

A INDIVIDUALIZAÇÃO DO CONSUMIDOR POR MEIO DO AVANÇO DAS REDES SOCIAIS E DO BIG DATA

CONSUMER IDENTIFICATION THROUGH SOCIAL NETWORKS AND BIG DATA

Rafael Fernandes Novo¹ (rfnovo@hotmail.com)
Marília Macorin de Azevedo² (marilia.azevedo@fatec.sp.gov.br)

^{1,2} FATEC – São Paulo

Resumo

Os sistemas de big data possuem o potencial de coletar e analisar grandes volumes de dados com diferentes especificidades, como padrões de navegação na internet e postagens em redes sociais. Esse contexto torna pertinente a discussão das implicações éticas do uso de tais informações por corporações e do discernimento dos usuários ao publicar informações em redes sociais. Este artigo discute metodologia e resultados de uma pesquisa de natureza descritiva e exploratória que produziu um estudo qualitativo sobre conscientização e percepção de privacidade de usuários do Facebook. Como resultado, foi possível identificar superexposição dos usuários e potencial utilidade dos dados para empresas e consumidores, ressaltando a necessidade de mais debate da questão ética a partir de aperfeiçoamentos no preparo das empresas e conscientização dos usuários.

Palavras-chave: Big data; Redes sociais; Informação; Ética; Consumidor.

Abstrac

Big data systems have the potential to collect and analyze large volumes of data with different features, such as internet browsing patterns and posts on social networks. Such context makes relevant the discussion on possible ethical issues on the use of this bulk of information by corporations and on users' judgment to publish information on social networks. This paper shows methodology and results of a descriptive piece of research whose objective was to conduct a qualitative study of awareness and perception of privacy in Facebook users. Excessive user exposure in the network was detected, as well as usefulness of data for businesses and consumers, as long as the ethical aspects are better discussed than presently, with more corporation preparedness and user awareness.

Keywords: Big data, Social Networks, Information, Ethics, Consumer.

1 Introdução

Redes sociais são definidas por Boyd (2008) como sistemas nos quais o usuário pode criar um perfil público ou semipúblico, manter uma lista de contatos com os quais estabelece diferentes tipos de conexão ou relação e acessar ou visualizar os contatos dos seus contatos, criando assim a sua rede de relacionamento. Com o advento destes sistemas, cresce o número de indivíduos que faz uso das redes sociais para expressar suas opiniões, comunicar e postar diversos tipos de informações. Estes dados tornam-se então disponíveis para outros indivíduos ou empresas que, por meio do uso de tecnologias como o big data, organizam e analisam quantidades massivas de dados, gerando informações relevantes sobre os usuários.

Torna-se possível a individualização do consumidor, gerando uma potencial vantagem para a empresa que passa a conhecer melhor seu mercado, além de beneficiar o próprio

consumidor, que pode receber ofertas direcionadas a suas necessidades. Porém este novo cenário abre espaço para uma importante discussão ética sobre o uso destes dados e sobre o risco ao qual as pessoas estão se expondo ao perderem o controle ou influência sobre o uso de informações pessoais.

O objetivo de qualquer empresa é dar lucro, contudo Murgel (2006), lembra que o caminho para o lucro passa pela satisfação do cliente. Para isso é necessário oferecer produtos e serviços relevantes ao cliente e conhecer a sua opinião de forma a fechar continuamente o ciclo com um processo ininterrupto de melhoria da oferta. A própria forma de ouvir o cliente está sujeita a uma avaliação ética. Ainda segundo Murgel (2006), a ética pode ser usada para definir ações certas ou erradas, o que certamente se aplica a relações de empresas com os seus clientes. A Ética existe em um plano distinto da legislação. Uma empresa pode atuar dentro da lei, porém tomar decisões que violem a ética.

2 Desenvolvimento

As redes sociais vêm crescendo e hoje estão presentes de modo significativo no cotidiano das pessoas. No ano de 2011, por exemplo, constatou-se que 13% dos norte-americanos adultos usavam o Twitter regularmente (SMITH, 2011). Outra evidência da crescente importância do conteúdo que circula por estas redes foi a decisão em 2010 da Biblioteca do Congresso Americano em arquivar e disponibilizar todas as postagens dos usuários do Twitter (LOC, 2013).

