

Repositório ISCTE-IUL

Deposited in *Repositório ISCTE-IUL*:

2021-06-15

Deposited version:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Chaves, A., Moura, I., Bernardino, J. & Pedrosa, I. (2020). The privacy paradigm : An overview of privacy in Business Analytics and Big Data. In Álvaro Rocha, Bernabé Escobar Peréz, Francisco Garcia Peñalvo, Maria del Mar Miras, Ramiro Gonçalves (Ed.), 2020 15th Iberian Conference on Information Systems and Technologies (CISTI). Sevilla: IEEE.

Further information on publisher's website:

[10.23919/CISTI49556.2020.9140831](https://doi.org/10.23919/CISTI49556.2020.9140831)

Publisher's copyright statement:

This is the peer reviewed version of the following article: Chaves, A., Moura, I., Bernardino, J. & Pedrosa, I. (2020). The privacy paradigm : An overview of privacy in Business Analytics and Big Data. In Álvaro Rocha, Bernabé Escobar Peréz, Francisco Garcia Peñalvo, Maria del Mar Miras, Ramiro Gonçalves (Ed.), 2020 15th Iberian Conference on Information Systems and Technologies (CISTI). Sevilla: IEEE., which has been published in final form at <https://dx.doi.org/10.23919/CISTI49556.2020.9140831>. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

O paradigma da privacidade

Um panorama geral sobre privacidade em *Business Analytics* e *Big Data*

The privacy paradigm

An overview of privacy in Business Analytics and Big Data

Aldair Chaves

Coimbra Business School |
ISCAC, Polytechnic of
Coimbra

Coimbra, Portugal

a2019106111@alumni.iscac.pt

Ítalo Moura

Coimbra Business School |
ISCAC, Polytechnic of
Coimbra,

Coimbra, Portugal

a2019123514@alumni.iscac.pt

Jorge Bernardino

Instituto Politécnico de
Coimbra – ISEC

i2A – Instituto de

Investigação Aplicada

Coimbra, Portugal

jorge@isec.pt

Isabel Pedrosa

Coimbra Business School |
ISCAC, Polytechnic of
Coimbra

Instituto Universitário de

Lisboa (ISCTE-IUL)

ISTAR-IUL, Portugal

ipedrosa@iscac.pt

Resumo — Nesta Nova Era, onde a informação tem um valor indispensável para as empresas e as tecnologias de mineração de dados crescem na área de Tecnologia da Informação, a privacidade continua a ser um assunto delicado na abordagem da exploração do grande volume de dados gerados e tratados pelas empresas. A forma com que os dados são recolhidos, tratados e destinados ainda não está claramente definida e é tema de constante debate em diversas áreas de atuação. Esta revisão de literatura traça um panorama sobre a privacidade na era do *Business Analytics* e *Big Data*, em diferentes linhas temporais, para ampliar discussões sobre um assunto que merece extrema atenção e tem como objetivo demonstrar que, apesar de terem sido criadas medidas para proteção de dados, há ainda muito a discutir entre as diferentes partes atuantes no processo, para se chegar a um ideal positivo tanto para utilizadores, quanto para as empresas.

Palavras Chave – *Big Data*, *privacidade*, *analytics*, *dados*, *RGPD*.

Abstract — In this New Age where information has an indispensable value for companies and data mining technologies are growing in the area of Information Technology, privacy remains a sensitive issue in the approach to the exploitation of the large volume of data generated and processed by companies. The way data is collected, handled and destined is not yet clearly defined and has been the subject of constant debate by several areas of activity. This literature review gives an overview of privacy in the era of *Business Analytics* and *Big Data* in different timelines, the opportunities and challenges faced, aiming to broaden discussions on a subject that deserves extreme attention and aims to show that, despite measures for data protection have been created, there is still a need to discuss the subject among the different parties involved in the process to achieve a positive ideal for both users and companies.

Keywords - *Big Data*, *privacy*, *analytics*, *dados*, *GPDR*.

