

A PERCEÇÃO DE VULNERABILIDADE E APLICAÇÃO ÉTICA DAS INFORMAÇÕES NAS REDES SOCIAIS PELOS SISTEMAS DE BIG DATA

VULNERABILITY PERCEPTION AND ETHICAL APPLICATION OF INFORMATION IN SOCIAL NETWORKS USING BIG DATA SYSTEMS

Rafael Novo¹ Marilia Macorin de Azevedo²

RESUMO

Sistema de *Big Data* possui o potencial de coletar e analisar grandes volumes de dados nos mais variados formatos, como padrões de navegação na Internet e postagens em redes sociais. Neste contexto, torna-se pertinente a discussão das possíveis implicações éticas do uso destas informações por corporações e da maturidade dos usuários ao publicarem informações em redes sociais. A presente pesquisa objetiva descrever e explorar como empresas podem conhecer melhor seus consumidores por meio da análise de dados postados em redes sociais e realizar um estudo qualitativo sobre a percepção de privacidade de usuários do Facebook no uso destas informações postadas. Foi possível identificar a existência de uma exposição dos usuários nesta rede e que o uso destes dados pode ser proveitoso para empresas e consumidores, considerando que a questão ética seja corretamente debatida, com preparo das empresas e maior conscientização dos usuários.

Palavras-chave: Análise de dados. Big Data. Ética. Informação. Redes Sociais.

ABSTRACT

Big Data system has the potential to collect and analyze large volumes of data in different formats such as browsing on the Internet and posts on social networks. In this context, it is relevant to discuss the possible ethical implications of information usage by corporations as well as the maturity of the users to publish information on social networks. This descriptive research intends to conduct a qualitative study of awareness and perception of privacy for Facebook users. As a result, it was identified that there is an exposure of users in this network and that the use of these data can be useful for businesses and consumers, as long as ethical aspects are properly discussed with more prepared corporations and more user awareness.

Keywords: Analytics. Big Data. Ethics. Information. Social Networks.

¹Graduando Centro Estadual de Educação Tecnológica Paula Souza, CEETEPS, São Paulo

²Doutora em Engenharia, Professora do Centro Estadual de Educação Tecnológica Paula Souza, CEETEPS, São Paulo.e-mail: marilia.azevedo@fatec.sp.gov.br

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 ele 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 nestes sites os mais diversos tipos de informações. Outro fator responsável por este aumento é a proliferação de dispositivos móveis, como *smartphones*, que permitem o acesso dos usuários a partir de qualquer lugar. Esses dados tornam-se então disponíveis para que outros indivíduos ou empresas façam a extração e análise destes por meio do *Big Data* - um conjunto de sistemas que possibilitam a análise de textos em quantidades massivas, gerando informações relevantes sobre os usuários das redes sociais.

Torna-se possível a individualização do consumidor, gerando uma potencial vantagem para a empresa ao conhecer melhor seu mercado, além de beneficiar o próprio consumidor, o qual pode receber ofertas direcionadas às suas necessidades. Porém este novo cenário abre espaço para uma importante discussão ética sobre o uso destes dados e sobre o risco a que as pessoas estão se expondo, ao perderem o controle ou influência sobre o uso destes.

No âmbito empresarial, temos o lucro como objetivo final de qualquer empresa, contudo Murgel (2006) lembra que o caminho para o lucro passa pela satisfação do cliente. Sendo necessário para isso, oferecer produtos e serviços relevantes ao cliente e ouvir a sua opinião de forma a fechar o ciclo com um processo contínuo de melhoria empresarial. Entretanto, a própria forma de ouvir o cliente também está sujeita a uma avaliação ética. Ainda segundo Murgel (2006) ética pode ser usada para definir ações certas ou erradas, o que se aplica às relações, por exemplo, de empresas com os seus clientes. Ética existe em um plano distinto da legislação. Uma empresa pode atuar dentro da lei, porém tomar decisões que violem a ética.

