SISTEMAS & GESTÃO
Revista Eletrônica

PRIVACIDADE EM *BIG DATA*: PANORAMA E AGENDA DE PESQUISA

Celina Rebello Silva
celina.rebello@coppead.ufrj.br
Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro, Rio de Janeiro, Brasil

Elaine Maria Tavares Rodrigues
elaine.tavares@coppead.ufrj.br
Universidade Federal do Rio de Janeiro – UFRJ, Rio de Janeiro, Rio de Janeiro, Brasil

RESUMO

Este artigo teve por objetivo mapear a produção acadêmica na área de privacidade em *Big Data*, para o domínio da Administração, levantando o estado da arte sobre o tema e algumas oportunidades de pesquisa oriundas dos *gaps* na literatura acadêmica. Privacidade é uma questão ainda em aberto e cada vez mais ameaçada por conta do efeito onipresente de dispositivos geradores de dados, como celulares, sensores e computadores, e de novas técnicas analíticas como *Big Data*. Foi realizado o levantamento bibliométrico extensivo da produção acadêmica sobre o tema nas principais bases de referência e análise de conteúdo aliada à mineração de textos, utilizando linguagem R. Os resultados apontaram não só áreas de concentração e temas de pesquisa mais abordados, mas também novas alternativas de pesquisa na área.

Palavras-chave: Privacidade, Big Data, Mineração de dados.

PROPPI / LATEC
DOI: 10.20985/1980-5160.2017.v12n4.1205

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp.491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

1. INTRODUÇÃO

Novos dispositivos e técnicas analíticas, como a Internet das Coisas e *Big Data*, ampliaram recentemente a capacidade das novas tecnologias de informação e comunicação. Sensores, mídias sociais e etiquetas RFID (Identificação por Rádio Frequência, do inglês *Radio-Frequency Identification*), por exemplo, aumentaram a sobrecarga de dados já existentes nas organizações. Como forma de lidar com o alto volume de dados, em diferentes formatos, desenvolveu-se novas técnicas analíticas.

Big Data está relacionado aos conjuntos de dados, cujo tamanho está além da capacidade de ferramentas típicas de software de banco de dados para capturar, armazenar, gerenciar e analisar (Minelli et al, 2013; Ohlhorst, 2013). Os principais atributos relacionados ao conceito de *Big Data* são o volume, a velocidade e a variedade (Simon, 2013).

O volume está relacionado à quantidade crescente de dados, que impactam diretamente nos processos organizacionais e influenciam métodos preditivos e estatísticos. Em um mercado altamente competitivo ou em contextos administrativos de alta complexidade, encontrar novas maneiras de interpretar os dados e processá-los de maneira mais rápida tem se mostrado uma capacidade importante. A variedade está relacionada à capacidade de analisar uma extensa gama de tipos de dados e fontes, incluindo dados estruturados, semi-estruturados e não estruturados (Ohlhorst, 2013) que, como *Big Data*, toma a forma de mensagens, imagens e outros tipos de dados em redes sociais, sensores, GPS de celulares, dentre outros. Por fim, a velocidade refere-se à capacidade de analisar dados de forma mais ágil, por vezes em tempo real (McAfee et Brynjolfsson, 2012).

A estes atributos soma-se o valor - resultado agregado a partir das análises das informações, como a qualidade das informações ou o valor financeiro extraídos (Beath et al., 2012; Maçada et Canari, 2014); e a veracidade – relativa à pureza e à autenticidade das informações (Ohlhorst, 2013).

Como qualquer tecnologia inserida no contexto organizacional, a implementação e o uso de *Big Data* passa por desafios, alguns deles explorados na literatura acadêmica. Um dos pontos mais nebulosos é que o *Big Data* levanta questões preocupantes para a ética, como, por exemplo, quais dados podem ser utilizados em uma análise (Tene et Polonetsky, 2013). Nesse sentido, alguns questionamentos emergem, dentre as quais: sob que condições alguém pode/deve ser considerado como parte de um grande conjunto de dados? E se alguma postagem no domínio 'público' é tomada fora do contexto e analisada de uma forma que o autor nunca imaginou? O que significa para alguém ser identificado, ou ser analisado sem saber? (Boyd et Crawford, 2012).

Faz-se necessário discutir quando e quais dados podem ser considerados como parte da estratégia de *Big Data*, tendo em vista que a dificuldade de garantir a segurança e a privacidade dos dados pode inviabilizar projetos (Boyd et Marwick, 2011). É essencial um questionamento ético constante não só sobre o uso, mas também sobre a coleta, o armazenamento e o controle de acesso a esses dados (Simon, 2013; Tene et Polonetsky, 2013).

Considerando que a privacidade está entre as principais preocupações em *Big Data*, esse artigo tem como objetivo mapear a produção acadêmica em Administração sobre a privacidade nesse domínio, revelando o estado da arte sobre o tema e identificando oportunidades de pesquisas futuras em gaps estabelecidos na literatura. Por meio de revisão de literatura e emprego de técnicas bibliométricas, o artigo levanta questões e indicadores ligados ao conteúdo das publicações e focos principais das investigações, tais como questionamentos sobre dados públicos e privados, individuais e coletivos, e como esses aspectos são abordados por pesquisadores da área de Administração.

A pesquisa bibliométrica atua como viabilizador do mapeamento da produção acadêmica sobre o tema de privacidade em *Big Data*, metrificado por contagem de artigos por periódicos e autores. Num segundo estágio, o conteúdo dos artigos mapeados é minerado, em sua totalidade, e analisado.

A relevância desse estudo se reforça pelo fato do conceito de privacidade ser difuso e as pesquisas sobre o tema privacidade em administração possuir viéses muito distintos. Alguns artigos, como o de Boyd et Crawford (2012), abordam questionamentos sobre *Big Data* nas áreas de ciências sociais;, enquanto outros, como o de Martin (2015), têm por foco aspectos éticos. Chen et al. (2012), por sua vez, mostram a evolução de *Big Data* no contexto tecnológico e estratégico para organizações. Esses trabalhos iniciam a abordagem dos desdobramentos do contexto de privacidade em *Big Data*, porém, nenhum elucida o estágio de produção científica alcançado para a área.

O artigo está estruturado em quatro sessões, sendo a primeira a apresentação dos conceitos e definições dos temas de pesquisa; e a segunda trata do processo de levantamento dos artigos, bases de investigação, tratamento das respostas e instrumentos de análise para mineração do conteúdo. A terceira sessão descreve a análise e interpretação dos resultados, apontando lacunas de pesquisa a serem exploradas; e, na quarta, encontram-se as conclusões.

1.1 Big Data

Muito se fala sobre *Big Data.* Reconhece-se seu potencial econômico, político e outras dimensões nas quais o fenô-

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp. 491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

meno está implicado. Todavia, segundo Boyd et Crawford (2012), Martin (2015) e Zuboff (2015), não existe consenso para a definição do termo. Neste sentido, Gandomi et Haider (2015) fizeram o esforço de estabelecer uma linha evolutiva das definições e seus conteúdos. Assim, segundo a literatura, *Big Data* é um fenômeno definido por diversas lentes.

Diebold (2012) coloca *Big Data* como assunto chave para todas as ciências e alega ser sua a primeira referência do termo num trabalho acadêmico. Porém, reconhecendo que a origem do termo remete a trabalhos não acadêmicos, de especialistas do domínio da ciência da computação, Gandomi et Haider (2015) reafirmam o caráter multifacetado da definição de *Big Data*, começando pela mais adotada, a de Doug Laney – o modelo dos 3V's – que apesar de não acadêmica, é citada em Chen et al. (2012). Gandomi et Haider (2015) acreditam que as definições do termo tenham evoluído rápido pelo incentivo de adoção vindo de empresas como a IBM, EMC, Teradata e SAP, dentre outros gigantes do setor de computação.

No modelo dos 3V's, volume se refere à quantidade de dados coletados e armazenados e que, por conta da magnitude, não existia ainda capacidade de processamento disponível para atuar sobre eles. A variedade se refere à diversidade de natureza dos dados: sensores, arquivos de vídeo, *tweets*, imagens, entre todas as formas de produção digital disponível. Velocidade remete ao fluxo de informações que vem crescendo de forma inegável graças à computação ubíqua e que permite muitas vezes análises em tempo real. Gandomi et Haider (2015) apontam que o modelo dos 3V's de Laney evoluiu para 6V's, onde o quarto V é definido pela IBM como veracidade; o quinto V é definido pela Oracle como valor; e o sexto V é definido pela SAS como variabilidade. A veracidade refere-se a quão fidedignos são os dados; o valor se refere ao valor que esse volume de dados agrega para a organização; e, por fim, a variabilidade refere-se às mudanças na estrutura dos dados e como usuários interpretam esses mesmos dados.