I. INTRODUÇÃO

A evolução da tecnologia alcançou patamares onde a informação tem valor inestimável e é utilizada para gerar uma infinidade de dados relevantes tanto para empresas, quanto para os utilizadores. A convergência das tecnologias que emergiram nos últimos anos fez com que a quantidade de detalhes e

possibilidades que podem ser exploradas através da avalanche de informação gerada a cada segundo, se tornasse um atrativo mercado e um fator definidor de posição e expansão [1] [X]. O fenómeno das redes sociais, os mais variados *gadgets*, telemóveis e seus aplicativos e *e-mails*, aliados às novas tendências no tratamento de um massivo volume de dados, trazem uma nova realidade num contexto democrático onde os direitos à privacidade e a evolução tecnológica estão a começar a entrar em conflito [2] [3].

A privacidade dos dados dos utilizadores e a forma como são utilizados ficaram em evidência e continuam a ser um foco de preocupação na atual era do *Big Data* e *Analytics*[4] [5]. As pessoas não têm acesso, de forma clara, a todas regras e detalhes de como seus dados são usados e como são conduzidos os estudos para os quais aceitamos ceder os nossos dados [6] [7]. E, embora haja um discurso benéfico sobre a utilização de dados e informação como uma forma de agregar valor ao quotidiano dos utilizadores, deve-se promover uma abordagem clara no que diz respeito às tecnologias envolvidas nas práticas de tratamento de dados para que não seja gerado ruído e entraves à evolução tecnológica.

O intuito deste trabalho é realizar uma revisão de literatura sobre a privacidade na utilização do *Business Analytics* e *Big Data*, além de sistematizar as oportunidades e desafios associados à Privacidade no contexto abordado.

Este trabalho está dividido em seis secções. Na primeira está a Introdução, trazendo o tema abordado e sua relevância. Na segunda está apresentado o Método de pesquisa, que traz sustentação ao estudo efetuado. Na terceira, um breve enquadramento sobre *Big Data*, *Business Analytics* e *RGPD*. Na quarta estão apresentados os Resultados das pesquisas. Na quinta, é levantada uma Discussão sobre a temática. E na sexta e última secção, está apresentada a Conclusão sobre a investigação.

II. MÉTODO DE PESQUISA

Este estudo foi concebido através de buscas eletrônicas para levantamento e escolha de publicações que abordam o tema de privacidade de dados, *Business Analytics* e *Big Data* utilizando

bases de dados multidisciplinares. As bases de dados que funcionaram como base da pesquisa foram: I) *Xplore Digital Library* (IEEE), II) Biblioteca do conhecimento online (*B-on*), III) Google Acadêmico (*Google Scholar*) e IV) *Computer Science Bibliography* DBLP. A estratégia utilizada para reunião dos materiais foi a busca utilizando os temas “*Big Data e privacidade*”, “*privacidade em Big Data*”, “*privacidade em Business Analytics and Big Data*”, “*Business Analytics, Big Data e privacidade*” no idioma português, no caso das plataformas de busca *Google Scholar* e *B-on*. No idioma inglês foram utilizadas buscas com palavras-chave “*Big Data, business analytics and privacy*”, “*privacy in Business Analytics and Big Data*” e “*privacy and protection in business analytics*”.

Alguns critérios foram estabelecidos para que o objetivo da pesquisa fosse atingido de forma satisfatória. Primeiramente, foi notada a escassez de artigos com esta temática, então foi aberto um leque de busca multidisciplinar, onde as Tecnologias da Informação não fosse somente o foco. O segundo critério foi a imposição de uma delimitação temporal em duas etapas, onde foram realizadas pesquisas para a comparação das discussões sobre o tema antes e depois de 2018. Esta divisão temporal foi criada para que pudéssemos ter uma ideia sobre as discussões antes e depois de acontecimentos relativos a perda de dados pessoais por grandes empresas como o Facebook e Sony, que protagonizaram escândalos que levaram à criação e aplicação de leis de proteção à privacidade como a GDPR (*General Data Protection Regulation*).

III. ENQUADRAMENTO

Para um acompanhamento mais completo do estudo feito nesta revisão de literatura, é necessário a abordagem para explicação de alguns termos utilizados no decorrer do trabalho.

Primeiramente, falaremos de *Big Data*. Apesar de se tratar de um termo que não é novo, trata-se, por definição, de um grande volume de dados, estruturados ou não, que tem sido utilizado para definir os dados gerados a partir de tecnologias, aplicativos, *sites web* entre outros [8], [9].