Neste cenário, é essencial que haja um debate sobre a segurança que as empresas oferecem sobre esses dados e a privacidade destes dados, por exemplo, das próprias redes sociais comentadas aqui – que apesar de serem gerados e disponibilizados por indivíduos que, em certo modo, optaram por tornar essas dados públicos, certamente não o fizeram com o intuito de serem usados por corporações.

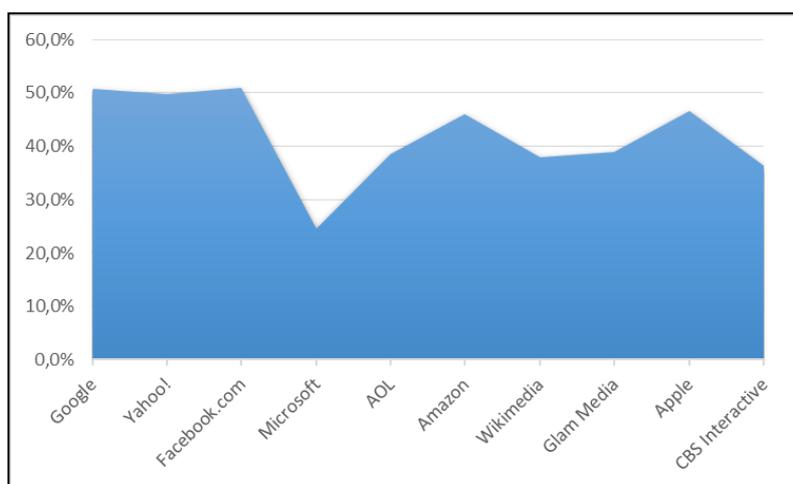
2 MATERIAL E MÉTODOS

As redes sociais vêm crescendo e hoje possuem elevada significância no cotidiano das pessoas. No ano de 2011, constatou-se que 13% dos norte-americanos adultos usavam o Twitter regularmente (SMITH, 2011). A decisão em 2010 da Biblioteca do Congresso Norte Americano em arquivar e disponibilizar todas as postagens dos usuários na rede social Twitter é outra evidência da crescente importância das redes sociais e seu conteúdo (LOC, 2013).

Esse é um fenômeno global e em franca expansão. A rede social exclusiva para contatos profissionais, chamada 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 este ainda é um número pequeno perto dos 900 milhões de usuários ativos do Facebook (AQUINO, 2012).

Um dos fatores de aumento 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. 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, dentro deste novo contexto, algumas redes sociais foram criadas especificamente para serem acessadas de dispositivos móveis, como o Foursquare, que tem como objetivo permitir que os usuários compartilhem experiências sobre lugares visitados.

Figura 1 – Percentagem de usuários de smartphones (Set. 2012)



Fonte: Adaptado de COMSCORE (2012)

Dados extraídos de redes sociais como LinkedIn ou Twitter podem trazer metadados que incluem diversas informações adicionais sobre os usuários, como por exemplo a geolocalização dos usuários, como demonstrada 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 postaram conteúdo com a *hashtag* (palavra-chave) #bigdata. Esses dados podem ser utilizados como mais um fator de decisão em uma campanha de marketing, permitindo uma abordagem mais assertiva para as empresas, direcionando a propaganda para o mercado consumidor específico.

Figura 2 – Amostragem de local de origem de *tweets* com a *hashtag* #bigdata



Fonte: Autor

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

- Local onde vivem;
- Outras pessoas vivendo no mesmo local;
- Padrões de deslocamento;
- Local que almoçam todos os dias;
- Com quem almoçam;
- Opções de entretenimento;

(Internet Explorer, Firefox, Chrome, etc) aceite armazenar cookies. Porém, segundo o autor, tornam-se mais frequentes as entidades (empresas) que executam a função de agregadores. 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 propaganda, marketing ou qualquer outra empresa que queira contratar seus serviços.