Do modelo mais tradicional de 3V's de Laney, em que o *Big Data* é definido por velocidade, variedade e volume, até hoje, as definições de *Big Data* se multiplicam e ganham um enfoque estratégico. Em Chen et al. (2012), o modelo dos 3V's é associado às técnicas analíticas e avaliado pelos autores como um oceano azul de oportunidades para negócios, pesquisa e diversas aplicações. *Big data* é definido como o conjunto de tecnologias e ferramentas (bancos de dados e ferramentas de mineração de dados), bem como técnicas (métodos analíticos), passíveis de emprego em larga escala, para dados complexos, num universo de aplicações, de modo a aprimorar o desempenho das organizações. Processos de armazenamento, gestão, capacidade de análise e visualização de dados fazem parte do arcabouço de *Big Data*.

Já para Boyd et Crawford (2012), a definição de *Big Data* é descrita como: "um fenômeno cultural, tecnológico e acadêmico, que se estabelece entre a dinâmica da tecnologia, análise e mitologia que provoca retórica utópica e distópica". Para os autores, o *Big Data* é um fenômeno sócio-técnico, cujos benefícios reais devem ser criticamente questionados e cuidadosamente examinados. Eles ressaltam que *Big Data* é visto como ferramenta de alto potencial para mazelas sociais, como identificação de células terroristas, cura do câncer e etc. e, ao mesmo tempo, ameaçadora pelo potencial para ferir questões como o direito do indivíduo à privacidade. Essa mesma visão crítica de desdobramentos do uso de *Big Data* pode ser observada em Martin (2015).

1.2 Privacidade

A palavra privacidade é apontada como um anglicismo de *privacy,* que tem raiz no termo latim *privare.* Assim como *Big Data*, a palavra privacidade não possui um conceito objetivo e único. Há vários posicionamentos doutrinários quanto ao seu significado, podendo ser descritivo ou normativo. Frequentemente, privacidade é definida em termos de controle da informação.

Definições de privacidade são usadas para denotar a maneira como as pessoas definem situações e condições de sigilo e como são avaliadas essas situações, ou, ainda, para indicar a necessidade de haver restrições sobre o uso da informação ou seu processamento. Newell (1995) analisa privacidade numa perspectiva multidisciplinar, definindo formas e razões para seu estabelecimento, compreendendo o termo como separação temporária do domínio público. Glenn (1966) coloca que não há consenso se privacidade é, ou não, um direito entre estudiosos do direito e da filosofia, e que na época de seu estudo, já era um tema controverso.

Privacidade informacional, em um sentido normativo, refere-se, normalmente, a um direito moral não absoluto de pessoas para terem controle direto ou indireto sobre o acesso a (1) informações sobre si mesmo, (2) situações em que outros poderiam adquirir informações sobre si, e (3) tecnologia que pode ser usada para gerar, processar ou divulgar informações sobre si mesmo.

A necessidade de se proteger a vida privada surgiu da conflitante relação entre o indivíduo e a ordem imposta pela sociedade. Nissembaum (1997) coloca a definição de privacidade por Charles Fried como: "A privacidade não é simplesmente a ausência de informações sobre nós nas mentes dos outros, ao contrário, é o controle que temos sobre informações sobre nós mesmos". A autora segue a definição de W.A. Parent: "Privacidade é a condição de uma pessoa não ter suas informações pessoais conhecidas irregularmente por terceiros", se referindo a fatos que a maioria das pes-

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp.491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

soas em uma dada sociedade escolhe não revelar sobre si (exceto para os amigos, a família, assessores, etc.) ou a fatos sobre os quais uma pessoa em particular é extremamente sensível, escolhendo não revelar. Já Westin (2003) diz que a privacidade é a afirmação de que indivíduos e grupos determinam por si próprios, quando, como, e em que medida as informações sobre eles são comunicadas aos outros. Beardsley (1971) sugere que as pessoas têm o direito de decidir quando e quanta informação sobre si será revelada a terceiros. Gavison (1980), por sua vez, tem oferecido variantes de definições de privacidade em três esferas, todas como aspecto central de restrição de acesso: a primeira como direito, a segunda como perda da privacidade e a terceira com as punições. A ideia central reside na questão de limitar dados e informações às pessoas, ou às informações sobre pessoas.

Warren et Brandeis (1890) já protestavam contra atividades intrusivas de jornalistas. À altura, levantavam o debate sobre os limites individuais de até onde a sociedade poderia saber algo a respeito de um cidadão. Eles ressaltavam que as mudanças políticas, sociais e econômicas demandam o reconhecimento de novos direitos, e a lei comum segue evoluindo no sentido de garantir as novas demandas da sociedade. A lei, segundo os autores, se atualiza de forma a restabelecer equilíbrio e inviolabilidade dos direitos fundamentais.

O direito tenta, através de doutrinas, limitar o caráter invasivo de algumas exposições da privacidade. Temos, em Fuster (2014), a questão da privacidade à luz da teoria das esferas de Robert Alexy. Essa teoria reforça conceitos de Parent (1983), definindo a existências de três diferentes níveis de proteção da vida privada, esferas com diferentes intensidades de proteção. Assume a **esfera mais interna** (âmbito último intangível da liberdade humana) como a mais íntima, intangível, extremamente reservada, com assuntos mais secretos que não devem chegar ao conhecimento de terceiros; a **esfera privada ampla** como o âmbito privado na medida em que não pertença à esfera mais interna, incluindo assuntos que o indivíduo leva ao conhecimento de outra pessoa de sua confiança, ficando excluído o resto da comunidade; e **a esfera social**, que engloba tudo o que não for incluído na esfera privada ampla: todas as matérias relacionadas com as notícias que a pessoa deseja excluir do conhecimento de terceiros. O direito à privacidade, então, se definiria como aquilo que nos preserva do conhecimento alheio, nos reservando a nossa própria vivência.

Existem razões morais para proteção dos dados pessoais e para a prestação de controle direto ou indireto sobre o acesso a estes dados por parte de terceiros. Van den Hoven (2008) os define como:

- acesso irrestrito por outros para senhas de um indivíduo, em que características e paradeiro podem ser usados para prejudicar o titular dos dados em uma variedade de maneiras;

- **Assimetria da informação:** dados pessoais tornaram-se commodities. Normalmente, os indivíduos não estão em uma boa posição para negociar contratos sobre a utilização dos seus dados e não têm os meios para verificar se os parceiros vão rezar nos termos do contrato. Leis de proteção de dados, regulamentos e governança se destinam a estabelecer condições justas para a elaboração de contratos sobre a transmissão dos dados pessoais, intercâmbio e fornecimento de dados com freios, contrapesos e garantias de reparação;

- **Injustiça informacional e discriminação:** as informações pessoais fornecidas em uma esfera ou de contexto (por exemplo, cuidados de saúde) podem mudar o seu significado quando usadas em outra esfera ou contexto (como as transações comerciais) e podem levar à discriminação e prejuízos para o indivíduo;

- **Invasão de autonomia moral:** a falta de privacidade pode expor os indivíduos a forças externas que influenciam as suas escolhas.

Para o direito, todas estas formulações fornecem boas razões morais para limitar e restringir acesso aos dados pessoais, fornecendo aos indivíduos controle sobre seus dados. Quando terceiros assumem postura invasiva, surgem episódios como o caso Carolina Dieckman, ocorrido no Brasil, que resultou em lei; e, em âmbito mundial, as revelações de Edward Snowden e suas implicações em relações econômicas e diplomáticas.

No objetivo de encontrar esse âmbito mais íntimo e interno do indivíduo, questiona-se a existência de algum conjunto de comportamentos que respeite os interesses da vida em comunidade. Talvez, a natureza da estrutura social tenha se desenvolvido de tal forma que o passado recente force o reconhecimento de que a privacidade, até agora presumida como ingrediente de ação moral, deva ser agora especificada como direito. A filosofia que descreve a estrutura política como essencialmente de natureza corporativa tem tradicionalmente derivado, ou resultou em uma descrição do estatuto moral do indivíduo que não só nega o direito à privacidade, mas designa um delito político e moral.

