Informação e Direito à Privacidade no contexto do Big Data: Uma Revisão Sistemática da Literatura (2012-2022)

Vinícius Barros Caixeta

Dissertação de Mestrado em Gestão e Curadoria da Informação

Setembro de 2023

Informação e Direito à Privacidade no contexto do Big Data: Uma Revisão Sistemática da Literatura (2012-2022)

Vinícius Barros Caixeta

Dissertação de Mestrado em Gestão e Curadoria da Informação

Orientadora: Professora Doutora Paula Alexandra Ochôa de Carvalho Telo

Coorientador: Professor Doutor Roberto Henriques

Setembro de 2023

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Gestão e Curadoria da Informação realizada sob a orientação científica da Professora Doutora Paula Alexandra Ochôa de Carvalho Telo e do Professor Doutor Roberto Henriques

AGRADECIMENTOS

Dedico esta dissertação de conclusão de mestrado a todos aqueles que contribuíram de forma significativa para a minha jornada acadêmica e profissional, em especial à minha família, aos meus valiosos orientadores, aos amigos e colegas de trabalho que estiveram ao meu lado.

À minha família, expresso minha profunda gratidão por seu apoio ao longo desta jornada. À minha companheira, que compartilhou comigo os desafios e triunfos desta caminhada, sua paciência e compreensão motivaram-me a persistir, e seu amor deu significado a cada passo dado em direção a esta realização. Agradeço à minha mãe em especial, seu apoio incondicional e amor foram fundamentais para que eu pudesse chegar até aqui. Seu exemplo de determinação e seu constante incentivo foram o alicerce sobre o qual construí o meu percurso acadêmico. Esta conquista é também sua, pois a sua participação e encorajamento estiveram sempre presentes.

Aos meus professores orientadores, Professora Doutora Paula Alexandra Ochôa de Carvalho Telo e Professor Doutor Roberto Henriques, cuja orientação e conhecimento foram fundamentais para o desenvolvimento deste estudo, sou imensamente grato. Este trabalho é uma reflexão direta da sua orientação e investimento em meu crescimento acadêmico.

Que esta dedicação transmita minha profunda apreciação por todos aqueles que desempenharam um papel nessa jornada. Seu apoio, orientação e amizade são as fundações que me permitem avançar e contribuir para um mundo digital mais seguro e ético.

Informação e Direito à Privacidade no contexto do Big Data: Uma Revisão Sistemática da Literatura (2012-2022)

Vinícius Barros Caixeta

Resumo

O papel do Big Data na Quarta Revolução Industrial ultrapassa as grandes quantidades de dados, provenientes de diferentes fontes e formatos, tornando evidentes padrões e possibilitando automatizar os processos de decisões, agregando, assim, alto valor económico e social. Com a proliferação da internet e a popularização da internet das coisas (IoT), a quantidade de dados aumentou as informações disponíveis para entidades públicas e privadas. Esta massificação dos dados, embora traga benefícios nas áreas económicas e sociais, levanta muitas questões éticas a ser debatidas e investigadas na área da Gestão e Curadoria de Informação. Uma das principais questões éticas diz respeito à privacidade, isto é, a ameaça à privacidade resultante do aumento e integração de informações pessoalmente identificáveis. Com base neste contexto, a presente dissertação utilizou a metodologia de Revisão Sistemática da Literatura com o objetivo principal de entender como a literatura especializada tem tratado o acesso à informação e o direito à privacidade no contexto do Big Data. Foi possível identificar seis assuntos principais, baseados em conceitos como a regulamentação, saúde, ética, educação, direito da criança e política criminal. Conclui-se que as soluções de big data trazem benefícios que podem remodelar as sociedades globais, os negócios e a vida humana em geral, mas emergem preocupações com a informação e o direito à privacidade e a possibilidade de o Big Data lesar liberdades, direitos humanos e garantias individuais.

PALAVRAS-CHAVE: Big Data; Gestão de Informação; Direito à privacidade; Revisão Sistemática da Literatura; Quarta Revolução Industrial

Information and Right to Privacy in the context of Big Data: A Systematic Literature Review (2012-2022)

Vinícius Barros Caixeta

Abstract

The role of Big Data in the Fourth Industrial Revolution goes beyond the vast amounts of data from different sources and formats, making patterns evident and enabling the automation of decision-making processes, thus adding significant economic and social value. With the proliferation of the internet and the popularization of the Internet of Things (IoT), the quantity of data has increased the information available to public and private entities. This massification of data, while bringing benefits in economic and social areas, raises many ethical issues to be debated and investigated in the field of Information Management and Curation. One of the main ethical issues concerns privacy, i.e., the threat to privacy resulting from the increase and integration of personally identifiable information. Based on this context, the present dissertation used the Systematic Literature Review methodology with the main objective of understanding how specialized literature has addressed access to information and the right to privacy in the context of Big Data. It was possible to identify six main topics, based on concepts such as regulation, health, ethics, education, children's rights, and criminal policy. It is concluded that big data solutions bring benefits that can reshape global societies, businesses, and human life in general, but concerns emerge regarding information and the right to privacy, and the possibility that Big Data may harm freedoms, human rights, and individual guarantees.

KEYWORDS: Big Data; Information Management; Right to privacy; Systematic Literature Review; Fourth Industrial Revolution.

Lista de Figuras

Índice

Introdução

A indústria, ao longo dos anos, passou por três revoluções indústriais que desencadearam profundas alterações nas estruturas sociais e nos sistemas económicos, trazendo um exponencial desenvolvimento não só para as indústrias, como também para a sociedade. Neste momento, acredita-se que estamos a viver uma quarta revolução industrial, que teve início na viragem do século e se baseia na revolução digital, quer dizer, é caracterizada por uma internet mais ubíqua e móvel, por sensores menores e mais poderosos, pela inteligência artificial e aprendizagem automática (ou aprendizado de máquinas) (Schwab, 2017, p. 158).

Segundo Schwab (2017), a quarta revolução não diz respeito apenas a sistemas e máquinas inteligentes e conectadas, apresenta, um âmbito muito mais amplo. Constatam-se descobertas que ocorrem paralelamente que vão desde o sequenciamento genético até a nanotecnologia. Ou seja, podemos afirmar que a quarta revolução industrial é essencialmente diferente de todas as anteriores, onde visivelmente se verifica a fusão dessas tecnologias e a interação entre o mundo físico, digital e biológico. Esta fusão originaria o que RAJKUMAR (2010, citado por Rossetti & Angeluci, 2021) denomina de *Cyber Physical Systems* (CPS).

De acordo com Rossetti e Angeluci (2021), num mundo Cyber Physical deduz-se que os objetos de interesse social e empresarial ficam disponíveis como dados digitais em redes, armazenados em tempo real, tornando-se, desta maneira, rastreáveis, explorados e analisados na rede. Esta situação leva a uma explosão de dados, referidos como Big Data. A presença destes nas relações económicas e sociais, sem dúvida, representa um *boom* para as áreas das ciências e empresariais. As ferramentas que fazem uso do Big Data são de grande importância para definir estratégias de marketing, diminuir os custos, aumentar a produtividade e tomar melhores decisões. Contudo, há que ter extremo cuidado na sua utilização, se os dados forem corretamente captados, analisados e apresentados, revelam perfis de comportamentos e preferências de enormes contingentes humanos. Por outro lado, se utilizados a favor de interesses particulares, esses mesmos dados podem oferecer às empresas a possibilidade de

induzir padrões de comportamentos e consumo, ou, até mesmo, influenciar nos meios públicos e levar a manipulação de grupos políticos a determinarem uma eleição de um candidato (Zwitter, 2014). Ou seja, dirigir sua estratégia comercial para excluir e/ou absorver clientes, como também, direcionar e manipular campanhas políticas a um perfil da população, que tornará decisiva a suas escolhas de voto.

Como podemos verificar, esta massificação dos dados constitui-se tanto uma oportunidade como um desafio. Um dos maiores desafios é o direito à privacidade. Este direito prevê o sigilo de informações pessoais e da própria vida pessoal, consagrado em 1948 no documento da Declaração Universal do Direito do Homem. No artigo 12, deste mesmo documento, o texto afirma que "ninguém será sujeito a interferências na sua vida privada, na sua família, no seu lar ou na sua correspondência, nem ataques à sua honra e reputação". Para tal, todas as pessoas têm direito a proteção da lei. A Comissão Europeia, no ano de 2016 (COMISSÃO EUROPEIA, 2016), ciente das novas configurações relacionais resultante da era digital, estabelece um pacote de medidas sobre a proteção da informação e a privacidade por meio do Regulamento Geral sobre a Proteção de Dados. Este regulamento representa uma medida essencial para legislar os direitos fundamentais da privacidade e intimidade na era digital e a tomada das decisões na cadeia de consumo.

No entanto, esta adequação legislativa apresenta algumas fragilidades quanto a sua aplicabilidade no Big Data, por um lado, o indivíduo tem pouco controle de como os seus dados estão sendo usados para gerar conteúdo sobre eles mesmos, em outro, o frenético crescimento do mundo tecnológico ocorre de forma mais rápida do que os normativos a essa proliferação digital.

Assim sendo, este estudo aborda as limitações dos normativos legais referentes à informação e a privacidade na era do Big Data. Sugere que a limitação deste discurso tem sua gênese na centralidade da responsabilidade moral dos agentes humanos individuais (Floridi, 1999). De facto, as novas configurações relacionais que estão surgindo no cenário contemporâneo, favorecida pelas redes digitais, resultante, por vezes, do que Floridi (1999) denomina de "a mão invisível de interações sistémicas",

envolvendo agentes globais e locais, humanos e artificiais, requer mais rigor, novos critérios e modelos de uma moralidade distribuída, na qual possa ser aplicada em diferentes níveis de ação e interação.