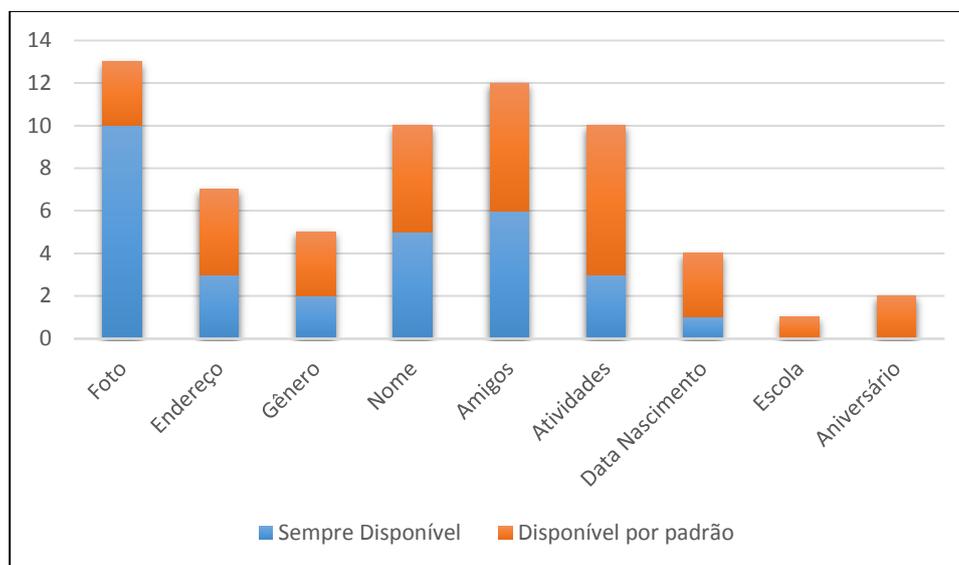
Geralmente, a única informação vinculada ao usuário gerador do cookie é o endereço de rede do computador que este está utilizando. Desta forma, após buscar na Internet por um novo produto como um refrigerador ou uma viagem de férias, o usuário passa a receber em outros sites que visita ou até mesmo nas redes sociais que frequenta anúncios direcionados sobre o tema. Entretanto, esta informação não é individualizada, 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 um problema, pois muitos usuários estão 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. Contudo existe um receio destes usuários em serem rastreados por empresas de propagandas, sem a compreensão completa de quão extensa é essa monitoração ou como estes dados são utilizados; por isso é necessário 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 assertivas na Internet, enquanto 43% afirmaram que percebem isso como uma invasão de privacidade.

Entretanto, como ressaltado por Krishnamurthy (2009), o uso de cookies de navegação, juntamente com as redes sociais e o poder computacional de soluções de *Big Data*, pode trazer novas consequências para os usuários, como 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, e-mail, amizades, 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 como isenção de responsabilidade para as empresas, em vez de balizadores das garantias de privacidade para os consumidores. Adicionalmente as redes sociais deixam disponível uma série de dados pessoais, os quais podem ser acessados por qualquer indivíduo ou corporação com acesso à rede, como demonstrado na 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, dar-se-á a necessidade de regulamentações globais que norteiem o uso da Internet entre países. Por exemplo, um usuário do Brasil que usa o Facebook tem os seus dados armazenados em servidores localizados nos Estados Unidos e, portanto, sujeitos a uma série de legislações locais. Uma empresa europeia que utilize estes dados obedece a outro grupo de normas que podem ser dissonantes das leis brasileiras e estadunidenses. É necessário um órgão mundial responsável por este controle, como a Organização Mundial de Comércio (THE ECONOMIST 2010).