Warren et Brandeis (1890) argumentam que o ato de publicar um determinado conteúdo faz com que o indivíduo abra mão de seu direito à privacidade. Já a calúnia, como é gerada por terceiros, deve ser tratada conforme os instrumentos legais. Apesar dos esforços argumentativos de Warren et Brandeis (1890), Glenn (1966) debate o contraste en-

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp. 491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

tre definições de privacidade a partir da distinção de Hegel de *moralität,* como sendo relativo ao julgamento particular, e *Sittlichkeit*, como obrigações definidas por ordens corporativas e institucionais, fundamentando que a reivindicação de privacidade é simplesmente trivialidade, por conta de uma irresponsabilidade praticada pelo indivíduo. Ele continua sua análise da privacidade apontando o indivíduo como inteiramente dependente do grau em que identifica os seus interesses e os direitos com a estrutura de valor apropriado para a ordem particular e corporativa em que se encontra. Uma vez que os direitos e deveres dos indivíduos são determinados pelas ordens existentes, em que ele também participa e cuja forma mais elevada é o estado, então, a redução da privacidade ou o seu limite é realizada, com a alegação de que, em última instância, o indivíduo deve aceitar a interpretação de um "árbitro" que discerne corretamente valores, deveres e obrigações do momento histórico.

Glenn (1966) estabelece o contraponto com a filosofia anglo-americana, pois esta define a estrutura política como legitimidade coletiva da qual depende a privacidade. Ela deriva e depende de julgamentos individuais daqueles que estão constituídos naquela coletividade, definição essa que se mostra alinhada com a definição de Newell (1995). Ele percebe os problemas que poderiam advir de uma confusão entre responsabilidade moral e de responsabilidade legal. Privacidade é assumida como sendo um direito justificado pela utilidade, se não pela natureza, ingrediente essencial da filosofia política anglo-americana, como todo o direito, que deve ser protegido pela lei. A persistência do pressuposto mal definido para a ideia de privacidade abre precedentes perigosos e ameaçadores e forma uma zona cinza entre a responsabilidade moral e responsabilidade legal, que surge do dilema de se punir a invasão de privacidade e, mesmo que por motivos puramente utilitários, se a punição causaria mais danos do que benefícios.

A questão da privacidade, em Glenn (1966), é contextualizada considerando a heterogeneidade geográfica, moral, religiosa, em contraponto à contextualização de Hegel. Hegel buscou a ambiguidade da dialética histórica para descrever um organismo político de tal homogeneidade, cuja estrutura societária iria resolver relações, conflitos morais, políticos, individuais, sociais. O modelo hegeliano cai em desuso pela complexidade das questões individuais, pelo julgamento privado e de oposição. A sociedade face a face, herança grega, rejeitava o conceito de privacidade.

Glenn (1966) aponta que Bentham reconhecia que o status moral do indivíduo necessitava de proteção contra uma organização social cada vez mais intrusiva e dominante. A rejeição da privacidade tomou forma de ataques agressivos, culminando em julgamentos privados e de opinião. Hegel conseguiu, ainda que em bases ambíguas, descrever a natureza da ordem política que foi rapidamente se tornando real, e, mesmo que suas premissas fossem inaceitáveis, muito se avançou graças à sua análise da estrutura institucional da ordem social. Bentham, ao se recusar a abandonar seu sistema de valores individuais, demonstrou que os padrões tradicionais de análise política eram inadequados para formular o status do indivíduo nos contextos político e legal da sociedade e do estado da época.

1.3 Privacidade em *Big Data*

A dinâmica da privacidade em *Big Data* possui pontos semelhantes quanto aos desafios de Warren et Brandeis (1890) sobre a facilidade com que informações são divulgadas. Porém, Matzner (2014) alerta que, atualmente, pessoas disponibilizam dados sem qualquer critério, aceitando como contrapartida benefícios de pouco valor ou mesmo nada em troca.

Kshetri (2014) e Zuboff (2015) apontam que a ubiquidade do Big Data também favorece a exacerbação das assimetrias de poder entre estados, indústrias, grupos e indivíduos. As tendências são de crescimento dos problemas de privacidade, por conta da produção de dados e a exposição de informações de teor privado, coletadas sem plena consciência dos indivíduos. McNeely et Hahm (2014) colocam questões como: quais dados são coletados e quais não? Por quê? O que é usado e o que não é? Quais são os desdobramentos dessa seleção? Como e por quê? O que não é mensurado? Quais fatores fundamentais ou críticos devem ser considerados para pleno entendimento de um fenômeno particular ou condição? Essas questões apontam para a necessidade de uma abordagem crítica sobre Big Data em termos do entendimento sobre a sua essência, uso e efeitos.

Essa mesma classe de questionamentos é abordada por Martin (2015), ao analisar a cadeia de suprimentos característica do Big Data, dos usos positivos e negativos da tecnologia, ressaltando aspectos como revenda de dados e o risco de mau uso da informação, com impacto significativo sobre os usuários como destruição de valor, redução de direitos das partes interessadas, e desrespeito a algum indivíduo envolvido no processo. A autora coloca, ainda, que a indústria de Big Data gera externalidade agregada negativa, por ampliar o sistema de vigilância através do qual uma gama de informações é coletada e reunida de forma invisível ao usuário.

Das questões de Boyd et Crawford (2012), aliada aos aspectos técnicos e ambientais apontados por Chen et al. (2012) e visão estratégica de Martin (2015), percebe-se que a realidade proporcionada pelo fenômeno Big Data gera implicações sobre a questão da privacidade, que não podem ser ignoradas ou negligenciadas. A vertente técnica aparece como condição ambiental, sendo necessária a contextuali-

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp.491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

zação de seu uso e aplicações, sendo necessária abordagem analítica profunda sobre os desdobramentos, não só sob o prisma da cadeia de Big Data como fez Martin, mas de um passo além: o de estabelecimento de regras e contornos, definições de limites para que, essa realidade que é o *Big Data,* gere externalidades positivas e não o que já se tem percebido recentemente como as práticas denunciadas por Edward Snowden. No atual estágio da dinâmica, mostra-se necessário o estabelecimento de requisitos que equilibrem a relação do acesso das informações e o direito à privacidade.

Os requisitos analíticos demandados por problemas de privacidade são basicamente dois: a definição de valores e a especificação de procedimentos. O primeiro é tarefa moral da filosofia, refletida na legislação; o segundo depende do processo legal, especialmente se ele opera no controle de funções administrativas em que o indivíduo é tema. A falha ou negação na execução dessas tarefas nos deixa com a única alternativa de aumentar o controle administrativo arbitrário que pode ativamente lograr as conquistas dos valores individuais. O controle dos processos legais é mandatório, mas deve ser justificado. Se privacidade for definida como requisito essencial para alcance da moralidade, então, privacidade é um direito que a lei deve não só proteger, mas prover. O homem moderno nasce acorrentado e só a lei poderia libertá-lo. O que se percebe, e será mostrado mais adiante, é uma tendência forte de pesquisas sobre políticas públicas e políticas digitais que colocam em pauta os limites ainda não delineados da dinâmica da privacidade em Big Data.

2. PESQUISA BIBLIOMÉTRICA

O início do processo bibliométrico se deu com a escolha das bases e a definição das ferramentas de controle de referências. Para controle de referências, foi utilizado o software Zotero em razão da sua capacidade de integração com navegadores *web*, bem como pelas funcionalidades de tratamento de redundâncias e interoperabilidade com os formatos "RIS", auxiliando na classificação de categorias e eliminação de redundâncias.

A busca se deu essencialmente pela conjunção das palavras chaves "*Big Data*" e "*Privacy*" no título, resumo ou campo de palavras-chaves. Definiram-se como critérios a busca em apenas periódicos acadêmicos revisados por especialistas, com disponibilidade de texto completo nas respectivas bases, no período de 2000 a 2015.

As bases foram selecionadas, no período de setembro a outubro de 2015, de modo a obter uma cobertura que se acredita ser razoável para uma revisão de literatura robusta. Foram retirados artigos sem autores e anônimos. Os resultados preliminares sem a leitura detalhada de resumos apresentaram a produção publicada nas seguintes bases: Ebsco, Emerald Insight, JStor, Proquest, Sage, Science Direct, Scopus, SpringerLink, Web of Science e Wiley Online.