A segunda expressão é *data analytics*, que compreende, basicamente, a análise dos dados. Este processo permite que os dados sejam examinados de forma a que conclusões possam ser retiradas sobre a informação gerada, sendo bastante utilizado por empresas para auxílio na tomada de decisão e posição nos negócios [10] [11][W]. E, por fim, outro termo bastante utilizado é o GDPR (*General Data Protection Regulation*). Com o passar dos anos os países viram a necessidade de criar ou modernizar leis específicas para a proteção de dados, como é o caso do GDPR na Europa, que se trata de uma lei de proteção à privacidade das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação dos mesmos [12]. No Brasil a lei que supre essa necessidade é a LGPD, em Portugal é a RGPD.

IV. RESULTADOS

Obedecendo os critérios estabelecidos na metodologia de pesquisa e após iniciar a busca pelos artigos, foi realizada a leitura dos títulos e resumos como forma de analisar se os trabalhos estavam de acordo com o tema, excluindo documentos que tinham foco apenas numa determinada região ou que, de alguma forma, não se encaixavam na ideia principal da pesquisa.

Foram escolhidos então sete artigos que traziam abordagens relevantes sobre o tema.

Para analisar os trabalhos escolhidos, foi executada a leitura na íntegra dos artigos e a confirmação do enquadramento de cada um. Os artigos escolhidos serão descritos adiante, nas suas ideias principais e abordagens relevantes ao tema proposto.

Para um melhor entendimento, a Figura1 foi criada com o objetivo de mostrar o sistema criado para tratamento e seleção dos artigos selecionados.

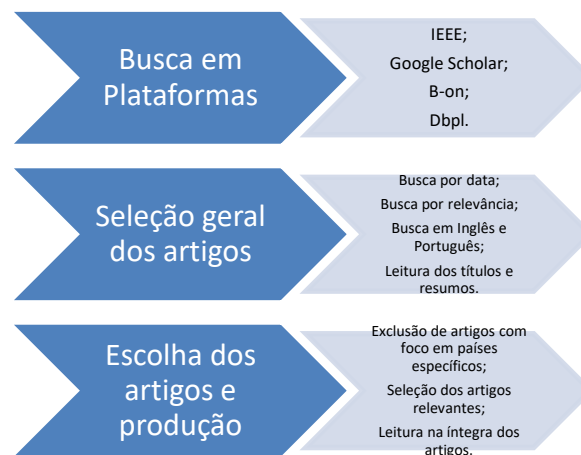


Figura 1 - Mapa mental de triagem dos trabalhos

Quanto às temáticas abordadas, todos os artigos escolhidos trazem uma visão geral sobre privacidade dos dados no contexto de *Business Analytics* e *Big Data*, abordando os problemas relacionados com privacidade de dados no contexto da atual sociedade, do grande volume de dados e a maioria apresentam sugestões teóricas sobre o que poderia ser praticado para um resultado satisfatório na questão da privacidade. Para um melhor entendimento, a Tabela 1 mostra a área de atuação de cada autor.

Tabela 1 - Lista dos artigos analisados

Título	Autor	Data	Área	País
Big Data Security and Privacy Protection	Dongpo Zhang	2018	IT	China
Big Data e Proteção de Dados: Uma Relação Possível(?)	Márcio M. Szinvelski, Taynara Silva Arceno & Lucas Baratieri Francisco	2019	Law	Brasil
Big Data e Open Data vs. Privacidade: Constituiria a Ferramenta Tecnológica em uma Invasão à Privacidade?	Maria Salles Coelho Mello Franco	2016	Law	Brasil
Freedom under the gaze of Big Brother: Preparing the grounds for a liberal defence of privacy in the era of Big Data	Henrik Skaug Sætra	2019	Politics	Noruega
Privacy preservation techniques in big data analytics: a survey	P. Ram Mohan Rao, S. Murali Krishna & A. P. Siva Kumar	2018	IT	Índia
Big Data for All: Privacy and User Control in the Age of Analytics	Omer Tene & Jules Polonetsky	2013	Law	Israel
Privacy in the Age of Big Data: A Time for Big Decisions	Omer Tene & Jules Polonetsky	2012	Law	Israel

Visando um acompanhamento melhor dos resultados da pesquisa, estes foram divididos em três categorias, de forma que poderemos acompanhar os assuntos mais relevantes abordados em cada artigo analisado. As categorias são: (a) preocupação com a privacidade; (b) benefícios do *Big Data* e *Business Analytics*; (c) soluções propostas e (d) síntese dos resultados.