Torna-se cada vez mais difícil ocultar ou restringir totalmente os rastros digitais de cada indivíduo. Focando no cerne desta pesquisa, as redes sociais, mesmo quando é feita a opção por não informar dados pessoais como sexo ou situação social, Kosinskia (2012) demonstrou que é possível inferir com elevado grau de assertividade estas e outras informações privadas dos indivíduos, tais como: orientação sexual; etnia;

religião; orientação política; inteligência; estado de espírito; uso de drogas; divórcio dos pais; idade e sexo. Para levantar esses dados, não foi analisada nenhuma informação postada pelo próprio indivíduo no Facebook. Utilizou-se apenas os *likes* ou curtidas destes indivíduos e análise dos textos que outras pessoas postaram, e que esses indivíduos curtiram. Este estudo foi feito com 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 privado 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 no tangente à privacidade e propriedade dos dados, na medida em que torna possível para governos, empresas ou até mesmo outras pessoas inferir informações privadas e não divulgadas dos indivíduos, 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, o *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 benefício próprio e os consumidores que querem proteger as suas informações pessoais.

É imperativo haver um balanço entre a privacidade à qual todos têm direito, e os benefícios que podem ser provenientes do uso destas informações, como questões de saúde e segurança ou coleta de dados estatísticos para entender padrões de serviços com o uso da Internet, 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).

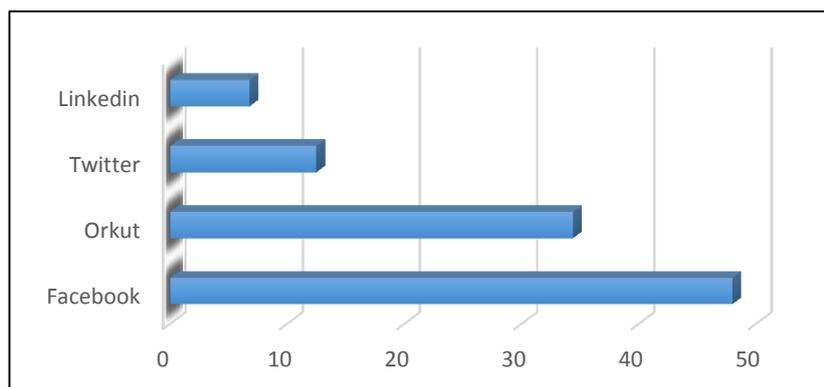
2.1 Método

A pesquisa descreve 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. Para tal, utilizam-se fontes de dados secundárias

com o objetivo de aprofundar o conhecimento sobre o tema. Quanto à abordagem de pesquisa, esta é quantitativa para pesquisa e classificação dos resultados.

Para a coleta de dados, optou-se por usuários do Facebook, devido a sua maior representatividade entre os usuários brasileiros de redes sociais, conforme demonstrado na Figura 5:

Figura 5 – Redes sociais com maior quantidade de usuários brasileiros (em milhões)



Fonte: Adaptado de VOLGARIN (2012)

Utilizando-se do próprio conceito de conexões existentes em uma rede social, o formulário de pesquisa com alternativas pré-fixadas descrito no

Quadro 1 foi enviado para todos os contatos do pesquisador, sendo retransmitidos por suas conexões, formando assim uma cadeia de comunicação aleatória de usuários da rede. Desta forma, não houve um controle sobre o perfil dos usuários que responderam à pesquisa, objetivando capturar a diversidade inerente à própria rede social.

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?

A coleta de dados ocorreu no período de setembro de 2013 até abril de 2014, com 143 participantes.