Algumas considerações foram feitas para pré-seleção dos artigos. Nem todas as bases apresentam as mesmas funcionalidades em suas respectivas interfaces de busca, sendo necessário o refinamento manual dos resultados. No caso da Springer, os filtros de disciplina aplicados foram: "*Business and Management*" e "*Social Sciences*". Ainda assim, foram retornados artigos do tipo entrevistas, notas de pesquisa, entre outros que não possuem estrutura de artigo científico. Quanto à Wiley Online, não foi possível a seleção da base referente aos artigos de administração, sendo feito o refinamento sobre todos os resultados retornados por título da sessão de periódicos acadêmicos de Administração de forma manual.

2.1 Resultados pelas bases (2000-2015)

As publicações que tratam de privacidade em Big Data apresentam um pico de produção acadêmica no ano de 2014, dentro do conjunto de publicações em Administração, tal como mostra a figura 1.

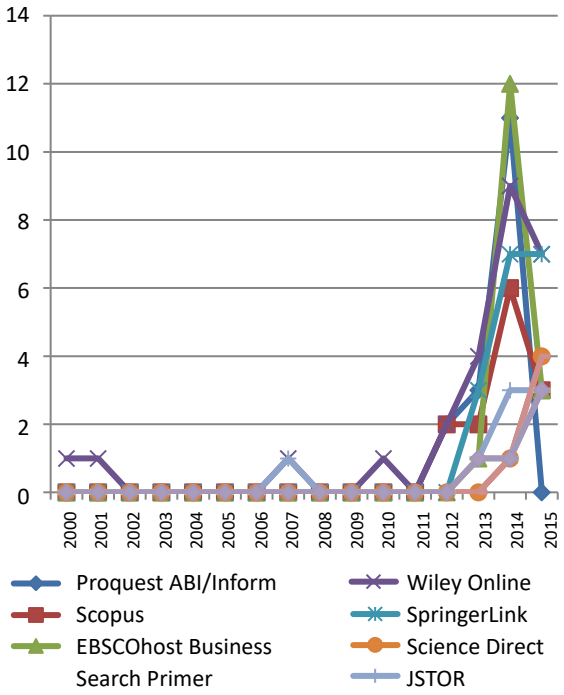

Figura 1. Séries temporais do número de publicações nas bases consultadas

Fonte: elaborado pelos autores(2015).

Os critérios de eliminação passaram por levantamento de artigos anônimos, sem referências em periódicos acadêmicos, artigos redundantes, com presença em mais de

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp. 491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

Tabela 1. Resultado do processo de seleção dos artigos a serem minerados textualmente

	Número de Artigos			
Bases	**Iniciais**	**Redundantes**	**Sem referências***	**Resultado Final**
EBSCO	16	8	4	**4**
EmeraldInsight	5	0	0	**5**
JSTOR	8	0	0	**8**
Proquest	16	2	7	**7**
Sage	5	0	0	**5**
Science Direct	5	0	0	**5**
Scopus	7	3	1	**3**
SpringerLink	17	0	1	**16**
Web of Science	5	0	0	**5**
Wiley Online	22	0	11	**11**
Total	116	13	24	69

Fonte: elaborado pelos autores (2015)

uma base, bem como leitura dos artigos em si. Artigos que apresentaram alguma das características citadas foram descartados para a fase de mineração de dados. Artigos cujas referências não continham periódicos acadêmicos revisados também foram eliminados. Os arquivos em duplicidade foram eliminados devido aos pré-requisitos do processo estatístico de mineração de dados, pois a redundância gera distorções matemáticas, e, por conseguinte, leva a resultados distorcidos do estado da pesquisa, temas e áreas de concentração. O resultado do processo de seleção dos artigos a serem minerados textualmente está apresentado na tabela 1.

Como a mineração de dados é o suporte bibliométrico empregado, o protocolo descreve o processo de associação de palavras, e, com isso, a redundância ocasionaria maiores pesos a termos em detrimento de outros.

Dentre as bases pesquisadas, a maior ocorrência de redundância se deu entre as que possuem portfólio com maior diversificação de periódicos, isto é, bases agregadoras da produção de editoras. As intersecções aconteceram essencialmente com EBSCO, Proquest e Scopus. Tomou-se como critério de permanência as bases que são editoras, e não as bases agregadoras.

3. MINERAÇÃO DE DADOS E ANÁLISE DE CONTEÚDO

Para o processo de mineração de dados, foi desenvolvido código fonte em linguagem R. Ao final do processo de seleção dos artigos foram selecionados 69 artigos de publicações revisadas por pares. A mineração consistiu em três etapas: a conversão dos arquivos para mineração; análise de conteúdo e resultados estatísticos; e a análise visual dos resultados estatísticos. A mineração de dados tem como objetivo identificar correlações entre termos chave dos estudos de privacidade em *Big Data* no domínio da administração, e foi executada considerando dicionário da língua inglesa, em sucessivas etapas, de modo a eliminar termos de alta frequência de ocorrência com baixo significado semântico, a saber: conectivos, preposições, pronomes, interjeições e advérbios, de modo a observar que, na segunda fase da mineração, fossem retiradas apenas classes de palavras que são substantivos, adjetivos, verbos, com baixa carga semântica e relevância, com emprego de dicionário de associações semânticas de termos. Na terceira fase foram computadas apenas palavras de frequência baixa, porém com alta carga semântica e relevância.

Após o corpo minerado ser composto, essencialmente, de substantivos, adjetivos e verbos, realizou-se a análise de frequência de termos e a correlação entre termos com associação semântica com as palavras-chave, palavras componentes do domínio de privacidade e *Big Data* e com aquelas de maior significado semântico nos textos retornados pelas bases.

Mapeadas essas relações, calculou-se a ordem com que os termos ocorreram em quantidade e relevância, e, numa segunda etapa, sua associação em clusters. A associação em clusters auxiliou na identificação visual das correlações entre componentes dos estudos, bem como na identificação de clusters.

3.1 Análise de termos e frequências

A terceira fase da mineração de dados, apresentada na figura 2, encontrou os seguintes resultados: *"social", "security", "access", "approach", "context", "control", "government", "development", "disclosure", "personal", "individuals", "google", "business", "facebook", "model", "mobile", "need", "personalization", "process", "services", "terms", "value", "risk", "scale", "role", "trust", "questions", "regulation".*

Esse conjunto indica que objetos de pesquisa relacionados à privacidade em *Big Data* passam por questões perti-

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp.491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

nentes às abordagens, desdobramentos, valores e regras, do que propriamente à tecnologia. A ordem de grandeza com que esses termos podem ser vistos, conforme figura 2, revela que os aspectos sociais são os mais importantes dentro da produção realizada sobre privacidade em *Big Data* na área de administração. A produção acadêmica tem voltado sua atenção para os papéis de atores, como o indivíduo, governo e companhias; políticas de controle e acesso; processos, serviços e termos propostos ao indivíduo por companhias como Google e Facebook; processos de tratamento, custódia, exposição, abertura, extração de valor dos dados, o risco e o futuro relacionados nesse contexto.

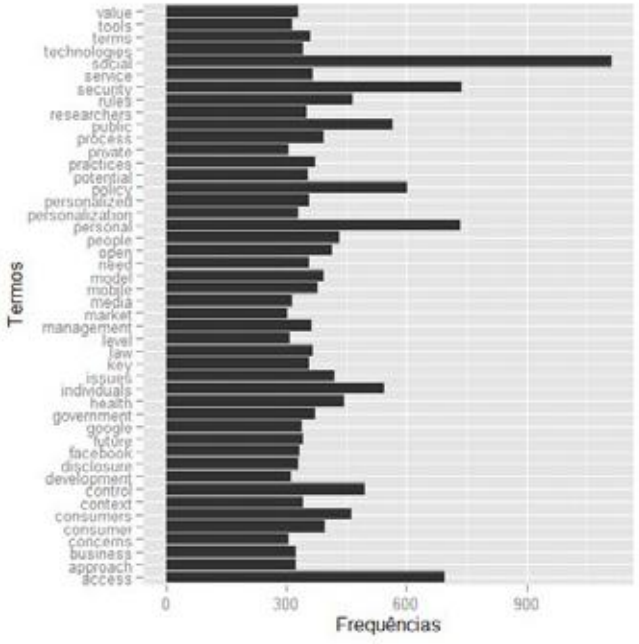

Figura 2. Análise de termos e frequências – primeira mineração da terceira fase

Fonte: elaborado pelos autores (2015)

A complementação da análise pela representação via dendograma (figura 3) reforçou o indicativo de que palavras como *"social", "personal" e "security"* devem ser eliminadas para aprofundamento da iteração seguinte. A eliminação desses termos deveria favorecer o melhor entendimento entre as relações de palavras pela proximidade de seus níveis na estrutura. Uma vez eliminados os termos de maior fequência, devem aflorar novos arranjos de palavras, com baixa frequência e forte significado semântico, significando o estágio das pesquisas na área de concentração, bem como sua combinação em novos arranjos, potenciais indicadores de lacunas de pesquisa.