A. Preocupações com a privacidade

A preocupação com a privacidade é um assunto referido constantemente em todos os trabalhos pesquisados de acordo com a estratégia definida, e mostra que, independentemente da discussão sobre a forma de utilização dos dados, é algo que precisa de ser avaliado com atenção. Segundo as ideias apresentadas por Zhang [13] ao mesmo tempo que o *Big Data* vem beneficiando as pessoas e seus estilos de vida, também traz preocupações referentes à privacidade das pessoas. Para o autor, o *Big Data* enfrenta alguns desafios no que diz respeito à segurança no tratamento dos dados, como a privacidade propriamente dita, a credibilidade da informação, a falta de mecanismos para proteção, privacidade e segurança dos dados.

Szinvelki, Arceno, Francisco [14] trazem uma visão mais jurídica no âmbito de proteção de dados e privacidade, onde há o desafio de harmonizar entre o direito e a tecnologia. Os autores abordam acontecimentos recentes que fomentaram (e ainda continuam a alimentar) debates sobre a utilização da tecnologia no processamento e manipulação de dados de forma nociva e que culminam na violação dos direitos de privacidade dos indivíduos. Além disso, sustentam a ideia de que os dados podem ser manipulados, e, com isso, também o serão pessoas, comportamentos e tendências, fazendo com que não saibamos se está a existir evolução cultural própria ou induzida. Esta visão é abordada, tendo como base a falta de transparência e clareza em relação as informações pessoais, desde a forma de consentimento, até o destino da informação.

Na pesquisa de Sætra [15] é efetuada a analogia entre *Big Data* e *Big Brother*, onde são levantadas questões a respeito da ameaça de privacidade na era do *Big Data*, formas de vigilância impostas sobre a informação que é disponibilizada, direta ou indiretamente, uma visão geral das alternativas (ou da falta delas) dadas com base na vida dos indivíduos e as interações tecnológicas. Além dessas analogias, o autor traça paralelos para discussão sobre o direito das pessoas à liberdade e ao anonimato, relacionando cenários de perda, roubo de informações e outros acontecimentos que podem significar na distribuição indevida de dados pessoais uma ameaça iminente. Para este autor, a evolução da tecnologia e a relação com o mundo *online* interferem nas nossas escolhas, pois a não aceitação às vezes implica estar fora de alguns nichos sociais e profissionais. Embora afirme que as tecnologias relacionadas com *Big Data* são neutras, culpa a forma com que são utilizadas, fazendo uma relação entre *Big Data* e vigilância, elucidando a forma com que esse comportamento pode ser nocivo, impactando na liberdade e privacidade das pessoas. O paradoxo entre *Big Data* e privacidade é um dos assuntos mais urgentes que devem ser tratados por governantes face a uma banalização do significado de liberdade quando são comparados com os benefícios trazidos pelo *Big Data*.

O trabalho realizado por Franco [16] faz o levantamento de alguns problemas relativos à utilização dos dados e privacidade, embora destaque a má utilização dos dados por parte das empresas. Alguns aspetos elencados pela autora já estão a receber atenção atualmente em alguns países, como a identificação do destino dos dados e formulários de consentimento, ao mesmo tempo que são trazidas colocações pertinentes quanto à falta de clareza nesses processos e a importância disso para que não haja excesso no âmbito jurídico, para não promover uma desaceleração tecnológica.

Na pesquisa de Tene, Polonetsky [17] é sugerido que a economia global está a ser impulsionada pelos avanços tecnológicos através da inovação e eficiência, mas ao mesmo tempo cria preocupação relativamente à privacidade, gerando conflito entre a inovação com a regulamentação de proteção de dados. O aumento na geração de dados através de novos dispositivos e mais pessoas utilizando as ferramentas digitais, aliados as novas tecnologias de *data mining*, análise de dados, armazenamento e poder computacional são os principais causadores deste conflito.