Trata-se de fenômeno global e em franca expansão. O LinkedIn alegou ter em março de 2012 mais de 160 milhões de usuários ativos, crescendo a uma taxa de 2 novos membros por segundo, e esse ainda é um número pequeno perto dos 900 milhões de usuários ativos do Facebook (AQUINO, 2012).

Um dos fatores de crescimento significativo no uso da internet e redes sociais é a proliferação de dispositivos móveis, como smartphones, que permitem o acesso a partir de qualquer lugar que possua sinal de telefone celular. A Figura 1, mostra que em alguns casos, mais da metade dos usuários dos principais sites da internet realizaram acessos via smartphones. Krishnamurthy (2010) lembra que algumas redes sociais, como o Foursquare, foram criadas especificamente para serem acessadas de dispositivos móveis.

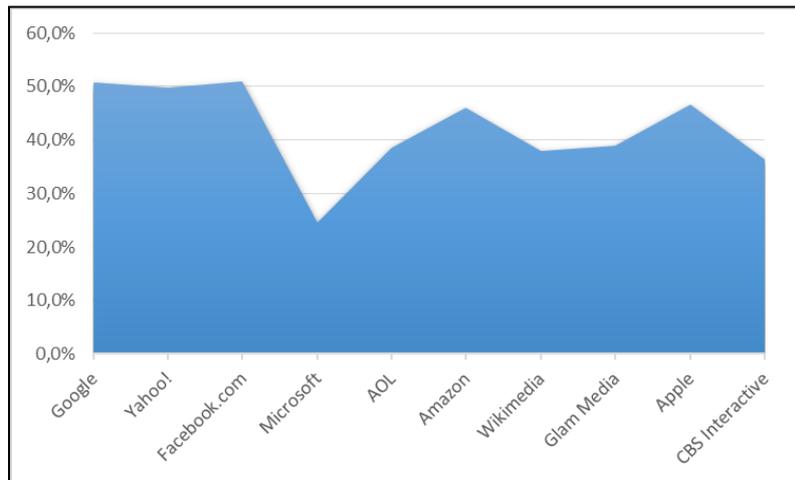


Figura 1 – % de usuários de *smartphones* (Set 2012)
Fonte: Adaptado de COMSCORE (2012)

Dados extraídos de redes sociais como LinkedIn ou Twitter podem trazer metadados (dados que dizem respeito a outros dados) que incluem diversas informações adicionais, como por exemplo a geolocalização dos usuários, como demonstrado em pesquisa representada na Figura 2, que correlaciona a localidade do usuário com o conteúdo postado por este no Twitter para identificar onde estão concentrados os usuários que postarem conteúdo com a hashtag (palavra-chave) #bigdata. Esses dados podem ser utilizados em uma campanha de marketing, permitindo uma abordagem mais segura para as empresas, por direcionamento da propaganda para mercado consumidor específico.



Figura 2 – Amostragem de local de origem de tweets com a hashtag #bigdata
Fonte: Os autores

Um dos exemplos dos riscos e exposição por meio do compartilhamento de informação nas redes sociais podia ser observado em <http://icanstalku.com/>. O site atualmente não está mais em operação, mas deixou um relato de como, por meio da análise dos metadados de fotos postadas no Twitter, foram informações sobre os usuários, como:

- Onde vivem;

Segundo Krishnamurthy (2009), tem aumentado o número de empresas que executam a função de agregadoras. Tais empresas coletam informações dos cookies armazenadas nos computadores dos usuários, executam processos de análises de dados e fornecem informações relevantes sobre comportamento e preferência dos consumidores para empresas de publicidade, marketing ou de qualquer outro ramo.

Geralmente, a única informação vinculada ao usuário gerador do cookie é o IP ou o navegador que este está utilizando. Desta forma, após buscar na internet por um novo produto como um refrigerador ou uma viagem de férias, o usuário passa a receber em outros sites que visita ou mesmo redes sociais que frequenta anúncios direcionados ao tema. Essa informação não é específica do indivíduo, ou seja, não existe uma relação entre o cookie e o nome, e-mail ou endereço residencial do usuário.