Assim, termos de baixa frequência e alta semântica serão *inputs* e tendem a participar do próximo dendograma. Eles devem formar novos clusters, após a remoção dos termos *"social", "personal"*, e *"security"*. Essa análise gerou a Tabela 2.

Eliminados os termos sugeridos pela análise prévia, tem-se o segundo nível de resultados da mineração. Nele, é possível percerber alguns termos com frequências muito próximas, como *"terms", "surveillance", "rules", "services",* num grupo de frequências próximas a 400 ocorrências, enquanto outros dois grupos contém *"default", "control",* e um terceiro com *"individuals", "management", "issues", "health"*, no gráfico de barras da figura 4, e no dendograma da figura 5.

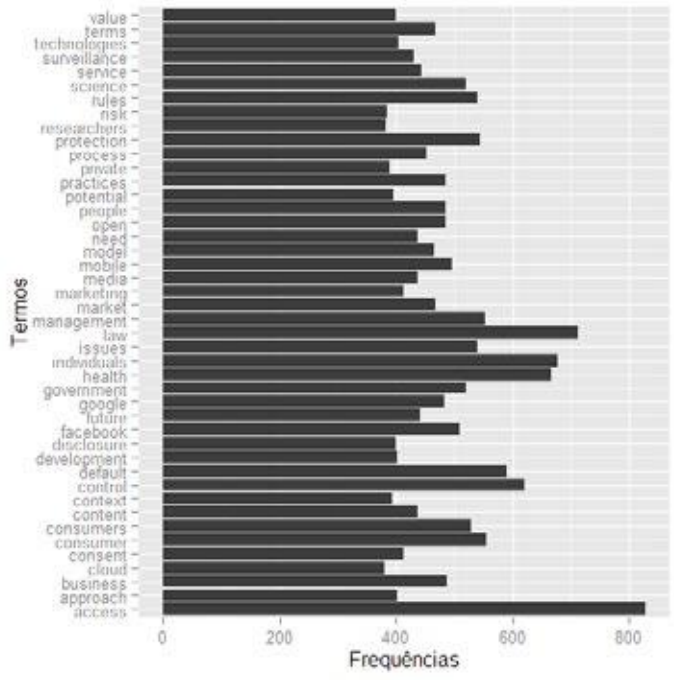

Figura 4. Análise de termos e frequências – primeira mineração da terceira fase.

Fonte: elaborado pelos autores (2015)

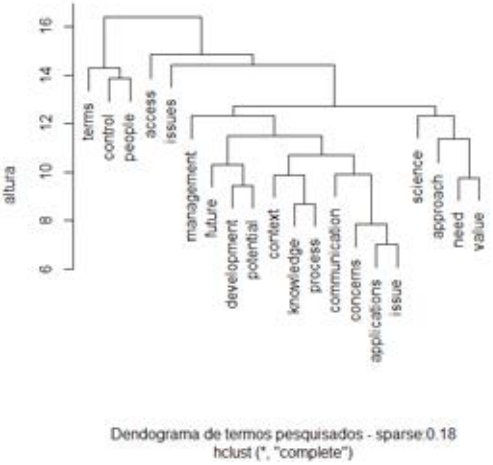

Figura 5. Dendograma da segunda mineração da terceira fase.

Fonte: elaborado pelos autores(2015)

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp. 491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

Tabela 2, Análise de termos de baixa frequência (n=280).

access	analytics	tools	business	challenges	cloud	collection	companies
people	concerns	model	policies	consumers	consent	content	researchers
control	default	terms	economic	Facebook	future	network	disclosure
mobile	industry	health	market	Context	key	media	knowledge
Law	process	trust	marketing	protection	issues	personalized	personalization
Need	condition	news	open	Google	sharing	government	management
press	potential	power	practices	Political	private	individual	development
quality	records	value	risk	Rules	science	service	technologies
Uses	society	Subject	support	approach	results	computing	surveillance

Fonte: elaborado pelos autores (2015)

Desta forma, o resultado da primeira mineração gerou um novo dendograma cuja estrutura reflete não só os termos da tabela 2, mas também aponta os indícios para áreas de concentração, como, por exemplo, para pesquisas sobre termos de controle sobre pessoas e questões de acesso. Os demais ramos do dendograma indicaram vertentes de pesquisa que consideram o balanço entre a abordagem científica do fenômeno de privacidade em *Big Data* com análises de necessidades e valor; o conhecimento dos contextos de processos; e o desenvolvimento de potencial futuro, em contraste com os aspectos gerenciais. São possíveis outras interpretações e arranjos dentro dos ramos, como o agrupamento sugerido abaixo, no dendograma da figura 6.

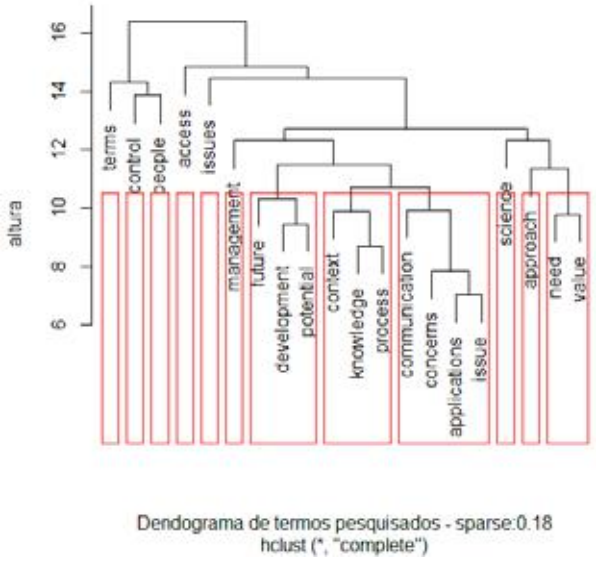

Figura 6. Agrupamentos para pesquisas potenciais
Fonte: elaborado pelos autores (2015)

Pelo dendograma da figura 6, foi possível visualizar algumas lacunas para pesquisa, como contrapontos entre a abordagem científica e abordagem gerencial às questões de privacidade em *Big Data*; questões de comunicação e contextualização entre usos das aplicações e processos de quem estaria de posse do conhecimento; como pessoas se comportam quando controladas; o que é considerado necessidade no contexto de *Big Data* e privacidade, entre outras opções.

A análise do mapa de palavras, na figura 7, construído a partir da mineração de dados, revela as principais questões tratadas pela produção acadêmica sobre privacidade com foco em *Big Data*, que vão desde relações econômicas sobre exposição das informações, passando pelos efeitos sobre governos, indivíduos e as relações de poder e confiança entre esses atores, até regras e direitos no contexto de uso de redes sociais. É interessante notar que seu resultado expressa o que foi levantado estatisticamente nos passos anteriores desse estudo: a questão central da privacidade em *Big Data* passa pelo balanço entre o acesso às informações dos indivíduos, o que são os direitos e como as leis tratam, se tratam, esses pontos.

Figura 7. Mapa de palavras. A relevância é expressa pela magnitude visual dos termos
Fonte: os autores (2015)

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp.491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

4. CONSIDERAÇÕES FINAIS

Dispositivos digitais transmitem informações diversas: algumas transparentes para o usuário, outras carregadas e compartilhadas por ele enquanto ator no próprio sistema. Mas, em algum momento do processo de aquisição de dados, há sempre intervenção humana para definir o que é transmitido ou qual uso será dado para essas informações; e, seja aceitando termos e condições contratuais, seja deixando um dispositivo exposto ou vulnerável, os efeitos dessa prática recaem sobre o indivíduo. São decisões tomadas, no âmbito corporativo ou governamental, ou por indivíduos, que, de forma indiscriminada, sem estabelecimento de limites, podem ter implicações ameaçadoras, como mostram Martin (2015) e Zuboff (2015), erodindo sistemas econômicos. A literatura e a análise apontam que esse aspecto compõe o núcleo de desafios para pesquisas de privacidade em *Big Data*, juntamente com o papel desempenhado pelas organizações do setor.