Abordando de uma forma um pouco diferente dos anteriores, Rao, Krishna, Kumar [18] afirmam que a análise de dados reflete na privacidade dos utilizadores, mas difere no sentido de que afirma também que parte da culpa no vazamento dos dados pessoais é por parte dos utilizadores, especialmente dos telemóveis, nos quais utilizam diversas aplicações e aceitam diferentes termos de privacidade sem saber ao certo o que está consentindo. Os autores são enfáticos ao dizer que os utilizadores precisam ser educados neste sentido, mas não exclui a necessidade da criação de políticas voltadas para a proteção dos dados. Outra diferente abordagem, é quando sugerem que a responsabilidade da privacidade dos utilizadores é de responsabilidade de quem está com as informações e são responsáveis por garantir a privacidade dos dados.

Num segundo trabalho no campo estudado, Tene, Polonetsky [19] sugerem que o *Big Data* se mostra cada vez mais relevante para os negócios, governos e na qualidade de vida das pessoas. A discussão sobre a proteção de dados individuais, principalmente por governantes e defensores da privacidade, aumenta, acompanhando o crescimento da recolha de dados, os usos cada vez mais diversificados das suas aplicações, sendo proporcionados pela melhoria e redução do custo da tecnologia, aprimoramento das técnicas e armazenamento quase ilimitado.

B. Benefícios do Business Analytics e Big Data

A evolução dos sistemas tecnológicos trouxe diversas possibilidades e são indiscutíveis os benefícios trazidos pelo *Big Data* e a análise de dados para a tomada de decisões, fazendo com que seja possível dinamizar relações sociais, profissionais e culturais. A pesquisa de Franco [16] destaca a importância da utilização da análise de dados e benefícios trazidos na área da saúde, prevenindo a proliferação de doenças e como poderiam ser utilizados em setores para amenizar crises económicas, planeamento energético e trânsito.

Segundo Tene, Polonetsky [17], diversos são os exemplos que demonstram a relevância da utilização analítica de dados gerados pela sociedade como a prevenção de acontecimentos, correção de falhas e ou melhorar serviços e traz à tona a discussão sobre a efetividade da minimização do nível de informação na recolha de dados.

Num outro estudo, Tene, Polonetsky [19] focaram-se em demonstrar as vantagens do uso do *Big Data* em aplicações diversas em vários setores da sociedade, como: lojas, saúde, meio ambiente, entre outros. Expuseram também as preocupações com privacidade, balanceando os argumentos e tentando encontrar o meio termo entre o retorno sobre a tecnologia e o risco da exposição e controlo dos dados. A defesa sobre o uso do *Big Data* baseia-se numa projeção do que se pode conseguir com a utilização das informações, de uma forma ainda não imaginada, mas que não será conseguido se não possuir o acesso e armazenamento dos dados.

C. Soluções propostas

Uma preocupação geral para todos os investigadores é de sugerir formas que possam garantir a privacidade dos utilizadores. Como mostra na pesquisa de Zhang [13], algumas medidas que poderiam melhorar a questão da privacidade, como uma forma de supervisão dos dados nas redes sociais, o aprimoramento de mecanismos legais de proteção de dados e a criação de um órgão para regulamentar as questões referentes a privacidade. Além das medidas mencionadas, é enfatizada também a importância da conscientização dos utilizadores como forma de contribuir para que mecanismos de segurança e privacidade sejam implementados.

Para Szinvelki, Arceno, Francisco [14], é indispensável que, para a manutenção da democracia e direitos das pessoas, exista transparência sobre quais dados que são recolhidos, métodos de recolha de dados, finalidade, programas de segurança e tolerância a falhas.