Há diversos estudos (Silva, 2019; Veldkamp & Chung, 2019; Neiva, 2020) que buscam avaliar a eficácia do Big Data em vários domínios. Estes estudos, na sua maioria, realçam as inúmeras controvérsias entre Big Data e o direito à privacidade, resultando, assim, um debate académico que reforça a necessidade de se criar uma legislação consistente que preveja o direito à privacidade no contexto do Big Data. Contudo, parecem não existir estudos que analisam diretrizes e/ou ações que possam harmonizar interesses pessoais de utilizadores e interesses económicos no contexto do Big Data, carecendo desta análise sobre o fenómeno.

Nesse sentido, o presente estudo utiliza uma Revisão Sistemática da Literatura, visando encontrar estudos de qualidade na literatura que correlacionem as três temáticas (informação, direito à privacidade e big data) e apresentem pontos relevantes dessa relação, para abordar a seguinte questão: como a literatura tem tratado a informação e o direito à privacidade no contexto do Big Data?

Esta dissertação está estruturada em seis capítulos. O capítulo 1 realizar-se-á a fundamentação teórica, em que serão fundamentados os conceitos de informação, Big Data e direito à privacidade suportado por autores mais eminentes. O capítulo 2, refere-se ao processo metodológico da pesquisa, no qual indicaremos o tipo de estudo e descreveremos os procedimentos metodológicos e a sua realização. No capítulo 3 apresentar-se-á a análise e discussão dos dados. São apresentadas as discussões do estudo no capítulo 4 e nelas são apontadas as limitações sentidas neste campo do estudo, manifestando igualmente as principais descobertas e possíveis debates futuros. Por fim, na última parte consta a conclusão do presente estudo, onde consolidamos as descobertas e perspetivas apresentadas ao longo da pesquisa.

Objetivo geral

Este trabalho visa, de maneira geral, entender como a literatura tem tratado a informação e o direito à privacidade no contexto do Big Data.

Objetivos específicos

I. Pesquisar nos principais periódicos internacionais estudos que relatem sobre a problemática da informação e a privacidade no contexto do Big Data;
II. Identificar os métodos de pesquisa, países de foco, contexto e ano de publicação dos estudos selecionados;
III. Identificar sugestões de diretrizes e/ou ações que possam harmonizar interesses pessoais de utilizadores e interesses económicos no contexto do Big Data;
IV. Identificar e relatar as lacunas ainda existentes na literatura no que concerne o acesso à informação e a privacidade no contexto do Big Data e propor novas direcções para os futuros pesquisadores.

1. Conceitos envolvidos

1.1. Definindo a palavra informação

O ato de definir, refere-se à formalização do significado de uma palavra de modo a reduzir elementos informais e imparciais. Contudo, esta formalização poderá apresentar-se incompleta se o significado for compreendido apenas pelo indivíduo que a definiu. Mas, mesmo assim, a definição ainda pode lançar nova luz sobre o objeto a ser definido (Polanyi, 1958). Tomando em consideração a afirmação de Polanyi, nesta secção, iremos percorrer de forma resumida o caminho da definição de informação no contexto das várias correntes teóricas desenvolvidas na Ciência da Informação, finalizaremos com a recente contribuição de Floridi do que venha a ser informação.

Segundo Saracevic (2009, pp. 1-2), a manifestação mais visível da Quarta Revolução Industrial é o fenómeno da "explosão de informação[1]", na qual se verificam registos de informações de toda espécie (publicações científicas, técnicas, expansão para outras áreas, como por exemplo, negócios, humanidades, direito, saúde, empreendedorismo, etc.). Até hoje, em todas estas áreas, o fenómeno da informação continua crescendo vertiginosamente, especialmente nos ambientes digitais e na *web*.

Como podemos verificar, a palavra informação é tão poderosa e, ao mesmo tempo, evasiva, pois pode ser usada em vários contextos e associar-se a diversas explicações, dependendo do conjunto de requisitos e das teorias que a oriente (Floridi, 2005).

Embora a palavra informação tenha-se tornado um assunto central, ainda não há uma definição única e unívoca da palavra. No ponto de vista de Floridi (2005) esta constatação, em parte, é compreensível se consideramos os diferentes usos da palavra em contextos diversos.

Na busca em definir uma determinada palavra, quase sempre recorremos à sua etimologia. É certo que a etimologia de uma palavra nunca pode garantir o seu significado. No entanto, ao examinar a história do uso de uma palavra, segundo Capurro & Hjorland (2007, p. 155), "[…] encontramos algumas das formas primitivas ou contexto subjacentes às práticas científicas de nível mais elevado", reduzindo, desta maneira, conceitos abstratos e, por consequência, ajudando-nos "[…] a entender como os usos atuais e futuros estão interligados".

---

[1] "[…] The term "information explosion" is a metaphor (as is "population explosion") because nothing really exploded but just grew at a high rate, even exponentially at times". (Saracevic, 2009, p. 2).

Portanto, antes de analisarmos o conceito da palavra informação e suas conexões com outras ciências, convém fornecer alguns fundamentos etimológicos da palavra.

1.1.1. A etimologia da palavra “informação”

De acordo com o *Dicionário Contemporâneo da Língua Portuguesa*, informação vem do latim *informatio, onis*, ("delinear, conceber ideia"), ou seja, dar forma ou moldar na mente (contexto epistemológico e ontológico), como em educação, instrução ou treinamento (contexto pedagógico).

Segundo Capurro & Hjorland (2007), o termo *informatio* teve uma importância fundamental na filosofia medieval e clássica. Aristóteles já se referia ao termo *informatio* imprimindo uma forma sob o material. Posteriormente, a ontologia de Aristóteles foi reinterpretada por Tomás de Aquino, que distingue *informatio* e *creatio*. Para Tomás de Aquino, Deus Cristão, na criação, não informa a matéria, mas a cria do nada.

Com o surgimento da ciência empírica, houve uma mudança do conceito de informação, no qual o contexto ontológico de dar forma foi abandonado a favor de premissas empíricas e epistemológicas, passando, assim, para o sentido de comunicar alguma coisa a alguém. Este novo sentido, revela-nos que o conceito de informação está relacionado como a visão de conhecimento, apresentando, desta maneira, “[…] uma conexão muito negligenciada entre as teorias da informação e as teorias do conhecimento”, haja vista que a informação significa dados processados sobre alguém ou sobre alguma coisa e conhecimento refere-se à informação útil obtida por meio da aprendizagem ou da experiência (Capurro & Hjorland, 2007, p. 159).

O termo informação, como usado no inglês, refere-se ao conhecimento comunicado. Este termo teve seu ápice na Segunda Guerra Mundial – com o desenvolvimento e disseminação de redes de computadores e, com a emergência da Ciência da informação (CI) como disciplina nos anos 50. Segundo Capurro & Hjorland (2007), embora o conhecimento e a sua comunicação sejam fenómenos inerentes a todas as sociedades,

são as Tecnologias de informação e comunicação (TICs) que caracterizam a nossa sociedade como sociedade da informação.

Como podemos verificar, a informação representa a base das sociedades atuais, na medida em que o controle das informações dos indivíduos – por meio de sistemas como o Big Data, exerce um poder considerável sobre as pessoas. Assim sendo, o acesso à informação tornou-se um direito humano. Este direito estende-se ao direito ao conhecimento das informações públicas acerca das ações governamentais, como também o direito ao acesso à informação veiculada nos meios de comunicação.

1.1.2.O caráter interdisciplinar da informação

De acordo com Capurro & Hjorland (2007, p. 173), a informação é caracterizada como uma área interdisciplinar, tendo em consideração que diversas disciplinas científicas usam o seu conceito dentro de seu próprio contexto e com relação a fenómenos específicos. Diante de tal situação, fica a seguinte questão: será que é possível um significado comum para a palavra "informação"?

Para Bogdan (1994, citado por Capurro& Hjorland, 2007, p. 160) é quase impossível um significado comum para a palavra informação. Esta constatação se dá pela multifuncionalidade da palavra "informação", visto que a mesma:

> […] tem sido usada para caracterizar uma medida de organização física (ou sua diminuição, na entropia), um padrão de comunicação entre fonte e recetor, uma forma de controle e feedback, a probabilidade de uma mensagem ser transmitida por um canal de comunicação, o conteúdo de um estado cognitivo, o significado de uma forma linguística ou a redução de uma incerteza. Estes conceitos são definidos em várias teorias como a física, a termodinâmica, a teoria da comunicação, a cibernética, a teoria estatística da informação, a psicologia, a lógica indutiva e assim por diante.

Capurro & Hjorland (2007) reconhece que há um esforço no campo da filosofia da informação em buscar um conceito comum para a palavra informação. O debate gira em torno de duas posições: se o conceito deveria remeter ao processo de conhecimento

– incluindo, como condição necessária um conhecedor humano ou, no mínimo um sistema interpretativo, ou se deveria excluir estados mentais e intenções relacionadas ao utilizador. Entre estas duas posições encontram-se diferentes teorias quantitativas da informação, como também uma teoria unificada (Capurro & Hjorland, 2007). O certo é que estas posições refletem a complexidade histórica do termo.

1.1.2.1.O conceito de informação nas ciências exatas

O conceito de informação nas ciências exatas pode ser compreendido por duas perspetivas teóricas: Teoria Matemática da Comunicação e Teoria Sistémica.