Aquino (2012) revela ainda que o uso de cookies ou redes sociais para direcionar propagandas não é vista por muitos como problema, pois é comum que usuários estejam dispostos a receber anúncios relevantes durante o uso das redes sociais. Os usuários entendem ser esta uma contrapartida pelo uso gratuito dos serviços destes sites e, como relatado por Tucker (2013), pode ser até interessante para o consumidor ter anúncios feitos sobre medida para a sua necessidade atual. Existe no entanto um receio desses usuários em serem rastreados por empresas de publicidade, sem que fique clara a extensão do monitoramento ou de como os dados são utilizados. É, portanto, necessária cautela por parte das empresas ao usarem dados de clientes provenientes das redes sociais. Esse conflito fica demonstrado em pesquisa realizada por Borba (2013) que constatou que 45% dos entrevistados acreditam que essa monitoração e análise de dados pode ser útil para eles, por exemplo, na geração de buscas mais eficientes na Internet, enquanto 43% afirmaram que percebem isso como uma invasão de privacidade.

No entanto, como ressaltado por Krishnamurthy (2009), o uso de cookies de navegação no contexto das redes sociais, considerando o poder computacional de soluções de big data, pode trazer uma consequência indesejada pelos usuários, a geração de informações pessoais identificáveis (IPIs), ou seja, o estabelecimento de vínculos entre informações coletadas pelos cookies (como sites visitados, padrão de navegação e interesses) com as informações da sua rede social (como nome, email, amigos, endereço).

Tene (2012) alerta que apesar de existirem políticas de privacidade nas redes sociais, inclusive apresentadas para os usuários no momento de inscrição nos serviços como Facebook ou LinkedIn, estas muitas vezes servem apenas para isentar de responsabilidade as empresas, e não como balizadores das garantias de privacidade. Adicionalmente, as redes sociais deixam disponíveis vários dados pessoais acessáveis por qualquer indivíduo ou corporação, como demonstrado pela Figura 4.

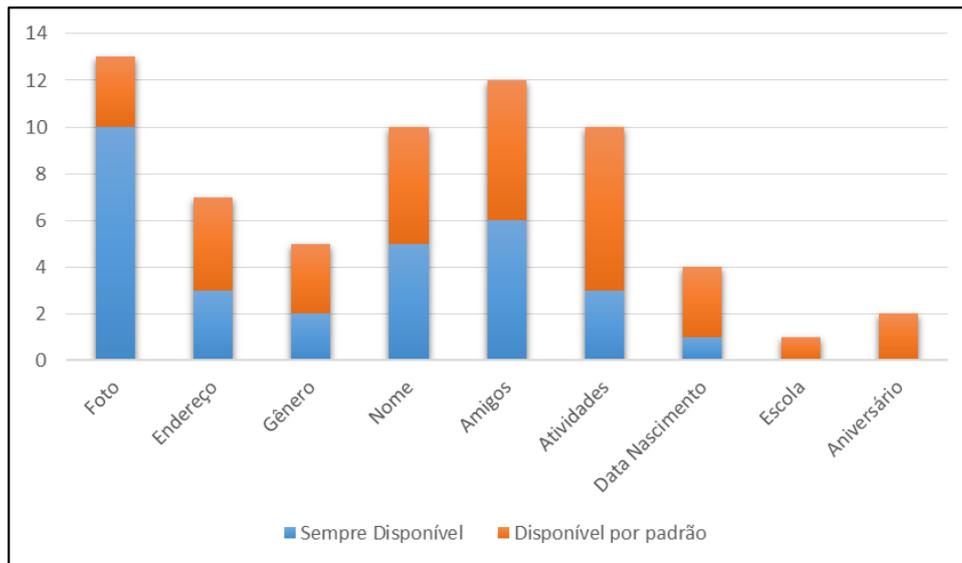


Figura 4 – Informações disponíveis em 13 redes sociais investigadas
Fonte: Adaptado de Krishnamurthy (2010)