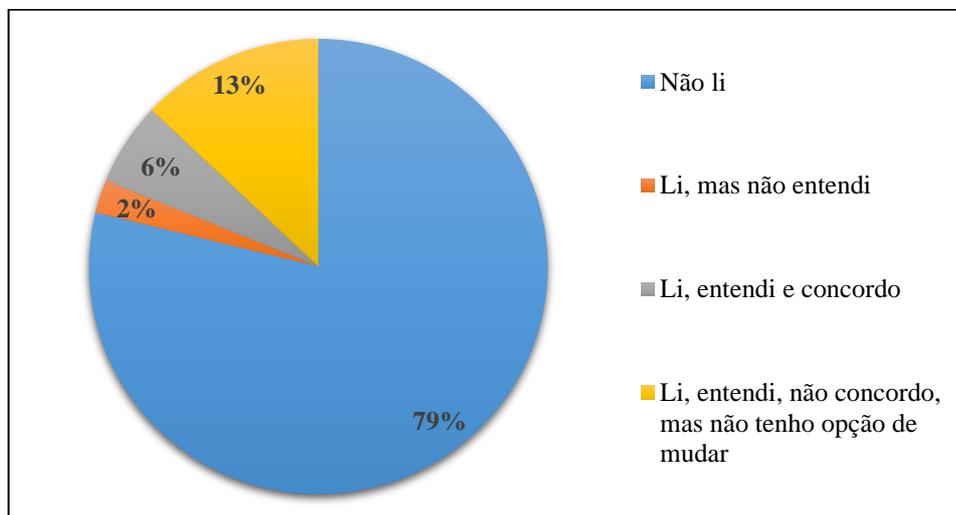
O número de participantes representa um percentual ínfimo dos usuários do Facebook, com isso esta pesquisa demonstra apenas tendências de comportamento dos usuários, sem, contudo, constituir massa crítica válida para um referencial estatístico.

3 RESULTADOS E DISCUSSÕES

As políticas de privacidade das redes sociais falham como instrumento prático de proteção e conscientização do consumidor. No estudo realizado, 79% dos entrevistados não se preocuparam em ler a política de privacidade, demonstrando indiferença com relação à relevância destas perante o uso da rede social. Do total, apenas 6% efetivamente entenderam e concordaram com as políticas de privacidade. Fatores que podem ter influenciado este resultado são o tamanho do texto, a

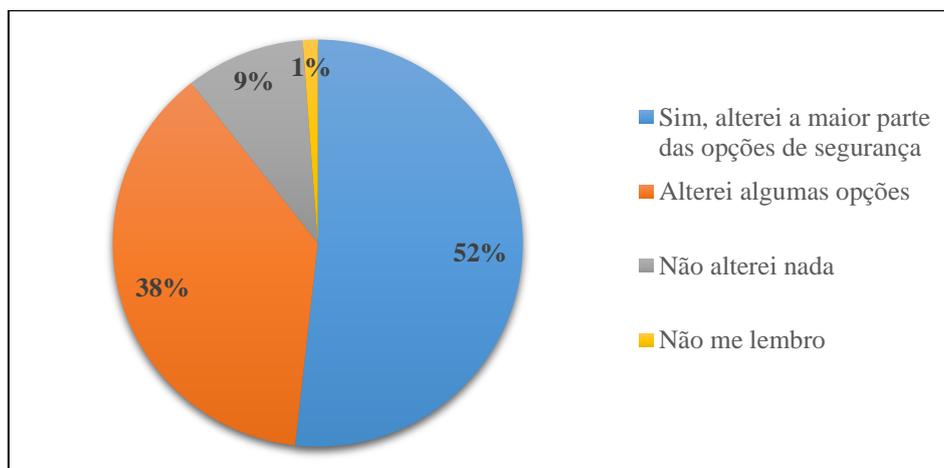
complexidade da linguagem e a impotência com relação a não ser possível alterar nenhuma cláusula do documento.

Figura 6 – Usuários e as políticas de privacidade do Facebook



Outro ponto relevante apontado por Krishnamurthy (2010) é a configuração de segurança e privacidade das redes sociais, que expõe por padrão várias informações dos usuários, conforme visto na Figura 4. No caso do Facebook, é possível configurar algumas restrições básicas com relação à segurança e privacidade do usuário. Averiguou-se que a maioria dos usuários (90%) alega ter realizado ao menos alterações básicas nestas políticas.