No contexto da Teoria Matemática da Comunicação, podemos afirmar que Shannon & Weaver (1949)foram os primeiros teóricos anunciarem um conceito científico de "informação" em seu livro "The Mathematical Theory of comunication". Esta teoria tratada no livro é normalmente conhecida como "Teoria da informação". Tal denominação se dá pelo fato de que os autores estão preocupados com a eficácia da comunicação e, portanto, privilegiam como conceito central de seu trabalho a informação.

Shannon & Weaver (1949) propõem e discutem uma análise tripartida da informação (ou comunicação). A primeira análise trata dos problemas técnicos, relativa à quantificação da informação (como, por exemplo, o volume do som numa conversa). A segunda análise se refere as questões semânticas, ou seja, relacionadas ao significado e à verdade. De acordo com Araújo (2009, p. 193):

> […] Enquanto o primeiro nível envolve apenas uma operação mecânica (reconhecer as letras num papel, captar os sons de uma fala), o segundo se relaciona a uma operação mental específica, a de depreender, de determinada materialidade (sonora, visual, etc.), um sentido, que pode se dar de maneira conotativa ou denotativa, literal ou irônica, metafórica, etc.

A outra análise é a influência[2], isto é, relaciona-se com o impacto e a eficácia da informação. Por exemplo, ao emitir informação acerca de um candidato à presidência, supostamente, desejo que seja provocado um comportamento e/ou causar alguma reação na pessoa que está ouvindo às informações de forma a convencê-la a votar no candidato.

Shannon & Weaver (1949, p. 6) reconhecem que: "[…] The effectiveness problem is closely interrelated with the semantIc problem, and overlaps It In a rather vague way; and there is in fact overlap between all of the suggested categories of problems".

Embora, Shannon & Weaver (1949) reconheçam que todas as análises estão interrelacionadas, ao produzirem uma teoria acerca dessa temática, direcionam-na apenas para a primeira análise, ou seja, para a análise técnica. Os autores, ao excluírem a análise semântica e a análise de influência, tornam possível a construção de um referencial teórico para os problemas técnicos, quer dizer, com o transporte físico da informação.

Na teoria proposta por Shannon e Weaver (1949), verifica-se que eles eliminam o significado do conceito de informação, descartando, assim, a subjetividade e elevando a informação para a dimensão objetiva. Ou seja, a informação é compreendida numa disposição linear. Assim sendo, os autores definem o conceito de informação (comunicação) da seguinte maneira: a fonte de informação, a partir de um transmissor, por meio de um canal, envia informação a um recetor, que a conduz a um destino. Este conceito pode ser simbolicamente visualizado na figura 1.

[2] A análise de influência é encontrada na literatura como análise pragmática (ver Capurro & Hjorland, 2007).

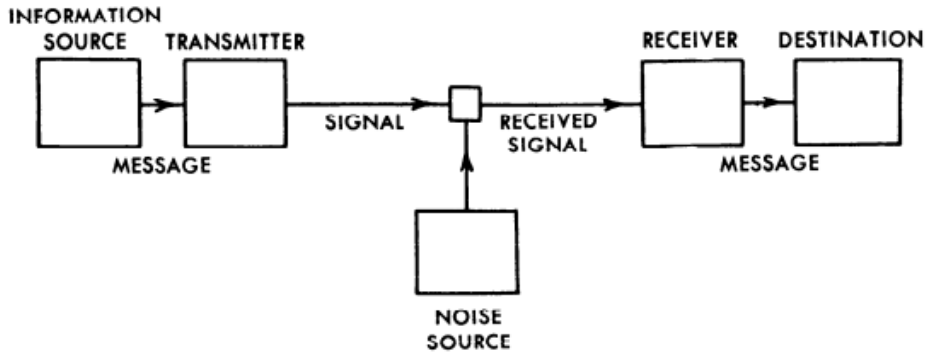

Figura 1- Sistema de Comunicação de Shannon e Weaver (1949, p. 07)

Os autores correlacionam a informação ao seu grau de incerteza – isto é, correlacionam ao número de escolhas possíveis a fim de criar uma mensagem. Quanto maior a liberdade de escolhas, maior a incerteza, quer dizer, a informação.

Este raciocínio evidencia dois outros importantes conceitos: entropia e probabilidade, ou seja, a informação é uma entidade da ordem de probabilidade, na qual a entropia [3]é um de seus atributos.

Contudo, estas teorias apresentam alguns problemas: como quantificar a informação, para determinar a quantidade ótima? Como alcançar o grau adequado de redundância de forma a prever a interferência do ruído e a capacidade do canal a ser transferida de um emissor a um recetor?

Por estas e outras questões, a Teoria Matemática da Comunicação recebeu algumas críticas, como por exemplo, Matsuno (1996, p. 111)sublinha que a informação é

[3] A entropia é o grau de casualidade, de indeterminação que algo possui. Ela está ligada à quantidade de informação. Quanto maior a informação, maior a desordem, maior a entropia. Quanto menor a informação, menor a escolha, menor a entropia (Shannon & Weaver, 1949).

intrinsecamente ambivalente em relação à dimensão temporal e que a temporalidade da informação é diacrônico ou sincrônico. A teoria da informação de Shannon "[…] refers exclusively to synchronic information in the sense that the source matrix of messages or information is given as a temporally invariant entity". Ou seja, esta teoria refere-se a um processo que existe num espaço de tempo finito, ignorando antecedentes históricos.

O certo é que esta teoria teve um impacto muito expressivo nos estudos que se seguiram *a posteriori*, principalmente no que diz respeito a transferência da informação e a recuperação da informação.

Segundo Saracevic (2009), a recuperação da informação surge na década de 50 e que, muitas vezes, foi considerada como sinónimo de CI (Ciências da Informação). Este campo está direcionado para a medição de procedimentos para a recuperação da informação.

De acordo com Shapiro (1995), cada disciplina possui uma terminologia, cujas vicissitudes, muitas vezes, legitima e eleva o *status* de uma profissão ou disciplina. Contudo, há a "tendência de se usar e de se definir termos para impressionar outras pessoas" (Capurro & Hjorland, 2007, p. 154). Tal atitude é denominada de definição persuasiva. Provavelmente, o uso da palavra "informação" que causou mais confusões e, inclusive, falta de respeito a disciplina tenha ocorrido no uso do termo "Ciências da Informação".

Schrader (1986), por meio de uma revisão sistemática da literatura, realizou um abrangente estudo sobre as variadas definições de CI e uma análise dos conceitos de informação relevantes para a CI. Ele encontrou cerca de 700 definições para a palavra CI do ano de 1900 a 1981 e chegou à seguinte conclusão:

> The literature of information science is characterised by conceptual chaos. This conceptual chaos issues from a variety of problems in the definitional literature of information science: uncritical citing of previous definitions; conflating of study and practice; obsessive claims to scientific status; a narrow view of technology; disregard for literature without the science or technology label; inappropriate analogies; circular definition; and, the multiplicity of vague,

> contradictory, and sometimes bizarre notions of the nature of the term 'information'. (Schrader, 1986, p. 94).

Como podemos observar, Schrader (1986) enfatiza o caos conceitual em relação às definições de Ciências da informação como também do seu próprio objeto. O autor, em seu artigo "The domain of information science: problems in conceptualization and in consensus-building" analisou vários usos do termo informação, concluindo que a maioria das definições reivindicam o conceito do termo como ´dados ou conhecimento´.

No ponto de vista de Lancaster (1993, citado por De Magalhães & Souza, 2019), um exemplo provavelmente óbvio de como o termo informação perdeu completamente o seu sentido é quando usamos a expressão recuperação da informação (RI), visto que, em verdade, não é realizada a recuperação da informação em si, mas um processo de pesquisa numa coleção de documentos a fim de identificar aqueles textos que tratam de um determinado assunto.

Para melhor entender o conceito de recuperação de informação, imaginemos uma biblioteca, com livros e periódicos das mais variadas áreas de conhecimento, e um utilizador com necessidade de informação sobre, por exemplo, a "4ª Revolução Industrial". Este utilizador poderá ler todos os livros e periódicos da biblioteca e identificar e selecionar os livros e capítulos relacionados com o tema/assunto desejado. Este processo é denominado de recuperação da informação, no entanto, seria humanamente impossível o utilizador pesquisar todo o acervo da biblioteca.

Com a utilização dos computadores e da internet o processo foi automatizado, trazendo mais agilidade e eficiência(Saias, 2003). Estes sistemas automatizados foram originalmente usados para gerenciar o grande volume de informação na literatura científica na segunda metade do século XX. Várias universidades, bibliotecas públicas, domínio público e portais de pesquisa de produções científicas como o RCAAP (Repositórios Científicos de Acesso Aberto de Portugal) até hoje usam o SRI (Sistema de Recuperação de Informação) para facilitar o acesso a livros, periódicos, documentos públicos, etc.

De acordo com Saias (2003) para recuperar a informação, o utilizador tem de partir de uma interrogação. Uma interrogação é composta por um conjunto de palavras que possam ser úteis ou relevantes para o utilizador. Tal informação (de interesse do utilizador) é normalmente chamada de "necessidade de informação" do utilizador.

Há inúmeros sistemas de recuperação de informação, como por exemplo a bibliometria e o *data mining*. Este último, será mais bem explorado na secção que trata sobre Big Data.

Relativamente à bibliometria, Araújo (2009)afirma que é anterior à teoria da matemática. Ainda segundo o autor, a bibliometria faz uso de métodos estatísticos e matemáticos para a contagem e estabelecimento de padrões de regularidade acerca da dinâmica e evolução da informação científica ou tecnológica como identificação de tendências e crescimento do conhecimento numa determinada disciplina, medição do impacto das publicações e dos serviços de disseminação da informação, avaliação de publicações científicas, etc. Para Roy & Basak (2013, p. 2)bibliometria "is the discipline where quantitative methods were employed to probe scientific communication process by measuring and analyzing various aspects of written documents". A bibliometria é uma área multidisciplinar, podendo ser aplicada em quase todas as áreas relacionadas com a comunicação escrita como seja a biometria, econometria e cienciometria.