Diante da natureza dinâmica da informação digital em um mundo altamente conectado, tornaram-se necessárias regulamentações globais que norteiem o uso da internet entre países. Por exemplo, um usuário do Facebook no Brasil tem os seus dados armazenados em servidores localizados nos Estados Unidos e portanto sujeitos à legislação local. Uma empresa europeia que utilize estes dados obedece a um outro grupo de normas que podem ser diferentes das leis brasileiras e americanas. Modelos de regulamentações como a dos oceanos (National Marine Sanctuary Act – NMSA, Endangered Species Act – ESA e outros) podem servir como base conceitual para a criação de normas globais para a internet. Além disso, é necessário um órgão mundial responsável por este controle, nos moldes da Organização Mundial de Comércio (THE ECONOMIST, 2010).

Torna-se cada vez mais difícil ocultar ou restringir totalmente os rastros digitais de cada indivíduo. Kosinskia (2012), ao escrever sobre as redes sociais demonstrou que é possível inferir com elevado grau de exatidão, mesmo quando o usuário tiver optado por não informar dados pessoais como sexo ou situação social, essas e outras informações privadas dos indivíduos, como orientação sexual, etnia, religião, orientação política, inteligência, uso de drogas, divórcio dos pais ou idade. Para levantar esses dados não é necessário analisar qualquer informação postada pelo próprio indivíduo no Facebook. Basta utilizar os *likes* (curtidas) destes indivíduos e análise dos textos que outras pessoas postaram, e o que essas outras pessoas curtiram. Esse estudo foi feito com dados de mais de 58 mil voluntários que consentiram na análise dos seus *likes* e preencheram um questionário para comparação dos resultados. Algumas das características analisadas, como etnia, obtiveram mais de 95% de acerto. Vale ressaltar que ainda não existe opção de tornar privados ou ocultar os *likes* do Facebook.

Kosinskia (2012) alerta para a grande quantidade de dados disponíveis online oriundos deste novo modelo de interação social e prevê que isto trará discussões sobre privacidade e

propriedade dos dados, na medida em que torna possível para governos, empresas ou mesmo outras pessoas inferir informações privadas e não intencionalmente divulgadas usando análises de dados.

Atualmente não existe um entendimento consolidado de que os dados gerados por um usuário, como informações postadas em redes sociais, sejam de propriedade do indivíduo. Isso cria abertura para as empresas usarem estes dados de forma ampla e irrestrita. Dentro deste contexto, a revista *The Economist* (2010) alerta para o fato de que as pessoas estão disponibilizando informações potencialmente confidenciais nas redes sociais e isso ocasionará uma tensão entre empresas que querem utilizar tais dados em benefícios próprio e consumidores que querem proteger suas informações pessoais.

Por outro lado, a restrição total aos dados tampouco é uma solução. Além do benefício mútuo já citado entre empresa e consumidor, outros exemplos de benefícios extraídos da análise de dados gerados por indivíduos também serão severamente afetados, como nos exemplos trazidos por Tene (2012):

- O remédio Vioxx foi banido do mercado depois de que seu uso foi relacionado com 27.000 mortes por parada cardíaca entre os anos de 1999 e 2003.
- Os padrões de propagação do vírus da influenza foram mapeados e foi possível prover uma resposta mais rápida de amparo à população.

É imperativo haver um balanço entre a privacidade à qual todos tem direito e os potenciais benefícios do uso destas informações, como questões de saúde e segurança ou coleta de dados estatísticos para entender padrões de serviços, permitindo a entrega de produtos de melhor qualidade ao próprio consumidor (TENE, 2012).

Levará algum tempo para que as organizações entendam completamente o que é viável em relação a capturar, gerenciar, analisar e, acima de tudo, criar inteligência social, a fim de agir sobre o conteúdo gerado pelo usuário (LESLIE, 2012, p. 9).

2.1 Método

A pesquisa descrita a seguir investiga o problema ético existente no uso das informações contidas em redes sociais e mede a conscientização dos usuários do Facebook sobre o uso destas informações postadas nas redes sociais. Com o objetivo de aprofundar a investigação foram incluídas fontes secundárias. A pesquisa é mista: qualitativa por não ter o objetivo de provar uma asserção e sim investigar um contexto, e quantitativa na apresentação dos resultados.