Na sua pesquisa, Franco [4] traz também algumas sugestões para melhora do ponto de vista tecnológico e jurídico. A primeira ideia da autora é transformar o *Big Data* em *Open Data*, que basicamente seriam os mesmos dados que são recolhidos e

tratados pelas empresas, só que públicos. Para ela, através deste conceito, as pessoas teriam acesso à sua informação e o conhecimento estariam ao alcance de todos. Do ponto de vista jurídico, primeiramente a autora afirma que é necessário a criação de medidas jurídicas, mas com cautela, haja vista que as leis evoluíram mais devagar que a tecnologia. E, em segundo, sugere a criação de leis baseadas em análises a respeito das opiniões dos utilizadores em relação ao que se acha invasivo.

Tene, Polonetsky [17] sugerem que as políticas de privacidade, juntamente com a expectativa dos utilizadores de se sentirem seguros, não é de toda forma tão eficaz, uma vez que políticas de privacidade muitas vezes não garantem o direito esperado, e sim uma garantia de não culpabilidade sobre as informações passadas. Além de que toda essa discussão sobre privacidade dos dados que está sendo gerada, pode criar crenças entre as pessoas e contrastar com os possíveis benefícios da análise de dados. Por fim, afirmam que não se trata de ser contra os direitos e privacidade dos utilizadores, mas defendem que, na medida em que cresce a recolha e utilização dos dados para soluções inovadoras e que tem impacto positivo, deve-se exigir a criação de um modelo de equilíbrio entre a privacidade e os benefícios da recolha de dados.

Rao, Krishna, Kumar [18], apresentam várias técnicas de preservação de dados como criptografia e ocultação de dados, explicando que nenhuma delas consegue garantir de forma concreta a efetiva proteção dos dados e pode ocorrer perda de dados e informação. Por fim, é apresentada uma técnica baseada em *Deep Learning* para aplicação de um modelo para melhorar a privacidade dos dados. Neste modelo, após a recolha dos dados, é utilizado um *software* para classificar os dados sensíveis (informações pessoais dos utilizadores) e um algoritmo baseado em *Deep Learning* é utilizado para separar tais dados, protegê-los dinamicamente, fazendo com que possam ser publicados sem nenhum dado sensível. Ainda segundo estes autores, este algoritmo seria constantemente treinado e melhorado, adicionando camadas adicionais para deteção e proteção das informações sensíveis.

Tene, Polonetsky [19] argumentam que o uso transparente no processamento de dados inibirá o seu uso sem ética, pois já foi provado por investigadores que quanto maior o número de dados e com o compromisso necessário os dados podem ser interligados, diminuindo assim a privacidade. Os indivíduos se sentirão mais dispostos a aceitar a partilha e recolha de dados de acordo com os benefícios que possam advir dessa partilha. Em contrapartida, isso fará com que as empresas trabalhem para usar os dados e trazer esse retorno às pessoas.

Para sistematizar os principais argumentos dos autores pesquisados, apresenta-se a Tabela 2, onde se enumeram os desafios e oportunidades enfatizados em cada um dos estudos.

Tabela 2 - Síntese dos resultados

Desafios	Oportunidades	Autor
<ol style="list-style-type: none"> 1. Proteção dos dados privados. 2. Credibilidade das informações 3. Falta de mecanismos de proteção. 	<ol style="list-style-type: none"> 1. Supervisão de dados dos utilizadores. 2. Aprimoramento dos mecanismos de proteção de dados, órgãos reguladores de leis. 	Dongpo Zhang
<ol style="list-style-type: none"> 1. Harmonia entre direito e tecnologia 2. Manipulação de opinião pública. 3. Falta de transparência no tratamento dos dados 	<ol style="list-style-type: none"> 1. Transparência quanto a coleta e tratamento de dados, a finalidade, programas de segurança e correção a falhas. 	Mártin M. Szinvelski, Taynara Silva Arceno, Lucas Baratieri Francisco
<ol style="list-style-type: none"> 1. Formas de controle e vigilância 2. Ameaça à liberdade das pessoas. 3. A falta de alternativas a não aceitar as políticas 	—	Henrik Skaug Sætra
<ol style="list-style-type: none"> 1. Má utilização dos dados por empresas 2. Falta de clareza 3. Cuidado com as leis para não gerar desaceleração tecnológica 	<ol style="list-style-type: none"> 1. <i>Open Data</i> 2. Criação de leis para proteção; 3. Medidas baseadas na opinião pública. 	Maria Salles Coelho Mello Franco
<ol style="list-style-type: none"> 1. Privacidade impactando nas inovações tecnológicas. 2. Importância da análise de dados para previsão de acontecimentos, melhoria de serviços ou correção de falhas. 3. Dúvida na efetividade da minimização na coleta de dados. 4. <i>Big Data</i> mais voltados para qualidade de vida das pessoas 	<ol style="list-style-type: none"> 1. Maiores debates para melhores definições sobre privacidade; 2. políticas de privacidade em paralelo com as expectativas dos usuários. 3. Transparência no processamento dos dados 	Omer Tene, Jules Polonetsky
<ol style="list-style-type: none"> 1. Análise de dados reflete em privacidade. 2. Problemas com privacidade também culpa dos utilizadores 	<ol style="list-style-type: none"> 1. Criação de políticas de proteção 2. Utilização de algoritmo baseado em <i>Deep Learning</i> 	P. Ram Mohan Rao, S. Murali Krishna, A. P. Siva Kumar