Embora os estudos bibliométricos não se refiram ao transporte de informação, define a informação da mesma maneira que a teoria da matemática, ou seja, "[…] que a informação pode ser quantificada e que, por meio dessa quantificação, seria possível prever suas manifestações futuras, já que, tal como os fenómenosda natureza, ela também obedeceria a leis que regem sua existência" (Araujo, 2009, p. 194). Tal confusão conceitual favoreceu, em meados de 1950, a aproximação da bibliometria com a recuperação da informação (Wormell, 1998).

De acordo com Wormell (1998), Eugene Garfiel foi o fundador da bibliometria. Ela, Eugene Garfiel, propôs à utilização deste método para a utilização de contagens de

citações para a recuperação da informação como para a medição bibliométrica de itens recuperados em processos de busca e seleção.

O certo é que este campo da bibliometria está-se afirmando como uma disciplina científica, na qual inclui todos os aspetos estatísticos e matemáticos acerca dos problemas da biblioteconomia, da documentação e da informação baseados na recuperação da informação.

As áreas empregadas na recuperação de informação têm avançado nas últimas décadas – principalmente com o rápido crescimento e desenvolvimento das tecnologias da informação, bem como, a ampliação do espaço de armazenamento, velocidade dos computadores e da internet. Essa nova área é denominadade *Big Data*, que será abordada na secção subsequente. De momento, podemos dizer que *Big Data* é a área do conhecimento que estuda como tratar, analisar e obter informação a partir de um grande volume de dados. No tocante a ciência, o *Big Data* levou ao surgimento de um novo paradigma (4ª Revolução Industrial), quer dizer, concebeu um novo método para ampliar as fronteiras da informação e, por conseguinte, do conhecimento.

Retomando o caráter interdisciplinar da informação, observa-se que as reivindicações acerca da naturalização do conceito de informação não couberam apenas a matemática e a física, mas também a biologia. De acordo com Titze (1971, citado por Capurro & Hjorland, 2007), a informação não é apenas um atributo da metafísica, mas expressa uma tendência para a ordem e evolução. Esta tendência ganha imensa expressão na CI com o trabalho pioneiro de Nobert Wiener e outros cibernéticos no final dos anos 40 e início dos anos 50 (Rifkin, 2003).

Para Wiener (1961, p. 132)[…] "information is information, not matter or energy. No materialism which does not admit this can survive at the present day". Este novo conceito de informação está relacionado com suas origens latinas, onde a informação é um processo dinâmico e não apenas um significado de uma mensagem.

Segundo Araújo (2009, p. 195) a ideia principal desta teoria é "[…] de que o todo é maior do que as partes e de que as partes devem ser estudadas, necessariamente, a partir da

função que desempenham para a manutenção e sobrevivência do todo". Nesta teoria, os princípios biológicos passam a ser referência geral para o estudo de qualquer fenómeno.

Enquanto o modelo matemático e físico pensava os processos da informação (comunicação) em termos lineares – quer dizer, do transporte de um ponto ao outro e da forma como otimizar tal transporte, no modelo biológico, a informação é tratada em todos os seus níveis (técnico, sintático e de influência), numa lógica de circularidade, quer dizer, todo processo sempre representa a saída de alguma entidade e, essa saída, vai provocar a formação de novos elementos de entrada. Esta dinâmica pode ser visualizada naquilo que denominamos *input* (entrada), *output* (saída) e retroalimentação (Araújo, 2009). Em outras palavras:

> Os sistemas de informação são sempre pensados a partir da lógica dos processos de entrada (entrada de dados, com a aquisição de itens informacionais, a seleção destes itens para a composição de determinado acervo), de processamento (os itens informacionais que dão entrada num sistema de informação precisam ser descritos, catalogados, classificados, indexados) e de saída (pelo acesso aos itens informacionais por parte dos usuários, na forma de disseminação, entrega da informação, empréstimo, etc) (Araújo 2009 p. 196).

Conforme Araújo (2009), relativamente à gestão de negócios, diversos modelos teóricos buscaram evidenciar os riscos, oportunidades, os pontos fortes e fracos de uma determinada empresa, ou seja, determinar as falhas em processos internos e externos, dos sistemas e de pessoas.

Contextualizando esta situação descrita acima com a teoria sistémica, podemos verificar que ela encadeia uma série de conceitos peculiares, tais como o de totalidade (uma organização governamental, uma empresa, uma equipa, etc), os objetos que compõe a totalidade (os membros de uma organização governamental, os membros de uma empresa, departamento e/ou equipa), os atributos destes objetos (características específicas que cada objeto tem de forma que os processos sejam eficientes e a empresa

ou uma organização governamental consiga manter seus sistemas otimizados, os processos (a "importação" ou entrada de algo, a "exportação" ou saída, e o processamento desse algo, entendido como as tarefas necessárias para a sobrevivência do sistema) e o ambiente (aquilo que é externo à totalidade, de onde ela retira os elementos de entrada e para onde dirige os elementos de saída) (Araújo, 2009).

No ponto de vista de Capurro & Hjorland (2007, p. 162), este processo de naturalizar a informação sob a perspectiva das ciências específicas, nomeadamente a matemática, a física, a biologia e a linguística demonstram uma tendência de "re-humanizar o conceito de informação", ou seja, colocá-lo no contexto cultural. Contudo, segundo o autor, verifica-se ainda a necessidade de uma reflexão mais profunda em que a informação e comunicação, sendo humanas ou não, sejam vistas com as suas correspondentes "*differentia specifica*" de acordo com o ponto de vista do género de interpretação ou seleção.

Mais se acrescenta, que esta reflexão representa não só a reaparição da dimensão ontológica das raízes gregas de *informatio* e da visão humanística restrita, como também da perspectiva moderna, desumanizada, onde a informação é considerada como conhecimento comunicado. Em outras palavras, dá origem ao que Capurro & Hjorland (2007) denomina de ontologia comunicativa em que não apenas seres vivos, mas todos os tipos de sistemas são considerados como produtores, processadores e compartilhadores de informação. Esta reflexão no contexto do Big Data, em que há uma enorme quantidade de informação sendo criada pelos seres humanos, onde estas se diferem da informação tradicional tanto em termos de formato quanto em conteúdo, é de suma importância.

1.1.2.2.O conceito de informação nas ciências humanas e sociais (teoria crítica)

Enquanto os modelos teóricos anteriormente discutidos, principalmente o sistémico, reivindicavam a estabilidade, a permanência e a integração da informação, contrariamente, o modelo na perspetiva das ciências humanas e sociais vão "enfatizar

o conflito, a desigualdade, o embate de interesses em torno da questão da informação – e para tanto, buscará explicar os fenómenosa partir de sua historicidade" (Araújo, 2009, p. 196).

O conceito de informação, nesta perspetiva, não está mais voltado para a transferência da informação, nem muito menos para o problema do equilíbrio social ou dos procedimentos funcionais para seu processamento no âmbito dos sistemas (Araújo, 2009). A Informação, nesta teoria, é compreendida como um recurso primordial para a condição humana. Neste contexto, ela não pode estar restrita a um pequeno grupo social. Portanto, os estudos desenvolvidos no âmbito dessa perspetiva são direcionados para as questões e problemas sociais, em outras palavras, preocupam-se com a democratização da informação, com o acesso à informação, com a criação de sistemas de informação e, também com as questões éticas da informação (Araújo, 2009).

No entanto, de acordo com Araújo (2003), as análises em torno desses estudos, pareciam conceber mais uma subárea dentro CI – quer dizer "a área de informação social", do que buscar compreender como a CI entende o seu objeto de estudo.

Ainda na perspetiva de Araújo (2003, p. 25), este entendimento do que venha a ser informação, só ocorrerá, de facto, quando a CI se aproxima da microssociologia, onde o conceito de informação é baseado em análises interpretativas ao invés de observações empíricas e/ou estatísticas propostas nas teorias anteriores, possuindo grande aproximação com a fenomenologia e etnografia e, portanto, com a perspetiva marxista.

Nesta teoria, entende-se "que a realidade é construída socialmente" (Berger & Luckmann, 2004, p. 11) "e não com uma existência em si mesma" (Araújo, 2003, p. 25). Assim sendo, o termo informação em CI, contrapondo a ideia de informação sob o prisma técnico-sistémico – no qual a informação é representada como um dado e/ou como uma coisa, que tem significado em si mesmo, lança uma nova visão ao termo, agora, orientada ao utilizador e ao ser humano. (Araújo, 2003). Esta perspetiva parte do princípio de que a informação é um conceito subjetivo, haja vista que ela pode ser percebida e compreendida de acordo como os sujeitos, por exemplo, uma pedra num

campo pode ser interpretada como um tipo de informação para o geólogo e um outro para o arqueólogo (Capurro & Hjorland, 2007).

Segundo Araújo (2003) vários estudos seguiram esta perspetiva no âmbito da CI, não no sentido de constituir uma linha de pesquisa distinta, mas de reformular o conceito do objeto de estudo da CI como um todo. Dentre estes estudos, destacamos os estudos de natureza cognitiva inspirados na teoria de Maturana & Varela, a abordagem hermenêutica da ciência da informação desenvolvida por, entre outros, Capurro e os estudos sobre os valores dos utilizadores originados com MacMullin & Taylor.