Os dados foram coletados entre usuários do Facebook, rede social mais utilizada por brasileiros, como mostra a Figura 5.

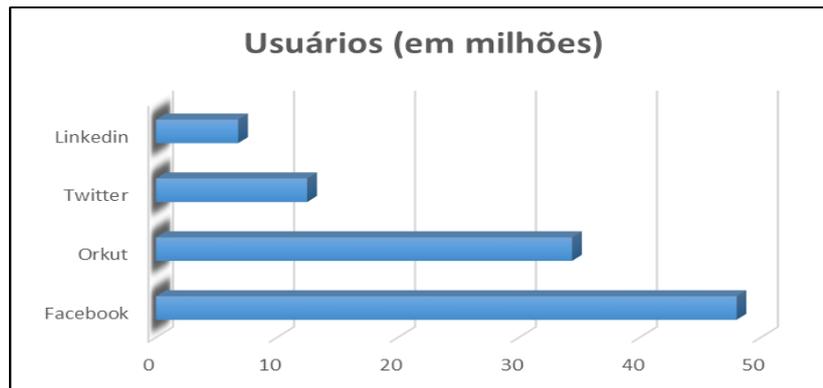


Figura 5 - Redes sociais com maior número de usuários no Brasil
Fonte: Adaptado de VOLGARIN (2012)

O formulário mostrado no Quadro 1 foi enviado para todos os contatos do pesquisador e retransmitidos a suas conexões e conexões delas, formando assim uma cadeia de comunicação de usuários da rede com algum grau de aleatoriedade – foi originada dos contatos de uma professora universitária e pesquisadora, mas não seguiu qualquer regra de propagação.

Quadro 1 – Formulário de pesquisa

Identificação	Qual o nome do seu perfil no Facebook? Qual o seu e-mail?
Uso e conhecimento das políticas do Facebook	Você leu as políticas de privacidade do Facebook? Você alterou as configurações de privacidade padrão do Facebook? Você utiliza algum critério ao postar conteúdo (fotos/mensagens/vídeos) no Facebook? Quais destes itens você considera informação pessoal?
Percepção sobre uso dos dados	O que você acha de empresas que fazem propaganda de produtos e serviços direcionadas a você com base nas suas publicações do Facebook? O que você acha de empresas que usam as informações disponíveis no Facebook para definir seu modo de interação com você como, por exemplo, calcular o seguro do seu carro? Você considera que suas informações estão expostas no Facebook?

Elaborado pelos autores com base em dados consultados

A coleta de dados ocorreu no período de setembro a novembro de 2013, com quase 50 respostas.

O número de respostas representa um percentual muito reduzido dos usuários do Facebook, o que limita a pesquisa a demonstrar tendências de comportamento de um grupo de usuários, sem massa crítica válida para uma análise estatística.

2.2 Resultados e Discussões

As políticas de privacidade das redes sociais, possuem um direcionamento predominantemente jurídico, com objetivo de prover salvaguardas e não de serem um instrumento prático de proteção e conscientização do consumidor. No estudo realizado, 70% dos entrevistados não chegaram sequer a tentar ler a política de privacidade, demonstrando indiferença com relação à relevância dela para o uso da rede social. Do total, apenas 11% declararam efetivamente

entender e concordar com as políticas de privacidade. Entre os fatores que podem ter influenciado este resultado estão o tamanho do texto, a complexidade da linguagem e a impossibilidade de alterar qualquer cláusula do documento.