Figura 7 – Usuários e as configurações de segurança e privacidade do Facebook

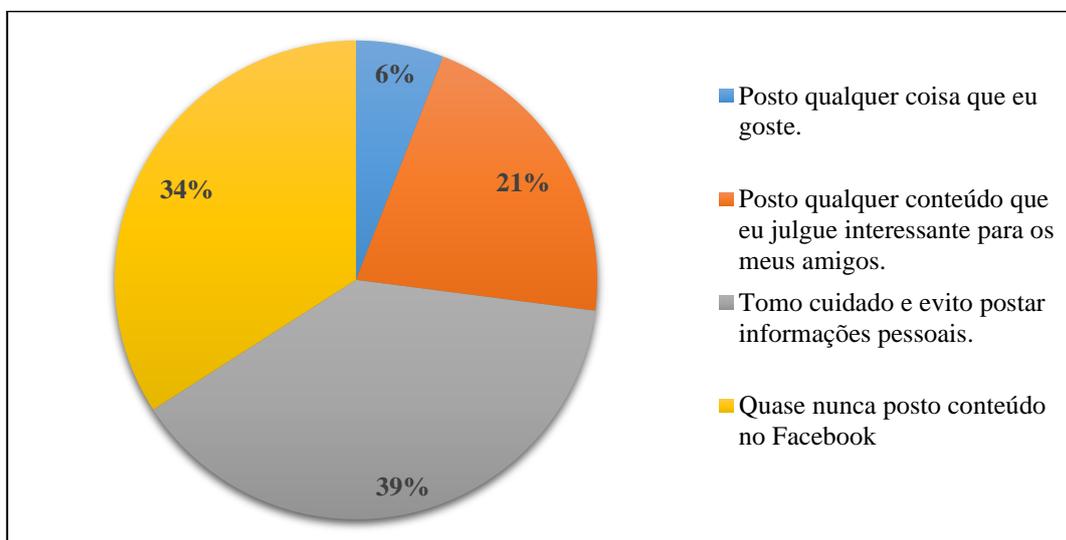


Entretanto, ao analisar os perfis dos usuários que responderam a pesquisa, verificando se estes haviam restringido o acesso de desconhecidos a duas informações

personais (e-mail e postagens), notou-se 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.

No que tange as postagens de conteúdo no Facebook, apenas 39% afirmaram usar como critério para postagens a cautela quanto à exposição de informações pessoais.

Figura 8 – Critérios ao postar conteúdos no Facebook



Ao questionar sobre o uso de postagens do Facebook para propaganda direcionada, descobriu-se que apenas 51% são tolerantes a esta prática, sendo que apenas 49% concordam da utilização das informações postadas para outros fins.

Esses números mostram que os usuários entrevistados não possuem uma total compreensão dos aspectos de privacidade e segurança das suas informações ao utilizarem o Facebook e possuem opiniões divididas sobre o uso destas por empresas, para criação de propagandas direcionadas ou quaisquer outros fins.

4 CONSIDERAÇÕES FINAIS

Dados como cookies ou postagens em redes sociais constitui fonte de informação disponível aos sistemas de *Big Data* e podem ser utilizados por empresas para que estas obtenham maior conhecimento sobre os consumidores, criando ofertas individualizadas e melhor relacionadas com os interesses de cada indivíduo. Por sua vez, o consumidor poderá ser beneficiado ao receber ofertas melhor ajustadas às suas necessidades. Nota-se, entretanto, um usuário das redes sociais desinformado sobre os possíveis usos das informações disponibilizadas nestes sítios. 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, porém o tema ainda carece de regulamentações que estabeleçam padrões de identificação e de segurança dos dados pessoais. Adicionalmente, mesmo sendo possível para as empresas utilizarem dados de redes sociais, estas devem avaliar os aspectos éticos envolvidos e criar políticas claras para os seus consumidores.

Esta pesquisa é um estudo inicial sobre o tema, necessitando ainda uma quantidade maior de respostas coletadas para permitir estudos estatísticos com maior grau de assertividade.

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**. 2013

BOYD, D; ELLISON, N. Social Network Sites: Definition, History, and Scholarship. **Journal of Computer-Mediated Communication**. 2008

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 out. 2013.

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 out. 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