Contudo, esta teoria – devido ao ritmo frenético da revolução da informação, que traz consigo questões importantes sobre criação, gestão e utilização de informações (Baumgaertner & Floridi, 2016)– sofre algumas críticas e, portanto, neste momento, é possível vislumbrar o advento de um novo paradigma que vem ganhando força constantemente na "Ciência pós-moderna".

Ao longo da explanação desta secção, verificamos que, no contexto de um ambiente tecnológico, diversas definições de informação foram construídas tanto no âmbito da abordagem quantitativa – como é o caso da Teoria Matemática da Comunicação e a Teoria Sistémica, quanto na abordagem qualitativa – como é o caso da Teoria Crítica, no entanto, estas diferentes abordagens ao conceito de informação causou ambiguidades na própria palavra "informação". Isto, provavelmente se deu pelo fato de que a maioria dos investigadores epistémicos da CI não realizaram uma fundamentação conceitual na perspetiva filosófica.

Baumgaertner e Floridi (2016) sublinham que a revolução da informação trouxe enormes benefícios e oportunidades, contudo, ultrapassou a nossa compreensão dos seus fundamentos e consequências, levantando questões conceituais que estão-se a exxpandir, evoluindo e ficando cada vez mais grave. Assim sendo, na mesma direção da teoria das ciências humanas, Floridi (2012)propõe um campo próprio para a informação, quer dizer, ele sugere uma Filosofia da Informação (FI). No ponto de vista de Floridi (2012) este campo pode apresentar-se como um paradigma inovador no sentido de

proporcionar uma área rica, útil e oportuna de investigações conceituais acerca da informação.

A FI, segundo Floridi (2012, p. s/n) é:

> […] the philosophical fieldconcerned with the critical investigation of the conceptual nature and basic principles of information, including its dynamics, utilization and sciences, and with the elaboration and application of information–theoretic and computational methodologies to philosophical problems.

Ao propor uma Filosofia da Informação, Floridi (2002)não pretende mapear os problemas específicos, tão pouco validar uma teoria unificada, mas analisar, avaliar e explicar sob uma perspetiva crítica, os princípios e os conceitos de informação, incluíndo as suas dinâmicas e utilização. Adicionalmente, propõe estar em constante vigilância para as questões sistémicas que surgem nos diferentes contextos de aplicação, sem descurar de outros conceitos essenciais para a filosofia como o ser, a vida, a verdade, o significado e o conhecimento.

Ao incluir a informação na agenda da filosofia, Floridi (2002) não ignora a importância das abordagens quantitativas, contudo, segue um caminho oposto, isto é, adota uma abordagem qualitativa. Esta abordagem traz importantes conceitos como o de informação semântica, significado e veracidade da informação. Estes conceitos abrirão caminhos para um esclarecimento do fenómeno da informação.

No contexto do *Big Data*, onde a informação em forma de dados, é coletada, armazenada e disseminada num ritmo acelerado, é importante a proposta de estudar o fenómeno da informação sob uma perspetiva filosófica de forma que as questões de privacidade sejam capazes de acompanhar as novas formas de violação decorrentes das novas configurações sociais que vem-se formando com os avanços tecnológicos da sociedade da informação. Assim sendo, discorrer-nos-emos, na próxima secção, sobre a informação na perspetiva filosófica de Floridi.

1.1.3. A Filosofia da Informação em Floridi

O termo Filosofia da Informação (FI) foi batizado por Luciano Floridi, professor e investigador da Universidade de Oxford, Inglaterra. A pesquisa de Floridi sobre a natureza, dinâmica e uso da informação levou a uma "tetralogia[4], chamada de *Principia philosophiae informationis,* composta pelas seguintes obras: The Philosophy of Information (2011, primeiro volume da tretalogia); The Ethics of Information (2013, segundo volume); The logic of information (2019, terceiro volume; The politics of information (não publicado, quarto volume).

Cabe ressaltar que para além da tetralogia, Floridi publicou mais duas outras obras relacionadas ao tema, a saber: "Information: a very short introduction", em 2010 e "The fourth revolution: how the infosphere is reshaping human reality", em 2014. As relações entre as obras podem ser visualizadas na Figura 1.

[4] Tetralogia, segundo Floridi (2020) é "fancy word to say 4-volume work, not my term". Disponível em: https://www.philosophyofinformation.net/about/. Consultado em 22/02/2023.

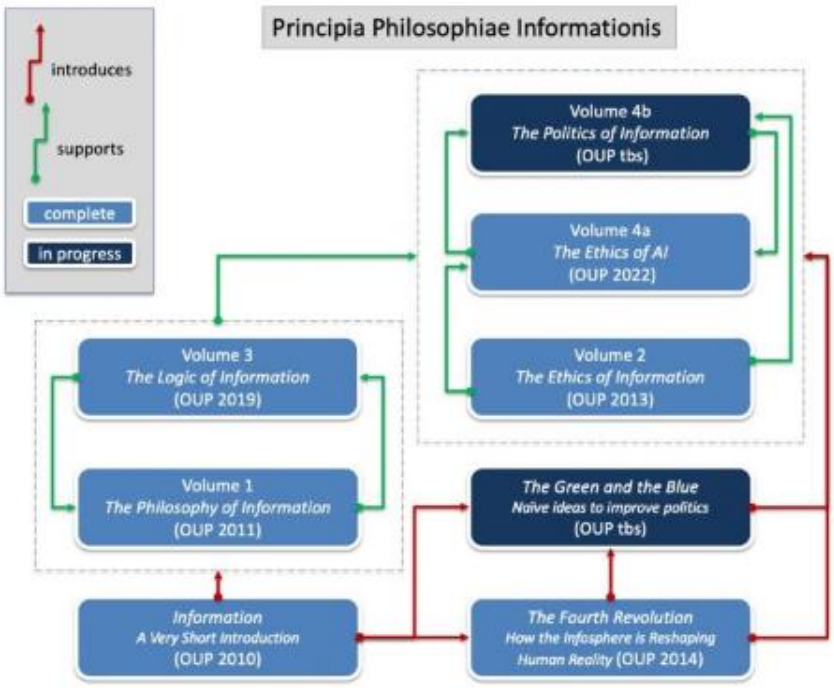

Figura 2- Principia Philosophiae Informationis de Luciano Floridi (2020)[5]

Assim sendo, Floridi (2020)tem trabalhado no desenvolvimento da Filosofia da informação como uma nova e independente área de pesquisa. O autor parte do princípio de que a informação é tão importante quanto os conceitos de verdade, significado, conhecimento, ser, e bem e mal.

Para Floridi (2002), o computador e as tecnologias de informação e comunicação apresentam-se como um símbolo do novo milénio. Estes, por sua vez, têm desempenhado um papel muito mais influente do que os moinhos na Idade Média, dos relógios mecânicos no século XVIII, do tear e da máquina a vapor na era da Revolução Industrial. Em outras palavras, o autor quer dizer que a atual sociedade da informação

[5]FLORIDI, Luciano. Research: The tetralogy project. 2020. Disponível em: Pesquisa – Luciano Floridi Filosofia da Informação (philosophyofinformation.net). Consultado em 22/02/2023.

tem gerado inovações, que até então ninguém acreditava que fosse possível. A força motriz destas inovações:

> […] is represented by the world of information, computation and communication phenomena, their corresponding sciences and technologies, and the new environments, social life, as well as the existential, cultural, economic and educational issues that they are bringing (Floridi, 2012, p. 3541).

Portanto, neste contexto dos fenómenos da informação, computação e comunicação, a informação não representa apenas um meio para obter conhecimento, mas se torna a base da discussão sobre tudo o que movimenta a sociedade atualmente. Assim sendo, é importante uma definição voltada exclusivamente para informação de modo que possibilite a demarcação de seu campo de atuação.

Floridi (2012) vai-se aventurar nesta empreitada por acreditar que é importante estabelecer um sentido técnico para a informação, haja vista que este servirá como base para julgar a informação. Em outras palavras, buscará estabelecer o grau de veracidade numa determinada informação.

Portanto, Floridi (2011)propõe uma Filosofia da Informação que se constitui de forma íntegra e sistemática pautada nas ciências, mas, principalmente sob um olhar crítico às novas configurações que esta vem contribuindo.

No intuito de subsidiar tal tarefa, Floridi (2011) por meio de um mapa conceitual apresenta alguns problemas essenciais para FI. O mapa conceitual inicia-se com uma pesquisa ontológica, ou seja, investiga a natureza e a existência da informação. Esta pesquisa ontológica tem como objetivo responder a três questões vitais: O que é informação, quais são as dinâmicas da informação e se é possível uma teoria unificada da informação.

De antemão, Floridi (2011), embora não explique a natureza especifica da informação, afirma que ela deve ser quantificável, plausível de adição, armazenável e transmissível. A partir disso, destaca três conceções importantes de cunho ontológico: a) informação como realidade; b) informação sobre a realidade e a c) informação para a realidade.

Contudo, o autor vai dedicar o seu estudo às definições da informação sobre a realidade, a qual a informação é abordada sob seis diferentes teorias: matemática, física, biológica, económica, ética e semântica.

Embora Floridi (2011) reconheça a importância dessas outras teorias, vai dedicar-se às definições de informações semânticas, dado a sua importância para o nosso estudo, será tratada na próxima subsecção.

1.1.4. Informação semântica

Nas últimas três décadas, a maioria das análises apoiou a formulação de uma definição bipartida de informação, ou seja, com dados bem formados (sintaxe) e significativos (semântica). Esta definição ganhou consenso entre os investigadores da área, tornando-se, desta forma, uma Definição Geral de Informação (DGI). Apesar de reconhecer a utilidade prática da definição bipartida, Floridi (2011) sublinha que tal definição não atende as discussões epistemológicas sobre a informação, acrescenta que o problema da DGI provém da sua neutralidade relativamente ao valor da informação, ou seja, a sua veracidade e, que este tipo de "neutralidade", poderá propiciar a falsa informação (desinformação).