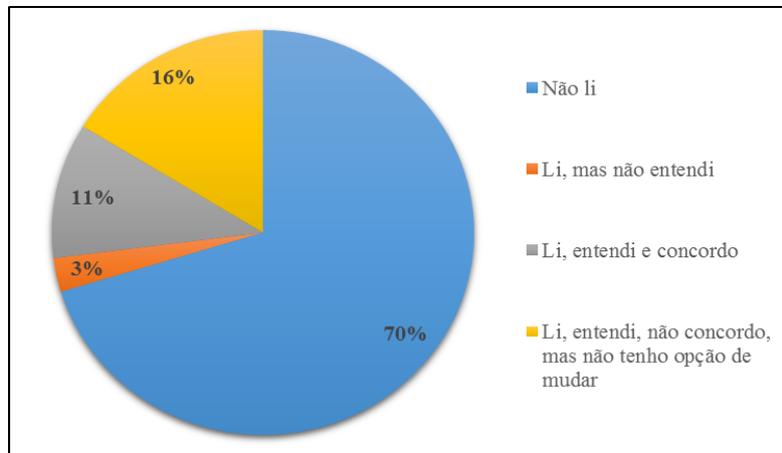


Figura 6 – Usuários e as políticas de privacidade do Facebook
Fonte: Os autores

Outro ponto relevante, apontado por Krishnamurthy (2010), é a configuração de segurança e privacidade das redes sociais, que por padrão expõe várias informações dos usuários, conforme visto na Figura 4.

No caso do Facebook é possível configurar algumas restrições básicas com relação a segurança e privacidade do usuário. Averiguou-se que a maioria dos usuários (87%) alega ter realizado ao menos alterações básicas na configuração default.

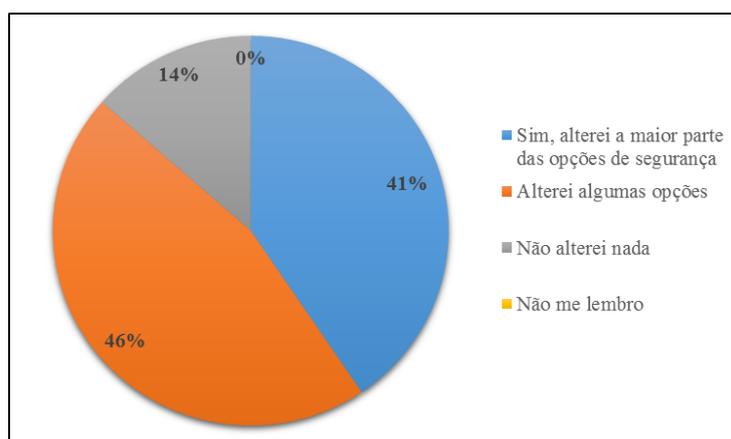


Figura 7 – Usuários e as configurações de segurança e privacidade do Facebook
Fonte: Os autores

Ao analisar os perfis dos usuários que participaram da pesquisa investigou-se se eles haviam restringido o acesso de desconhecidos a duas informações pessoais (e-mail e postagens). A compilação dos resultados revelou que 68% dos usuários analisados deixaram as postagens disponíveis para todos os demais usuários da rede social, e 54% deixaram o endereço de e-mail exposto. Essas duas informações pessoais básicas que estavam disponíveis na maioria das contas dos usuários podem representar uma falta de conhecimento técnico dos usuários nas

configurações de segurança do Facebook ou uma percepção por parte dos usuários que a exposição pública destas informações não constitui um risco.

Verificou-se também que 46% dos usuários alegaram postar qualquer conteúdo que julguem de interesse dos amigos conectados à rede social. Em contrapartida, apenas 32% afirmaram usar como critério para postagens a cautela quanto a exposição de informações pessoais.

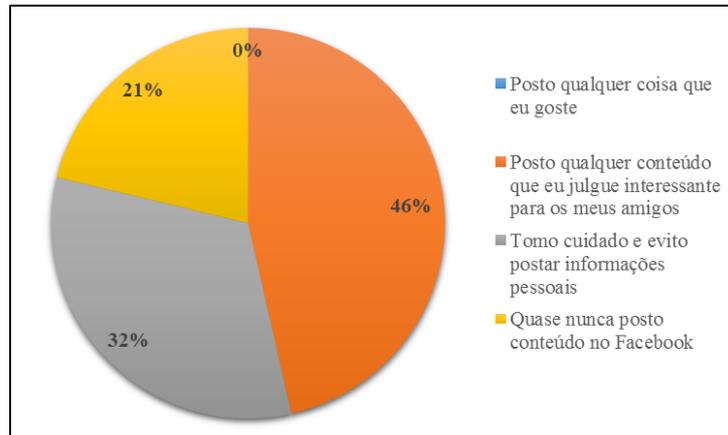


Figura 8 – Critérios ao postar conteúdos no Facebook
Fonte: Os autores

É relevante notar que mesmo entre o grupo de usuários que alegou não ter como principal critério evitar postagens com informações pessoais, a maioria (54%) discorda da utilização desses dados por empresas para propaganda ou para qualquer outro uso.