Big data é uma realidade e questões de privacidade sempre existiram. O campo se mostra não só um manancial de oportunidades, mas ao mesmo tempo desafiador para novas propostas de pesquisa, justamente no sentido do aprofundamento do debate de ideias de políticas sociotécnicas que fomentem o uso positivo de *Big Data*, preservando a privacidade e gerando externalidades positivas aos indivíduos e demais atores nesse ambiente. A produção acadêmica analisada concentra-se em questões de éticas e processos, muito mais levantando questões do que apontando soluções em si.

REFERÊNCIAS

Beardsley, E.L. (1971), "Privacy: Autonomy and Selective Disclosure", in Pennock, J.R.; Chapman, J.W. (ed.), Privacy: Nomos XIII, Atherton Press, New York, pp. 56-70.

Beath, C.; Becerra-Fernandez, I.; Ross, J, et al. (2012), "Finding value in the information explosion", MIT Sloan Management Review, Vol. 53, No. 4, pp. 18-20.

Boyd, D.; Crawford, K. (2012), "Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon", Information Communication and Society, Vol. 15, No. 5, pp. 662-679.

Boyd, D.; Marwick, A. (2011), "Social privacy in networked publics: teens attitudes, practices, and strategies", Privacy Law Schoolars Conference, 2 jun. 2011, disponivel em: http://www.danah.org/papers/2011/SocialPrivacy-PLSC-Draft.pdf (acesso em 20 jul. 2014).

Chen, H.; Chiang, R.H.L.; Storey, V.C. (2012), "Business Intelligence and Analytics: from big data to big impact", MIS Quarterly Executive, Vol. 36, No. 4, pp. 1165-1188.

Diebold, F.X. (2012), "A Personal Perspective on the Origin(s)and Development of 'Big Data': The Phenomenon, the Term, and the Discipline, Second Version", PIER Working Paper No. 13-003, disponível em: http://ssrn.com/abstract=2202843 (acesso em 11 nov. 2017).

Fuster G.G. (2014), The Emergence of Personal Data Protection as a Fundamental Right of the EU, Springer International Publishing, Switzerland.

Gandomi, A.; Haider, M. (2015), "Beyond the hype: Big data concepts, methods, and analytics", International Journal of Information Management, Vol. 35, No. 2, pp. 137-144.

Gavison, R. (1980), "Privacy and the limits of the law", The Yale Law Review Journal, Vol. 89, No. 3, pp. 421-471.

Glenn, N. (1966), "Philosophical Views on the Value of Privacy", Law and Contemporary Problems, Vol. 31, pp. 319-325.

Kshetri, N. (2014), "Big data's impact on privacy, security and consumer welfare", Telecommunications Policy, Vol. 38, No. 11, pp.1134–1145.

Maçada, A.C.G.; Canary, V.P. (2014), "A Tomada de Decisão no Contexto do Big data: estudo de caso único", XXXVIII Enanpad 2014, Rio de Janeiro, 13-17 set. 2014.

Martin, K.E. (2015), "Ethical Issues in the Big Data Industry", MIS Quarterly Executive, Vol. 14, No. 2, pp.67–85.

Matzner, T. (2014), "Why privacy is not enough privacy in the context of "ubiquitous computing" and "big data"", Journal of Information, Communication Ethics in Society, Vol. 12, No. 2, pp. 93-106.

Mcafee, A.; Brynjolfsson, E. (2012), "Big Data: The Management Revolution", Harvard Business Review, Vol. 90, No. 10, pp. 60-68.

Mcneely, C.L.; Hahm, J. (2014), "The Big (Data) Bang: Policy, Prospects, and Challenges", Review of Policy Research, Vol. 31, No. 4, pp. 304–310.

Minelli, M.; Chambers, M.; Dhiraj, A. (2013), Big Data Big Analytics: Emerging Business Intelligence and Analytic trends for today's businesses. John Wiley & Sons, Hoboken, New Jersey.

Newell, P. (1995), "Perspectives on Privacy", Journal of Environmental Psychology, Vol. 15, No. 2, pp. 87–104.

Nissembaum, H. (1997), "Toward an approach to privacy in Public: Challenges of Information Technology", Ethics and Behavior, Vol. 7, No. 3, pp. 207-219.

Ohlhorst, F. (2013), Big Data Analytics: Turning Big data into Big money, [S.l.]: Wiley.

Parent, W.A. (1983), "Privacy, Morality and the Law", Philosophy and Public Affairs, Vol. 12, No. 4, pp. 269-288.

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp. 491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

Simon, P. (2013), Too Big too Ignore, John Wiley & Sons, Hoboken, New Jersey.

Tene, O.; Polonetsky, J. (2013), "Big Data for All: Privacy and User Control in the age of Analytics", Northwestern Journal of Technology and Intellectual Property, Vol. 11, No. 5.

Van den Hoven, J.; Weckert, J. (2008), "Information technology, privacy, and the protection of personal data", Information Technology and Moral Philosophy, Cambridge University Press, Cambridge.

Warren, S.D.; Brandeis, L.D. (1890), "The Right to Privacy", The Harvard Law Review, Vol. 4, No. 5, pp. 193-220.

Westin, A.F. (2003), "Social and Political Dimensions of Privacy", Journal of Social Issues, Vol. 59, No. 2, pp.431-453.

Zuboff, S. (2015), "Big other: Surveillance capitalism and the prospects of an information civilization", Journal of Information Technology, Vol. 30, No. 1, pp. 75–89.

Recebido: 22 jul. 2016.

Aprovado: 11 out. 2017.

DOI: 10.20985/1980-5160.2017.v12n4.1205

Como citar: Silva, C.R.; Rodrigues, E.M.T. (2017), "Privacidade em Big Data: panorama e agenda de pesquisa", Sistemas & Gestão, Vol. 12, No. 4, pp. 491-505, disponível: http://www.revistasg.uff.br/index.php/sg/article/view/1205 (acesso: dia mês abreviado ano).

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp.491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

Anexo 1 - Artigos levantados

Serviram de base para mineração de dados os seguintes artigos:

Ambrose, M.L.; Ausloos, J. (2013), "The Right to Be Forgotten Across the Pond", Journal of Information Policy, Vol. 3, pp.1-23.

Barbu, A. (2013), "Eight contemporary trends in the market research industry", Management& Marketing, Vol. 8, No. 3, pp. 429-450.

Barocas, S.; Nissenbaum, H. (2014), "Big Data's End Run Around Procedural Privacy Protections", Communications of the ACM, Vol. 57, No. 11, pp. 31–33.

Bates, D.W.; Saria, S.; Ohno-Machado, L., et al. (2014), "Big data in health care: Using analytics to identify and manage high-risk and high-cost patients", Health Affairs, Vol. 33, No. 7, pp. 1123-1131.

Bertot, J.C.; Gorham, U.; Jaeger, P.T., et al. (2014), "Big data, open government and e-government: Issues, policies and recommendations", Information Polity, Vol. 19, No. 1/2, pp. 5–16.

Bozdag, E. (2013), "Bias in algorithmic filtering and personalization", Ethics and Information Technology, Vol. 15, No. 3, pp. 209–227.

Bragge, J.; Sunikka, A.; Kallio, H. (2012), "An exploratory study on customer responses to personalized banner messages in the online banking context", Journal of Information Technology Theory and Application, Vol. 13, No. 3, pp. 5-18.

Brennan, N.; Oelschlaeger, A.; Cox, C., et al. (2014), "Leveraging the big-data revolution: CMS is expanding capabilities to spur health system transformation", Health Affairs, Vol. 33, No. 7, pp. 1195-1202.

Cate, F.H.; Mayer-Schönberger, V. (2013), "Notice and consent in a world of big data", International Data Privacy Law, Vol. 3, No. 2, pp. 67-73.

Cate, F.H.; Cate, B.E., "The supreme court and information privacy", International Data Privacy Law, Vol. 2, No. 4, pp. 255-267.

Chow-White, P.A.; Macaulay, M.; Charters, A., et al. (2015), "From the bench to the bedside in the big data age: ethics and practices of consent and privacy for clinical genomics and personalized medicine", Ethics and Information Technology, Vol. 17, No. 3, pp. 189–200.

Chung, T.S.; Wedel, M.; Rust, R.T. (2015), "Adaptive personalization using social networks", Journal of the Academy of Marketing Science, Vol. 44, No. 1, pp. 66-87.