Assim sendo, na perspetiva de Floridi (2011), os dados sendo rigorosos na sua validade, jamais permitirá a informação falsa no seu objectivo. Desta maneira, o autor formula uma definição de informação semântica, fazendo uma ampliação da DGI, onde afirma que a informação semântica se constitui de dados bem formados, que possuem significados e são verídicos – salientamos que é esta a definição de informação usada neste estudo, onde "bem-formados" refere-se a ocorrência da semântica, isto é, os dados estão de acordo com o significado do sistema ao qual está vinculado, "significativo" é quando os dados estão de acordo com os significados (semântica) do escolhido sistema, código ou idioma em questão e "verídicos" é quando possui um significado naquilo que está sendo informado, visto que tem uma relação direta com a realidade.

De forma resumida, Floridi (2003, 2005) faz a distinção dos tipos de informação por meio de um mapa conceitual. Esta, por sua vez, estrutura-se a partir do dado. O dado é classificado como analógico ou digital. Em relação a este último ainda tem como característica o dado binário, isto é, aquele codificado com linguagem binária, que se expressa de 0 a 1.

Segundo Floridi (2003, 2005), a informação pode ser constituída por cinco tipos diferentes de dados – como pode ser visto em seu mapa conceitual (Figura 3), são eles: a) dados primários (são aqueles que constituem diretamente a informação, ou seja, os dados contidos nas páginas de um livro); b) dados secundários (são o inverso dos dados primários, constituído pela ausência destes, ou seja, a inexistência de dados, também ocasiona informação. Por exemplo, quando o seu automóvel não emite nenhum ruído ao girar a ignição, gera dados secundários sobre o estado da bateria do automóvel); c) os metadados (são indicações secundárias sobre a natureza dos dados primários). Eles permitem que um sistema de gestão de banco de dados cumpra suas tarefas descrevendo propriedades essenciais dos dados primários, por ex. localização, formato, atualização, disponibilidade, restrições de direitos autorais, etc.); d) dados operacionais (são dados sobre o uso dos próprios dados, as operações de todo o sistema de dados e o desempenho do sistema) e, por fim, e) os dados derivados (são dados que podem ser extraídos, sempre que estes sejam utilizados como fontes em busca de padrões, pistas ou inferências provas, por ex. para análises comparativas e quantitativas (ideometria).

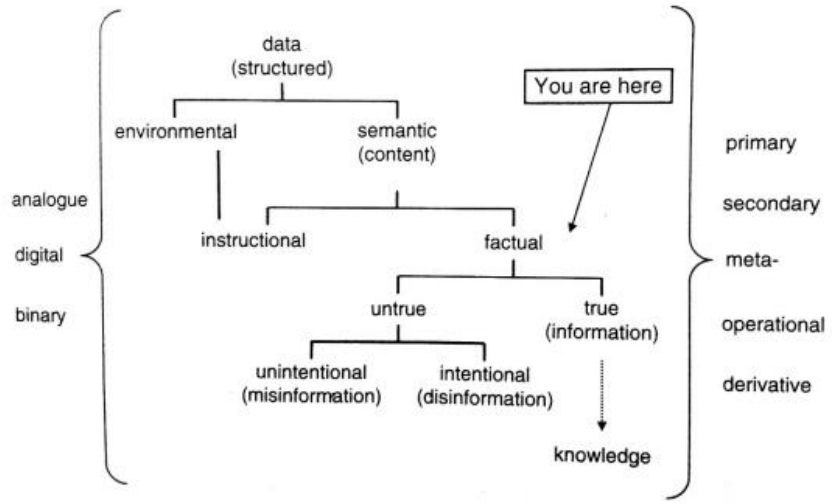

Figura 3- Mapa conceitual de Floridi

Fonte: Ripoll & Matos, 2021, p. 219

Por seu turno, estes dados podem ser do ambiente ou de conteúdo semântico. Os dados do ambiente, como o próprio nome indica, são dados que podem ser encontrados no ambiente e que mesmo não sendo de um sujeito produtor, podem ser significativos. Os dados de conteúdo semântico geram dois tipos de informação: a) instrucional (são aquelas que contêm instrução de como proceder diante de um fato) e factual (são aquelas que somente representam o fato). Floridi (2003) enfatiza que informação instrucional, ao contrário da factual informação, não pode ser considerada verdadeira ou falsa.

No que lhe toca, a informação factual é dividida em *untrue* e *true*. Aquela *(untrue*) é caracterizada por pseudo-informação, que podem ser intencionalmente falsas (*disinformation*) ou apenas falsa por motivos alheios (*misinformation*). Esta (*true*) corresponde a dados factuais semânticos verdadeiros, isto é, produzem informação, logo, geram conhecimento.

Segundo Floridi (2003), informação e conhecimento tem uma estreita relação, visto que este encapsula a verdade porque está baseado na informação semântica factual, ou seja, constituída por dados bem formados, significativos e verdadeiros.

No contexto do nosso estudo, onde os dados dos indivíduos são usados em aplicações como o *Big Data*, a validade ou veracidade da informação, torna-se essencial de modo a fazer um uso responsável dos dados e assim, aumentar a confiança e manter relacionamentos fortes com os clientes e/ou consumidores.

Na próxima secção iremos abordar o *Big Data* de forma que seja possível entendermos como as ferramentas do Big Data por meio da criação, gestão e utilização de informações consegue obter o controle das diversas atividades desenvolvidas pelos indivíduos, nas inúmeras situações de sua vida, permitindo, desta forma, o conhecimento nos mínimos detalhes de sua conduta pública e privada.

1.2. Construindo a definição de big data

Sem dúvida, estamos vivendo uma Nova Era caracterizada pelo grande volume de dados e informações, denominado *Big Data*. Este termo, De acordo com De Mauro et al. (2016) – devido a nova natureza das tecnologias, somado a imensa variedade de aplicativos que dependem de dados, tornou-se comum em trabalhos acadêmicos e científicos tanto nas áreas das Tecnologias de Informação e Comunicação (TIC) quanto em outras áreas do conhecimento, como: sociologia, medicina, biologia, economia e gestão.

No entanto, como observa De Mauro et al. (2016, p. 122), "[…] the degree of popularity of this phenomenon has not been accompanied by a rational development of an accepted vocabulary […]". No ponto de vista desses autores, "[…] the term 'Big Data' itself has been used with several and inconsistent meanings […]", necessitando, desta forma, de uma definição formal.

Assim sendo, De Mauro et al. (2016) buscam clarificar o fenómeno *Big Data* de forma que as principais conexões ao termo sejam identificadas e, desta maneira, alcançar uma definição abrangente e assertiva. Para tal, os autores, por meio de uma revisão da literatura, procuraram analisar as ocorrências das palavras mais significativas relacionadas ao termo *Big Data*, agrupando estas palavras conforme as suas conexões conceituais. Deste procedimento, os autores captaram quatro temas fundamentais, quer dizer, conceitos predominantes que representam os componentes essenciais da

temática estudada, são eles: informação, tecnologia, método e impacto. Nas próximas linhas,buscar-se-á examinar cada um destes temas e sua conexão ou interconexão com o *Big Data*.

Podemos afirmar que uma das razões do rápido crescimento do *Big Data* está relacionado com a extensa **velocidade** na qual os dados são criados, compartilhados e utilizados atualmente (De Mauro et al., 2016). Para Floridi (2010) estamos vivendo um novo paradigma, caracterizado por um fenómeno social, denominado de revolução da **informação**, onde as TICs proporcionam uma nova organização da sociedade, que funciona com base em serviços e produtos informacionais. "[…] Hence, we can conclude that information, not data, is the fundamental fuel of the current Big Data phenomenon" (De Mauro et al., 2016, p. 123). De Mauro et al. (2016), ao analisarem os elementos comuns em todos os aplicativos de *Big Data*, constataram que há uma hierarquia dados-informações-conhecimento-sabedoria, na qual a **informação** aparece como dados estruturados de forma a serem úteis e relevantes para um propósito específico. Assim, nesta perspetiva, segundo Cricelli e Grimaldi (2008, citado por De Mauro et al., 2016), a **informação** poderá se tornar um ativo de conhecimento e, assim, gerar valor para as empresas. Por exemplo, no caso da digitalização das bibliotecas, os métodos usados pelo *Big Data* podem identificar os vários processos de catalogação dos ativos de **informação** do acervo de uma determinada biblioteca ao longo dos anos, bem como detetar novas inconsistências nos dados. Contudo, para tal procedimento é necessário uma tecnologia avançada em gestão de dados. Esta tecnologia, justifica-se porque o conteúdo digital pode estar em diferentes níveis de sintaxe e abstração semântica. No ponto de vista de De Mauro et al. (2016), as aplicações de Big Data oferecem flexibilidade suficiente para lidar com esses ativos de informações intrinsecamente heterogêneos, visto que o *Big Data* é uma ferramenta especializada para manipular as três dimensões dos dados: volume, velocidade e variedade. Segundo, De Mauro et al. (2016), outra causa importante sobre o crescimento exponencial da informação diz respeito ao grande crescimento dos dispositivos pessoais conectados à internet, equipados com sensores digitais (gravadores de áudio, circuitos de câmaras de vigilância

e localização geográfica; e/ou cedência de dados pessoais). Estes sensores possibilitam a digitalização, visto que a conexão de rede permite que os dados sejam coletados, transformados e, por fim, organizados como informação. A IoT Analytics , consultoria que atua na área de *Internet of Things* (IoT na sigla em inglês, ou Internet das Coisas), estima que até 2025 haverá, em todo mundo, cerca de 27 bilhões de dispositivos conectados à rede mundial de internet. No ano de 2022 a expetativa é que os números de novos dispositivos aumentassem em 18%, chegando a 14 bilhões de dispositivos[6]. Este cenário em que objetos do dia a dia (smartphones, veículos, aspirador de pó, frigorífico, prédios e outros providos de **tecnologias** com capacidade computacional e de comunicação) conectados à internet, são denominados de Internet das Coisas (IoT)[7]. Estas representam uma fonte profícua da informação no contexto do *Big Data*. Contudo, os dados que são produzidos e utilizados atualmente têm uma **variedade** de formato, como os vídeos, imagens e textos produzidos pelos humanos, que representam fontes de dados menos estruturadas. Esta variedade de tipos e formatos de dado representa um dos principais desafios para as aplicações do *Big Data*. Portanto, a tecnologia corresponde uma condição necessária para usar *Big Data*.