Ao se questionar sobre o uso de postagens do Facebook para fins que não a propaganda, observou-se que apenas 38% são tolerantes ao uso de seus dados para propaganda e 62% discordam da utilização das informações postadas para outros fins (sem que esses outros fins fossem explicitados).

3 Considerações Finais

Dados como cookies ou postagens em redes sociais constituem fontes de informação disponível aos sistemas de big data e podem ser utilizados por empresas para que estas obtenham maior conhecimento de mercados-alvo. O consumidor poderá ser beneficiado ao receber ofertas melhor ajustadas às suas necessidades, mas verifica-se, mesmo em uma pesquisa limitada, um usuário das redes sociais desinformado sobre os possíveis usos das informações disponibilizadas. Surge a questão ética no uso desses dados, que precisa ser debatida com maior profundidade com o objetivo de garantir o respeito ao consumidor. É necessário que o indivíduo tenha maior controle sobre quem pode acessar suas informações e qual o uso que as empresas estão fazendo destes dados, mas o tema ainda carece de regulamentações que estabeleçam padrões de identificação e de segurança dos dados pessoais.

Esta pesquisa é um estudo inicial sobre o tema, necessitando ainda uma quantidade maior de respostas coletadas para permitir estudos que apresentem resultados estatisticamente

significativos. Por exemplo, para uma margem de erro de 2% e nível de confiança de 95% será necessário entrevistar cerca de 2500 usuários, com cuidados em relação a uma distribuição de originadores representativa de níveis de renda e instrução.

Referências

- AQUINO, J. Transforming social media data into predictive analytics. **Destination CRM**. 2012
- BORBA, D. K; SANTOS, L. F. A; KAWAMOTO JR, L. T. **Big Data: percepção dos usuários sobre vantagens e invasão de privacidade**. VIII Workshop de pós-graduação e pesquisa do Centro Paula Souza. São Paulo, 2013
- BOYD, D; ELLISON, N. **Social Network Sites: Definition, History, and Scholarship**. Journal of Computer-Mediated Communication, 2008
- KRISHNAMURTHY, B; WILLS, C. E. **On the leakage of personally identifiable information via online social networks**. Proceedings of the 2nd ACM workshop on Online social networks. ACM, 2009
- _____. **Privacy leakage in mobile online social networks**. Proceedings of the 3rd conference on Online social networks. USENIX Association, 2010
- LESLIE, A. **Social Analytics & Intelligence: Converting Contextual to Actionable Insights**. Hypatia Research & Advisory, 2012
- LOC. **Library of Congress. Update on the Twitter Archive at the Library of Congress**. Janeiro, 2013. Disponível em: <http://www.loc.govtoday/pr/2013/files/twitter_report_2013jan.pdf>. Acesso em 10 outubro de 2013.
- COMSCORE Announces Beta Release of Media Metrix® Multi-Platform, Providing a Unified View of Web, Smartphone and Tablet Audiences. Novembro de 2012. Disponível em: <http://www.comscore.com/Insights/Press_Releases/2012/11/comScore_Announces_Beta_Release_of_Media_Metrix_Multi-Platform>. Acesso em 10 outubro de 2013.
- MURGEL, D; SILVA, J; NEVES, J. **A ética nos negócios como diferencial competitivo**. ENEGEP, 2006
- SMITH, A. **13% of online adults use Twitter**. Pew Research Center, 2011
- TENE, O; POLONETSKY, J. **Privacy in the age of big data: a time for big decisions**. Stanford Law Review Online, 2012
- TUCKER, W.; THALER, R. H. **Informação mais inteligente, consumidor mais forte**. Harvard Business Review, 2013
- VOLGARIN, R; et al. **A aplicação de redes sociais no contexto mercadológico**. VIII Workshop de pós-graduação e pesquisa do Centro Paula Souza, 2013.

Recebido em 07/03/2014

Aceito em 16/08/2014