Cohen, I.G.; Amarasingham, R.; Shah, A., et al. (2014), "The legal and ethical concerns that arise from using complex predictive analytics in health care", Health Affairs, Vol. 33, No. 7, pp. 1139-1147.

Cranor, L.F.; Sadeh, N. (2013), "Privacy engineering emerges as a hot new career", IEEE Potentials, Vol. 32, No. 6, pp. 7–9.

Crosas, M.; King, G.; Honaker, J., et al. (2015), "Automating Open Science for Big Data", The ANNALS of the American Academy of Political and Social Science, Vol. 659, No. 1, pp. 260–273.

Curtis, L.H.; Brown, J.; Platt, R. (2014), "Four health data networks illustrate the potential for a shared national multipurpose big-data network", Health Affairs, Vol. 33, No. 7, pp. 1178-1186.

Daries, J.P.; Reich, J.; Waldo, J., et al. (2014), "Privacy, Anonymity, and Big Data in the Social Sciences", Communications of the ACM, Vol. 57, No. 9, pp.56–63.

Duan, R.; Hong, O.; Ma, G. (2014), "Semi-Supervised Learning in Inferring Mobile Device Locations", Quality and Reliability Engineering International, Vol. 30, No. 6, pp. 857–866.

Einav, L.; Levin, J. (2014), "The Data Revolution and Economic Analysis", Innovation Policy and the Economy, Vol. 14, No. 1, pp. 1–2.

Fabian, B.; Ermakova, T.; Junghanns, P. 92015), "Collaborative and secure sharing of healthcare data in multi-clouds", Information Systems, Vol. 48, pp. 132–150.

Fleurence, R.L.; Beal, A.C.; Sheridan, S.E., et al. (20140, "Patient-powered research networks aim to improve patient care and health research", Health Affairs, Vol. 33, No. 7, pp. 1212-1219.

Frizzo-Barker, J.; Chow-White, P. (2014), "Research in brief: From patients to petabytes: Genomic big data, privacy, and informational risk", Canadian Journal of Communication, Vol. 39, No. 4, pp. 615-625.

Gehrke, J. (2012), "Quo vadis, data privacy?", Annals of the New York Academy of Sciences, Vol. 1260, No. 1, pp. 45–54.

Genov, N. (2015), "The future of individualization in Europe: changing configurations in employment and governance", European Journal of Futures Research, Vol. 2, No. 1, pp. 1–9.

Gleibs, I.H. (2014), "Turning Virtual Public Spaces into Laboratories: Thoughts on Conducting Online Field Studies Using Social Network Sites", Analyses of Social Issues and Public Policy, Vol. 14, No. 1, pp. 352–370.

Habte, M.L.; Howell, C.; Warren, A. (2015), "The Big Data Dilemma: Compliance for the Health Professional in an Increasingly Data-Driven World", Journal of Health Care Compliance, Vol. 17, No. 3, pp. 5–12.

Haggerty, K.D.; Ericson, R.V. (2000), "The surveillant assemblage", The British Journal of Sociology, Vol. 51, No. 4, pp. 605–622.

Hardin, S. (2013), "ASIS&T annual meeting plenary speaker: Edward chang: Mobile opportunities", Bulletin of the American Society for Information Science and Technology, Vol. 39, No. 3, pp. 46–48.

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp. 491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

Heffetz, O.; Ligett, K. (2014), "Privacy and Data-Based Research", The Journal of Economic Perspectives, Vol. 28, No. 2, pp. 75–98.

Helles, R.; Lomborg, S. (2013), "Regulatory response? Tracking the influence of technological developments on privacy regulation in Denmark from 2000 to 2011", Policy & Internet, Vol. 5, No. 3, pp. 289–303.

Hirsch, P.B. (2013), "Corporate reputation in the age of data nudity", Journal of Business Strategy, Vol. 34, No. 6, pp. 36–39.

Hofman, W.; Rajagopal, M. (2014), "A technical framework for data sharing", Journal of Theoretical and Applied Electronic Commerce Research, Vol. 9, No. 3, pp. 45–58.

Hogan, M.; Shepherd, T. (2015), "Information Ownership and Materiality in an Age of Big Data Surveillance", Journal of Information Policy, Vol. 5, pp. 6–31.

Holt, J.; Malčić, S. (2015), "The Privacy Ecosystem: Regulating Digital Identity in the United States and European Union", Journal of Information Policy, Vol. 5, pp. 155–178.

Hull, G. (2015), "Successful failure: what Foucault can teach us about privacy self-management in a world of Facebook and big data", Ethics and Information Technology, Vol. 17, No. 2, pp. 89–101.

Johnson, J.A. (2014), "From open data to information justice", Ethics and Information Technology, Vol. 16, No. 4, pp. 263–274.

Krishnamurthy, R.; Desouza, K.C. (2014), "Big data analytics: The case of the social security administration", Information Polity: Vol. 19, No. 3/4, pp. 165–178.

Kshetri, N. (2014), "Big data's impact on privacy, security and consumer welfare", Telecommunications Policy, Vol. 38, No. 11, pp. 1134–1145.

Kwon, O.; Lee, N.; Shin, B. (2014), "Data quality management, data usage experience and acquisition intention of big data analytics", International Journal of Information Management, Vol. 34, No. 3, pp. 387-394.

Kyunghee Y.; Hoogduin, L.; Zhang. (2015), "Big Data as Complementary Audit Evidence", Accounting Horizons, Vol. 29, No. 2, pp. 431–438.

Laat, P.B. (2014), "From open-source software to Wikipedia: "Backgrounding" trust by collective monitoring and reputation tracking", Ethics and Information Technology, Vol. 16, No. 2, pp.157–169.

Leonard, P. (2014), "Customer data analytics: Privacy settings for 'big data' business", International Data Privacy Law, Vol. 4, No. 1, pp. 53-68.

Lesley, W.S.; Shmerling, S. (2015), "Risks and Opportunities of Data Mining the Electronic Medical Record", Physician Leadership Journal, Vol. 2, No. 4, pp. 40–45.

Libaque-Sáenz, C.F.; Wong, S.F.; Chang, Y., et al. (2014), "Understanding antecedents to perceived information risks an empirical study of the Korean telecommunications market", Information Development, Vol. 32, No. 1, pp. 1-16.

Liu, D.; Wang, S. (2013), "Nonlinear order preserving index for encrypted database query in service cloud environments", Concurrency and Computation: Practice and Experience, Vol. 25, No. 13, pp. 1967–1984.

Liu, Y. (2014), "User control of personal information concerning mobile-app: Notice and consent?", Computer Law & Security Review, Vol. 30, No. 5, pp. 521–529.

Martin, K.E. (2015), "Ethical Issues in the Big Data Industry", Mis Quarterly Executive, Vol. 14, No. 2, pp. 67–85.

Matzner, T. (2014), "Why privacy is not enough privacy in the context of "ubiquitous computing" and "big data"", Journal of Information, Communication & Ethics in Society, Vol. 12, No. 2, pp. 93-106.

Mcneely, C.L.; Hahm, J. (2014), "The Big (Data) Bang: Policy, Prospects, and Challenges", Review of Policy Research, Vol. 31, No. 4, pp. 304–310.

Medina, E. (2015), "Rethinking algorithmic regulation", Kybernetes, Vol. 44, No. 6/7, pp. 1005–1019.

Mohammadpourfard, M.; Doostari, M.A.; Ghoushchi, M.B.G.; Shakiba, N. (2015), "A new secure Internet voting protocol using Java Card 3 technology and Java information flow concept", Security and Communication Networks, Vol. 8, No. 2, pp. 261–283.

Montgomery, K.C. (2015), "Youth and surveillance in the Facebook era: Policy interventions and social implications", Telecommunications Policy, Vol. 39, No. 9, pp. 771-786.

Morris, T.H.; Nair, V.S.S. (2010), "Private computing on public platforms: portable application security", Wireless Communications and Mobile Computing, Vol. 10, No. 7, pp. 942–958.

Navarro, J.M.B.; Villaverde, J.C. (2014), "The future of counter-terrorism in Europe the need to be lost in the correct direction", European Journal of Futures Research, Vol. 2, No. 1, pp. 1–12.

Nickerson, D.W.; Rogers, T. (2014), "Political Campaigns and Big Data", The Journal of Economic Perspectives, Vol. 28, No. 2, pp. 51–73.

Norberg, P.A.; Horne, D.R. (2013), "Coping with information requests in marketing exchanges: an examination of pre-post affective control and behavioral coping", Journal of the Academy of Marketing Science, Vol. 42, No. 4, pp. 415–429.