Como podemos verificar, a **tecnologia** é o segundo tema encontrado na literatura que faz conexão com o *Big Data*. Ou seja, para lidar com o fenómeno do *Big Data* é necessário a utilização de tecnologias adequadas de forma que esta seja capaz de processar um grande **volume** de dados na **velocidade** certa em que estes são produzidos. Adicionalmente, para além do processamento, outro ponto primordial – devido à natureza dispersa das máquinas, é a questão da transmissão. As redes de

[6]https://www.terra.com.br/noticias/ate-2025-mundo-tera-cerca-de-27-bilhoes-de-dispositivos-iot-conectados,ba609033b546442d18d06660dbc7bb2d709p6o4w.html.

[7]https://pt.wikipedia.org/wiki/Internet das_coisas.

comunicação precisam sustentar transferências de conjunto de dados e, portanto, os sistemas vão necessitar de técnicas específicas de *benchmarking* para avaliar seu desempenho geral (Xiong et al., 2013, citado por De Mauro et al., 2016). Gordon Moore (1965) previu que o poder computacional de um circuito integrado de um processador dobraria a cada dois anos. Isto implica dizer que vão surgir, a cada dois anos novos processadores com capacidades duas vezes mais potentes de armazenamento. No entanto, paralelamente a esta constatação, os dados também crescem de forma exponencial, assim, o armazenamento de uma grande quantidade de dados, torna-se um desafio técnico para o *Big Data* e, portanto, necessitando de métodos adequados para realizar às análises.

O **método** é o terceiro tema encontrado na literatura que faz conexão com o *Big Data*. Desta forma, para lidar com um grande volume de dados o *Big Data* necessita de **métodos** mais complexos para processá-los do que os habituais métodos estatísticos. Conforme De Mauro et al. (2016) o desenvolvimento do *Big Data* transformou o **método** de tomada de decisão de um processo estático para um dinâmico, portanto, já era presumível que a análise das relações entre os diversos eventos derivados de dados de informação seria substituída por **métodos** de análise mais complexos, como Data Mining e Business Intelligence. Por outro lado, conforme Loh (2019) as ações complexas realizadas pelo homem foram transferidas para a Tecnologia da Informação (TI). Atualmente, podemos armazenar dados não estruturados, como vídeos, imagens, sons etc. Somado a estes benefícios, há a possibilidade de análises mais complexas com o desenvolvimento de *software* com funções de inteligência artificial. Todavia, não há mão de obra qualificada para utilizar os novos tipos de análises requeridos pelo *Big Data*, acarretando, desta maneira, um grande obstáculo para o *Big Data*. Assim sendo, especializar-se na aplicação dos **métodos** de *Big Data* é muito importante para as instituições públicas e privadas na realização estratégicas de tomada de decisões. Além do mais, no que concerne possibilidades futuras avançadas pelas aplicações de Big Data devem ser cuidadosamente analisadas, levando em consideração o seu alto grau de complexidade com o máximo conhecimento.

O quarto e último tema conectado como o *Big Data*, que de certa forma está relacionado ao nosso estudo, diz respeito ao **impacto** dos avanços do *Big Data* sob a sociedade. Como foi possível apurar, o *Big Data* não é caracterizado apenas pelo volume, variedade e velocidade dos dados. Ele representa também um **mecanismo estratégico de análise**. Quer dizer, ao coletar, organizar e transformar os dados em informações úteis e valiosas, gera *insights* importantes para as organizações (privadas e públicas) de forma a melhorar seus processos (identificação de novas oportunidades, redução de custos, análise do perfil de consumidores etc.). Contudo, a utilização do *Big Data* também tem **impactado** na vida privada das pessoas. Com a utilização de tecnologias sofisticadas, cria uma relação humana algoritima, modelando a forma como os cidadãos são abordados e identificados. Ou seja, por meio da utilização de um *software*, "[…] o comportamento humano é transformado em algoritmos e, sob um número, é reproduzido graficamente para ser codificado e analisado" (Neiva, 2020), gerando, desta forma, uma grande controvérsia entre o *Big Data* e o direito à privacidade, que será abordado na próxima secção.

Na figura 4, é possível visualizar os principais temas relacionados com o Big Data e suas conexões encontrados na literatura através do estudo de De Mauro et al. (2016).

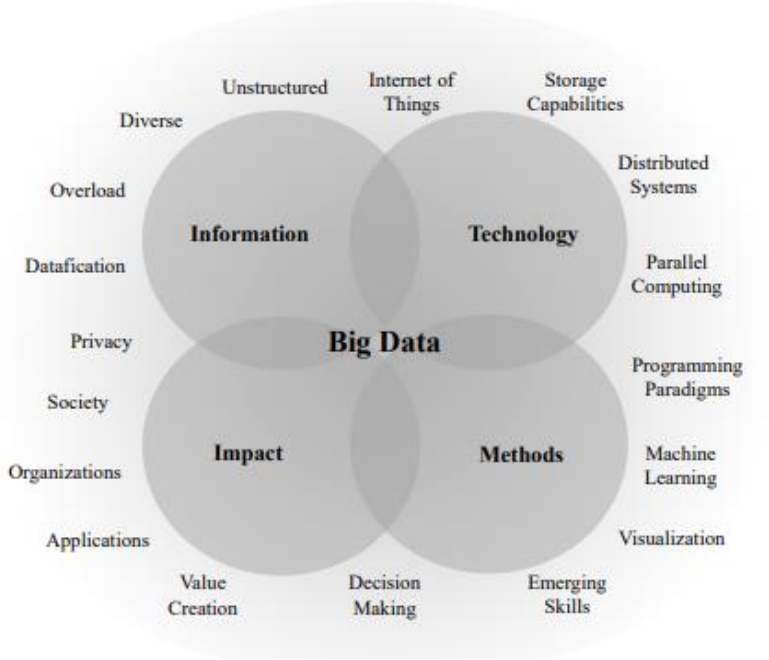

Figura 4- Temas do Big Data e tópicos relacionados na literatura existente

Fonte: De Mauro et al., 2016

Retomando a definição de *Big Data*, percebe-se que a evolução rápida e, às vezes até caótica sobre a literatura do *Big Data*, provavelmente tenha dificultado uma definição formal e universal. Vários autores têm proposto as suas próprias definições, resultando, assim, numa falta de concordância e homogeneidade para o termo (De Mauro et al., 2016). Nesta secção, pretende-se analisar algumas definições encontradas na literatura e vinculá-las aos quatro temas previamente discutidos acima, principalmente o tema relacionado ao atributo dos dados, de forma a possibilitar os pontos críticos que devem ser considerados para obter uma definição consensual do termo *Big Data*.

Conforme De Mauro et al. (2016), as definições de Big Data fornecem variadas perspetivas: descrito como um fenómeno social, ativo de informação, conjuntos de dados, tecnologias de armazenamento, técnicas analíticas, processos e infraestruturas. Estas definições, podem ser conectadas aos quatro temas acima discutido. Percebe-se que as definições dependem do foco de onde é colocado na descrição do fenómeno: I.

Atributos de Dados, II. Necessidades tecnológicas, III. Superação de Limites e IV Impacto Social.

Nos **atributos de dados**, segundo De Mauro et al. (2016), o foco está no conjunto de dados e recai sobre uma das definições mais populares para *Big Data*, ou seja, baseada nos 3 Vs de Laney (2001) que nada mais é do que as três dimensões dos dados: Volume, Velocidade e Variedade. Tendo em consideração a importância do valor e da veracidade das informações fornecidas pelo Big Data (Floridi, 2012), alguns pesquisadores (Erl et al., 2016; Ge et al, 2018; Loh, 2019) incluíram nas suas definições a Veracidade e o Valor. A veracidade foi incluída por entender que de nada adianta ter um grande volume de dados e não ter qualidade. Portanto, antes de qualquer análise, faz-se necessário garantir a veracidade dos dados. Relativamente ao valor. Os autores acreditam que nada adianta ter acesso a uma grande quantidade de informação/dados se ela não puder agregar valor. Vejamos, de forma, resumida a representatividade de cada uma desses cinco Vs.

O volume refere-se à quantidade de dados, que pode variar de terabytes a petabytes, o que são valores bem acima daqueles gerados por utilizadores usuais a qual estamos familiarizados (megabytes e gigabytes). A velocidade refere-se à rapidez com que os dados são gerados, transmitidos e processados, tudo em tempo real. A variedade refere-se à diversidade de fontes de dados, incluindo dados estruturados e não estruturados, como texto, áudio, vídeo e imagens. A veracidade refere-se à qualidade e confiabilidade dos dados. O valor respeita ao valor que os dados têm para as empresas ou entidades detentoras desses dados. Encontramos assim a equação final dos cinco Vs (ver figura 5), que caracteriza o Big Data (Loh, 2019).