V. DISCUSSÃO

A partir das ideias dos autores, é importante perceber que não deve haver uma oposição sobre a evolução e tendências tecnológicas, mas é notório que medidas gerais precisam de ser tomadas e as que já existem precisam ser adaptadas, de forma que não haja prejuízo para os utilizadores das mais diversas tecnologias, nem para as empresas de *data mining* e suas vertentes [20]. Claro que alguns acontecimentos recentes envolvendo perda de dados já fizeram com que algumas entidades internacionais tomassem medidas com o intuito de preservar os direitos das pessoas. Porém, as sanções e leis aplicadas ainda não estão claras no ponto de vista dos utilizadores, gerando desconforto e aumentando preocupações e especulações relacionadas com o controlo, monitorização e influência da opinião pública. Hoje já se fala de algoritmos inteligentes desenvolvidos por empresas que têm como função tratar os dados de forma que as informações dos utilizadores sejam protegidas através de criptografia, principalmente no que diz respeito à sua identidade ou dados que vão de encontro de diretivas de acesso atuais, mas é necessário que as pessoas percecionem uma maior clareza quanto às medidas de segurança.

Mecanismos mais complexos do que apenas “aceitar” ou “não aceitar” (ação que muitas vezes provoca o não acesso a

determinados conteúdos) precisam de ser criados. Talvez algoritmos certificados por entidades especializadas em segurança de dados (que, se não existem, poderiam ser criadas) pudesse criar modelos gerais ou certificar os modelos de algoritmos das empresas, assim como acontecem com os certificados de segurança dos sites, trazendo mais segurança para as pessoas que, hoje em dia, estão cada vez mais preocupadas com a forma como as técnicas de análise de dados evoluíram e com os resultados que podem trazer.

VI. CONCLUSÕES

Uma abordagem ao debate sobre privacidade na era do *Big Data* e *Business Analytics* sob o ponto de vista de diferentes áreas de atuação foi um dos principais propósitos deste trabalho. Além de trazer os aspetos mais importantes encontrados nos artigos pesquisados, abordando a privacidade na utilização do *Business Analytics* e *Big Data*, além de sistematizar as oportunidades e desafios associados à privacidade no contexto abordado. Através desta pesquisa, é possível perceber que a relevância do assunto privacidade ainda está presente e vem crescendo, principalmente após casos de perda de dados e criação de leis para proteção de dados.

Percebe-se a necessidade de manter e expandir a discussão sobre privacidade e o direito das informações veiculadas, de forma a se gerar uma visão crítica sobre a temática como uma forma de agir como colaboradores para a evolução da tecnologia e os benefícios trazidos pelas informações que geramos, tendo a certeza de que não está em causa a perda da nossa individualidade.