Nunan, D.; Di Domenico, M.L. (2013), "Market research and the ethics of big data", International Journal of Market Research, Vol. 55, No. 4, pp. 2–13.

Ohlhausen, M.K. (2014), "Privacy challenges and opportunities: The role of the federal trade commission", Journal of Public Policy and Marketing, Vol. 33, No. 1, pp. 4–9.

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp.491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

Olsson, N.O.E.; Bull-Berg, H. (2015), "Use of big data in project evaluations", International Journal of Managing Projects in Business, Vol. 8, No. 3, pp. 491–512.

Ossorio, P.N. (2014), "The Ethics of Translating High-Throughput Science into Clinical Practice", Hastings Center Report, Vol. 44, No. 5, pp. 8–9.

Parham, A.G.; Mooney, J.L.; Cairney, T.D. (2015), "When BYOD Meets Big Data", Journal of Corporate Accounting & Finance, Vol. 26, No. 5, pp. 21–27.

Park, Y.J.; Skoric, M. (2015), "Personalized Ad in Your Google Glass? Wearable Technology, Hands-Off Data Collection, and New Policy Imperative", Journal of Business Ethics, p. 1–12.

Parry, I.W.H.; Walls, M.; Harrington, W. (2007), "Automobile Externalities and Policies", Journal of Economic Literature, Vol. 45, No. 2, pp. 373–399.

Phillips, K.A.; Trosman, J.R.; Kelley, R.K., et al. (2014), "Genomic sequencing: Assessing the health care system, policy, and big-data implications", Health Affairs, Vol. 33, No. 7, pp. 1246-1253.

Poole, A.H. (2014), "How has your science data grown? Digital curation and the human factor: a critical literature review", Archival Science, Vol. 15, No. 2, pp. 101–139.

Porat, A.; Strahilevitz, L.J. (2014), "Personalizing Default Rules and Disclosure with Big Data", Michigan Law Review, Vol. 112, No. 8, pp. 1417–1478.

Portmess, L.; Tower, S. (2014), "Data barns, ambient intelligence and cloud computing: the tacit epistemology and linguistic representation of Big Data", Ethics and Information Technology, Vol. 17, No. 1, pp. 1–9.

Qin, B.; Wang, L.; Wang, Y., et al. (2015), "Versatile lightweight key distribution for big data privacy in vehicular ad hoc networks", Concurrency and Computation: Practice and Experience, Vol. 28, No. 10, pp. 2920-2939.

Rastogi, N.; Gloria, M.J.K.; Hendler, J. (2015), "Security and Privacy of Performing Data Analytics in the Cloud", Journal of Information Policy, Vol. 5, pp. 129–154.

Richards, N.M.; King, J.H. (2014), "Big Data Ethics", Wake Forest Law Review, Vol. 49, No. 2, pp. 393–432.

Robbin, A.; Koball, H. (2001), "Seeking explanation in theory: Reflections on the social practices of organizations that distribute public use microdata files for research purposes", Journal of the American Society for Information Science and Technology, Vol. 52, No. 13, pp. 1169–1189.

Roski, J.; Bo-Linn, G.W.; Andrews, T.A. (2014), "Creating value in health care through big data: Opportunities and policy implications", Health Affairs, Vol. 33, No. 7, pp. 1115-1122.

Rubinstein, I.S. (2013), "Big data: The end of privacy or a new beginning?", International Data Privacy Law, Vol. 3, No. 2, pp. 74-87.

Samuels, J.G.; Mcgrath, R.J.; Fetzer, S.J., et al. (2015), "Using the Electronic Health Record in Nursing Research Challenges and Opportunities", Western Journal of Nursing Research, Vol. 37, No. 10, pp. 1284–1294.

Schadt, E.E. (2012), "The changing privacy landscape in the era of big data", Molecular Systems Biology, Vol. 8, No. 1.

Schatzmann, J.; Schäfer, R.; Eichelbaum, F. (2013), "Foresight 2.0 - Definition, overview& evaluation", European Journal of Futures Research, Vol. 1, No. 1, pp. 1–15.

Schintler, L.A.; Kulkarni, R. (2014), "Big Data for Policy Analysis: The Good, The Bad, and The Ugly", Review of Policy Research, Vol. 31, No. 4, pp. 343–348.

Schnell, R. (2014), "An efficient privacy-preserving record linkage technique for administrative data and censuses", Statistical Journal of the IAOS, Vol. 30, No. 3, pp. 263–270.

Schnell, R. (2014), "An efficient privacy-preserving record linkage technique for administrative data and censuses", Statistical Journal of the IAOS, Vol. 30, No. 3, pp. 263–270.

Selinger, E.; Hartzog, W. (2015), "Facebook's emotional contagion study and the ethical problem of co-opted identity in mediated environments where users lack control", Research Ethics.

Siemelink, A.J. Digital forensics as a service: Game on. Digital Investigation. (no prelo)

Smith, S. (2014), "Data and privacy: Now you see me; New model for data sharing; Modern governance and staticians", Significance, Vol. 11, No. 4, pp. 10–17.

Stough, R.; Mcbride, D. (2014), "Big Data and U.S. Public Policy", Review of Policy Research, Vol. 31, No. 4, pp. 339–342.

Sukumar, S.R.; Natarajan, R.; Ferrell, R.K. (2015), "Quality of Big Data in health care", International Journal of Health Care Quality Assurance, Vol. 28, No. 6, pp. 621–634.

Taylor, L.; Cowls, J.; Schroeder, R., et al. (2014), "Big Data and Positive Change in the Developing World", Policy & Internet, Vol. 6, No. 4, pp. 418–444.

Taylor, L. (2015), "No place to hide? The ethics and analytics of tracking mobility using mobile phone data", Environment and Planning D: Society and Space, Vol. 34, No. 2, pp. 319-336.

Terry, N. (2015), "Navigating the Incoherence of Big Data Reform Proposals", The Journal of Law, Medicine & Ethics, Vol. 43, No. s1, pp. 44–47.

Tse, J.; Schrader, D.E.; Ghosh, D., et al. (2015), "A bibliometric analysis of privacy and ethics in IEEE Security and Privacy", Ethics and Information Technology, Vol. 17, No. 2, pp. 153–163.

Ulltveit-Moe, N.; Oleshchuk, V. (2015), "A novel policy-driven reversible anonymisation scheme for XML-based services", Information Systems, Vol. 48, pp. 164–178.

Revista Eletrônica Sistemas & Gestão
Volume 12, Número 4, 2017, pp. 491-505
DOI: 10.20985/1980-5160.2017.v12n4.1205

Ulltveit-Moe, N. (2014), "A roadmap towards improving managed security services from a privacy perspective", Ethics and Information Technology, Vol. 16, No. 3, pp. 227-240.

Unsworth, K. (2014), "Questioning trust in the era of big (and small) data", Bulletin of the American Society for Information Science and Technology, Vol. 41, No. 1, pp. 12–14.

Van Den Hoven, J.; Weckert, J. (2008), Information technology, privacy, and the protection of personal data, Information Technology and Moral Philosophy, Cambridge University Press, Cambridge.

Varian, Hal R. (2014), "Beyond big data", Business Economics, Vol. 49, No. 1, pp. 27–31.

Vezyridis, P.; Timmons, S. (2015), "On the adoption of personal health records: some problematic issues for patient empowerment", Ethics and Information Technology, Vol. 17, No. 2, pp. 113–124.

Wang, X.; Liang, Q.; Mu, J., et al. (2015), "Physical layer security in wireless smart grid", Security and Communication Networks, Vol. 8, No. 14, pp. 2431–2439.

Weaver, S.D.; Gahegan, M. (2007), "Constructing, visualizing, and analyzing a digital footprint", Geographical Review, Vol. 97, No. 3, pp. 324–350.

Zhang, X.; Liu, C.; Nepal, S., et al. (2013), "SaC-FRAPP: a scalable and cost-effective framework for privacy preservation over big data on cloud", Concurrency and Computation: Practice and Experience, Vol. 25, No. 18, pp. 2561–2576.

Zhou, W.; Piramuthu, S. (2014), "Information Relevance Model of Customized Privacy for IoT', Journal of Business Ethics, Vol. 131, No. 1, pp. 19–30.

Zuboff, S. (2015), "Big other: Surveillance capitalism and the prospects of an information civilization", Journal of Information Technology, Vol. 30, No. 1, pp. 75–89.