empresas antes de contratar um novo funcionário, analisam o perfil do candidato nas redes sociais.

Para Fábiana Barros, gerente de projetos da Across, consultoria de desenvolvimento organizacional com foco em gestão de pessoas, as mídias sociais tornaram horizontal a relação entre empresas e profissionais. A triagem de currículos por meio das redes sociais é cada vez mais comum, destaca. Muito vantajoso para as empresas, com base nas informações, seria mais fácil analisar o perfil do candidato a uma vaga, um processo que não é complicado de fazer. Seria necessário apenas entrar em uma rede social e acompanhar todo o histórico do candidato e, com base na análise dos dados, fica muito fácil traçar um perfil completo do candidato e, definindo se esse candidato estaria apto ou não a trabalhar na empresa de acordo com os próprios critérios da mesma.

Mas, em contrapartida, estão os usuários que têm sua privacidade violada porque tudo que é postado pode ser visto por qualquer pessoa do mundo e fatos isolados poderiam gerar conclusões erradas a respeito de determinada pessoa. Por isso, os termos dos três 'Vs' (volume, velocidade e variedade) propostos no Big Data são estendidos por diversos autores a cinco 'Vs', em que são incluídas a veracidade e a variabilidade ou o valor dos dados que trata da importância em verificar se o dado é verdadeiro e se ele é válido no tempo.

Outro grande assunto a ser discutido é a segurança e privacidade dos dados em rede. Confiar que dados estão extremamente seguros é uma questão duvidosa, há riscos de uso indevido dessas informações. Confiar que seus dados digitais estão armazenados em algum lugar do mundo, em um servidor sabe-se lá onde? este está fisicamente instalado, pode gerar receio e dúvidas em relação à segurança. Seria uma difícil decisão para o dono de uma grande empresa decidir em curto prazo, se é ou não viável hospedar seus dados em rede sabendo-se do risco do vazamento dessas informações. O risco de vazamento de informações sigilosas e estratégicas a empresas rivais, por exemplo, seria um grande prejuízo dessa empresa em relação aos concorrentes causando problemas, muitas vezes, irreversíveis, pois dados sendo mantidos nos equipamentos de outras empresas (Os

servidores) como, por exemplo, Google, Yahoo, Outlook, pode-se estar à mercê de quaisquer medidas de quebra de privacidade uma vez que este servidor tem acesso aos dados hospedados em seus centros de armazenamento, os Data Centers.

Segundo A Convergência digital, no mercado há uma grande carência de profissionais no ramo de segurança digital. Com uma massa de dados invadindo os servidores, e dados cada vez mais pessoais espalhados, é indispensável à discussão acerca da garantia de segurança de dados na rede. Big Data, indiscutivelmente aumenta o risco de brechas na privacidade, uma vez que todas as informações não estruturadas estão à disposição. Esse este risco é realmente preocupante ou não, eles estão à solta.

CONCLUSÃO

Os desafios pelo Big Data em relação aos dados podem ser divididos em como adquirir, armazenar, processar, transmitir e disseminar, gerenciar e manter, arquivar por longo prazo e garantir a segurança e o mais relevante de todos: treinar pessoas para usá-los. A utilização do Big Data é tão promissora que a IBM criou a Big Data University, com o objetivo de formar o novo profissional denominado o cientista de dados, com responsabilidades de estudar matérias relacionadas ao assunto e assim, se tornar apto a trabalhar com esse sistema. Para isso, novas tecnologias e conceitos fazem-se necessários no mundo moderno. Tecnologias que supram necessidades atuais e ao mesmo tempo sejam viáveis em meio à grande invasão em massa de dados que vêm ocorrendo e o crescimento exponencial previsto para os próximos anos. Preocupar-se de que forma essas informações serão usadas também é um conceito a se levar a discussão.

Percebeu-se por este estudo que o acesso a essa grande quantidade de informações pode acarretar a invasão de informações privilegiadas sem que o próprio usuário esteja ciente do ocorrido e isso pode nos levar a questionar a diferença entre público e privado, dando margem a novas pesquisas.

REFERÊNCIAS

CHEDE, Cezar. Você sabe realmente o que é Big Data? Disponível em: <https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/voce_realmente_sabe_o_que_e_big_data?lang=en>. Acessado em 10/03/2014.

ADAMI, Ana. “Big Data”. Revista InfoEscola. Disponível em: <<http://www.infoescola.com/informatica/big-data/>>. Acessado em 19/03/2014.

SANTANA, Otávio Gonçalves de. Introdução de Big Data. Disponível em <<http://www.devmedia.com.br/introducao-ao-conceito-de-big-data/27066>>. Acessado em 15/03/2014

TAUFER, Patrícia. Massa de informações digitais pode ser usada em benefício da população. Disponível em: <<http://g1.globo.com/jornal-da-globo/noticia/2013/12/massa-de-informacoes-digitais-pode-ser-usada-em-beneficio-da-populacao.html>>. Acessado em 17/03/2014

Vídeo. Conheça as aplicações reais do Big Data. Disponível em: <<http://olhardigital.uol.com.br/pro/video/39376/39376>>. Acessado em 17/03/2014

XEXEO, Geraldo. *Big Data: Computação para uma sociedade conectada e digitalizada*. Cienciahoje, vol. 306, 2013, p. 18-23.