Por estes motivos, considera-se fundamental aproximar tecnologia e utilizadores, fazendo com que ambos possam coexistir sem a visão de uma sombra obscura e um cenário de controlo apocalíptico. Para além disso, acredita-se que a fórmula perfeita para a problemática apresentada ainda é complexa, devido à rápida evolução de tecnologias e fatores que envolvem análise de dados. Os debates trarão maturidade para um consenso.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] P. Neves and J. Bernardino, “Big data issues,” in *ACM International Conference Proceeding Series - IDEAS'15 July 13–15, 2015, Yokohama, Japan*, 2015, pp. 200–201.
- [2] Y. Gahi, M. Guennoun, and H. Mouftah, “Big Data Analytics: Security and privacy challenges,” in *2016 IEEE Symposium on Computers and Communication (ISCC)*, Messina, Italy, 2016, pp. 952–957.
- [3] P. Jain, M. Gyanchandani, and N. Khare, “Big data privacy: a technological perspective and review,” *J. Big Data*, vol. 3, no. 1, 2016.
- [4] J. Bernardino and P. C. Neves, “Decision-making with big data using open source business intelligence systems,” in *Human Development and Interaction in the Age of Ubiquitous Technology*, 2016, pp. 120–147.
- [5] J. Almeida, P.D.C.D., Bernardino, “Big Data Open Source Platforms,” in *2015 IEEE International Congress on Big Data, BigData Congress 2015*, 2015, pp. 268–275.
- [6] Tamane, S., V. K. Solanki, and N. Dey, *Privacy and security policies in big data*. 2017.
- [7] R. J. Santos, J. Bernardino, and M. Vieira, “Balancing security and performance for enhancing data privacy in data warehouses,” in *Proc. 10th IEEE Int. Conf. on Trust, Security and Privacy in Computing and Communications, TrustCom 2011, 8th IEEE Int. Conf. on*

- Embedded Software and Systems, ICESS 2011, 6th Int. Conf. on FCST 2011*, 2011, no. November, pp. 242–249.
- [8] M. Chen, S. Mao, and Y. Liu, “Big data: A survey,” *Mob. Networks Appl.*, vol. 19, no. 2, pp. 171–209, 2014.
- [9] J. Bernardino and P. C. Neves, “Big Data in the Cloud,” *Open J. Big Data*, vol. 1, no. 2, pp. 507–536, 2018.
- [10] K. Kambatla, G. Kollias, V. Kumar, and A. Grama, “Trends in big data analytics,” *J. Parallel Distrib. Comput.*, vol. 74, no. 7, pp. 2561–2573, 2014.
- [11] L. Duan and Y. Xiong, “Big data analytics and business analytics,” *J. Manag. Anal.*, vol. 2, no. 1, pp. 1–21, 2015.
- [12] P. Voigt and A. Von dem Bussche, *The EU General Data Protection Regulation (GDPR). A Practical Guide*. 2017.
- [13] D. Zhang, “Big Data Security and Privacy Protection,” in *2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS)*, 2019, vol. 77, pp. 87–89.
- [14] M. M. Szinvelski, T. S. Arceno, and L. B. Francisco, “Big data e proteção de dados: uma relação possível(?),” in *5º Congresso Internacional de Direito e Contemporaneidade*, 2019, p. 15.
- [15] H. S. Sætra, “Freedom under the gaze of Big Brother: Preparing the grounds for a liberal defence of privacy in the era of Big Data,” *Technol. Soc.*, vol. 58, no. July, p. 101160, 2019.
- [16] M. S. C. Mello Franco, “Big Data e Open Data Vs. Privacidade: Constituiria a Ferramenta Tecnológica em uma Invasão à Privacidade?,” *Rev. Estud. Jurídicos UNESP*, vol. 20, no. 32, pp. 295–308, 2018.
- [17] O. Tene and J. Polonetsky, “PRIVACY IN THE AGE OF BIG DATA: A TIME FOR BIG DECISIONS,” *STANFORD LAW Rev. ONLINE*, vol. 64, pp. 63–69, 2012.
- [18] P. Ram Mohan Rao, S. Murali Krishna, and A. P. Siva Kumar, “Privacy preservation techniques in big data analytics: a survey,” *J. Big Data*, vol. 5, no. 1, 2018.
- [19] O. Tene and J. Polonetsky, *Big Data for All: Privacy and User Control in the Age of Analytics*, vol. 11, no. 5. 2013.
- [20] L. C. Borges, V. M. Marques, and J. Bernardino, “Comparison of data mining techniques and tools for data classification,” in *2013 - ACM International Conference Proceeding Series*, 2013, pp. 113–116.