Cabe salientar que existem outras definições na literatura para a conceitualização dos diferentes Vs do Big Data, onde diferentes autores defendem outros fundamentos de descrição e até apresentam modelos com mais de cinco Vs. Mas para o presente estudo, defendemos o conceito inicial dos três V´s de Laney (2001), a qual foi desenvolvido por Erl et al. (2016).

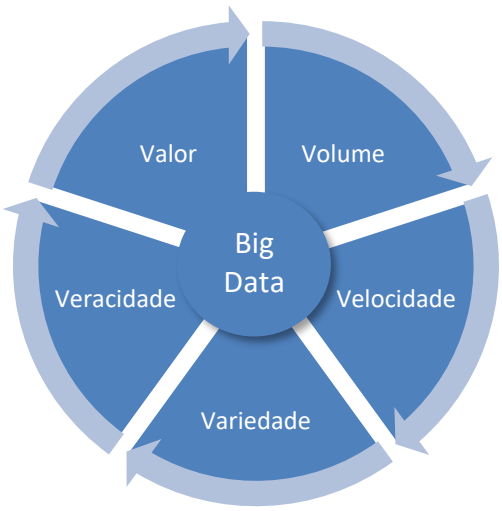

Figura 5- Os cinco Vs do Big Data.

Fonte: Autor com base em Erl (2016, p.29)

Após a extensa revisão da literatura sobre as definições para o *Big Data*, bem como os principais temas que se associam ao conceito do termo realizada por De Mauro et al. (2016), conclui-se que a essência do conceito do Big Data está intimamente relacionado com as seguintes perspetivas: a) Volume, Velocidade e Variedade para descrever as características da informação; b) Tecnologias e métodos analíticos para identificar requisitos essenciais de forma a realizar o uso adequado da informação e c) Fenómeno social para descrever as transformações tanto positivas, quanto negativas que a explosão de informação nos meios digitais vem trazendo para a sociedade como um todo. Por um lado, a informação serve como atributo para gerar *insights*, que podem oferecer valor económico para as empresas, de outro, a inserção de dados pessoais dos cidadãos em bancos de informações, poderá ferir o seu direito à privacidade.

Assim sendo, a conceitualização formal do *Big Data* para este estudo tem como base a definição proposta por De Mauro et al. (2016, p. 126), a qual defineo Big Data como um "[…] Information asset characterised by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value."

Uma outra definição que nos parece interessante registar é a definição proposta por Neiva (2020, p. 3), onde ela, ao conceituar o termo, coloca-o na esfera de um fenómeno social. Deste modo, a autora afirma que:

> *Big Data* pode ser perspetivado enquanto fenómeno cultural, tecnológico e académico que funde tecnologia (exacerba o poder computacional e a precisão algorítmica), com análise (visa identificar regularidades entre conjuntos de dados) e mitologia (crença generalizada de que oferece maiores formas de inteligência e conhecimento).

Antes de passarmos para o próximo tópico, acreditamos que vale a pena realçar um outro conceito de *Big Data* referido por Erl et al. (2016, cap. 1), no qual os autores afirmam que *Big Data* "[…] is a field dedicated to the analysis, processing, and storage of large collections of data that frequently originate from disparate sources".Erl et al. (2016) acrescentam que as soluções e aplicações de *Big Data* são requeridas para os casos em que as técnicas tradicionais não conseguem resolver. O *Big Data* atende a variados requisitos, quer dizer, tem a capacidade de combinar vários conjuntos de dados não relacionados, processar uma grande quantidade de dados não estruturados, bem como coletar informações ocultas de maneira sensível.

Embora a prática do *Big Data* possa parecer algo novo, há anos que se está a desenvolver. Ou seja, o problema da gestão e análise de um grande conjunto de dados não é um problema apenas da pós-modernidade, os governos dos estados e/ou dos países já faziam a utilização do censo para melhorar a saúde das pessoas e trazer avanços para a Ciência Social. As seguradoras de automóveis utilizavam as análises de grandes dados para prever os riscos sob os prémios dos segurados. Portanto, segundo Erl et al. (2016), a Ciência do Big Data evoluiu a partir destes contextos. Atualmente, o *Big Data*, em adição as técnicas analíticas tradicionais, traz novas abordagens para executar algoritmos analíticos através de recursos computacionais. Estas novas técnicas, devido a explosão da informação, onde conjuntos de dados estão, continuamente, tornando-se maiores, mais diversificados, mais complexos, é de suma importância.

1.2.1 Os desafios do *Big Data*

Seguindo o pensamento de Chen et al. (2014), a geração de dados é o primeiro passo do *Big Data*. Se focarmos no exemplo dos dados da Internet, há um sem-fim de dados gerados diariamente: registos de conversas ("chat"), de pesquisa, entradas em fóruns da internet e mensagens em microblogs. Estes dados coletados estão intimamente ligados à vida de cada uma dessas pessoas e, apesar de individualmente não apresentarem grande valor, através da exploração de grandes volumes de dados acumulados, tornam-se valiosos – podendo ser usados para prever comportamentos, estados emocionais e até escolhas dos utilizadores. Veldkamp & Chung (2019) defendem que o surgimento da economia de dados proveniente da *Big Data* tem promovido a fluidez económica global, e que esta era digital pode apresentar oportunidades para novos países enquanto potenciais líderes na economia. Concomitante a esta tendência, há também uma crescente preocupação sobre a melhor forma de regular o compartilhamento de dados entre fronteiras.

Mas o que significa a economia de dados? Silva (2019, p. 159)afirma que:

> [...]dentro de um sistema econômico maior, [a economia de dados] reflete um modelo de negócios pujante e robusto que possui estruturas inovadoras que estão sendo sedimentadas com seus respetivos impactos sociais e culturais no médio e longo prazo. Isso inclui a emergência e extração de novas matérias-primas; o surgimento e funcionamento de novos maquinários industriais; além do fortalecimento de novas formas de se racionalizar consumo ou, em outras palavras, gerenciar o comportamento do usuário/consumidor, adequando-os às dinâmicas da cadeia produtiva.

Esse modelo de negócios traz a datificação do tempo como *commodity* e a positivação do tempo como mercadoria. A datificação nada mais é do que o registo de uma ação ou fenómeno na forma de um dado, tornando-se um aspecto fundamental da economia de dados, devido, por exemplo, ao uso quotidiano de aparelhos digitais. Observa-se que mesmo nos países com baixo acesso à internet, a datificação tem uma popularização

expressiva, refletindo, desta forma, essa tendência global de coleta massiva de dados (Silva, 2019).

Contudo, os dados gerados a partir dessa datificação voluntária acarreta um problema para o *Big Data*. Os dados gerados são traduzidos em valores que, para o Big Data, são convertidos numa espécie de vigilância – o *dataveillance*. Esse *dataveillance* é definido por Silva (2019) como a produção de um tipo específico de matéria-prima, digamos, uma espécie de moeda principal da economia de dados. Isso, entretanto, acaba por ultrapassar o limite da privacidade dos indivíduos – pois a matéria-prima resultante é oposta à finalidade original do dado enquanto registo.

A datificação como *commodity,* portanto, acaba por ser posta em evidência aquando do debate sobre leis de proteção de dados, como foi o caso do Regulamento Geral da Proteção dos Dados (RGPD)[8] da União Europeia (2016/679), aprovada em 2016 e que entrou em vigor em Maio de 2018. Esse tipo de legislação, embora seja celebrada em caráter de proteção à privacidade do utilizador vem, na verdade, como uma forma de regular a obtenção da datificação desses dados, tendo em conta que o tratamento desses dados representa a nova forma de geração de riqueza para este milénio (Silva 2019).

Um estudo que vale ressaltar, pela pertinência de abordagem e junção deste conceito com as questões éticas e de privacidade, é o "Big Data ethics", do autor Zwitter (2014). Este estudo explora de que forma o Big Data impacta nas concepções morais. O estudo

[8] O Regulamento Geral sobre a Proteção de Dados (RGPD) é uma legislação da União Europeia que define as regras para o tratamento de dados pessoais de cidadãos na UE por empresas, organizações ou indivíduos. Esta regulamentação entrou em vigor em maio de 2018 e tem como objetivo proteger a privacidade e os direitos dos indivíduos em relação ao tratamento de seus dados pessoais (Parlamento Europeu, 2016).

sublinha, entre outras coisas, os princípios éticos da filosofia contemporânea e sua transformação devido a era do Big Data. Igualmente, o autor resgata a filosofia, a ética profissional, a elaboração de políticas e a publicação e investigação nesta temática, pois, de acordo com o autor, a ética está sempre em evolução para acompanhar os problemas dos dias atuais. Ainda neste estudo, são apresentadas quatro qualidades do Big Data com pertinência ética: a qualidade da quantidade de dados, sua organização, sua magnitude global e seu foco de destacar as correlações sobre a casualidade. Zwitter (2014) conclui o artigo afirmando que Big Data pode estar induzindo certas alterações a conjuntura tradicional da ética em relação à individualidade, ao livre-arbítrio e ao poder, colocando em evidência o risco as questões da privacidade da informação. Desta forma, aponta como alternativa: a) uma abordagem na esfera da educação, para conscientizar e educar sobre as consequências dos históricos digitais deixadas por crianças e adolescentes;b) nas esferas de leis e políticas, de modo que se possa realizar o desenvolvimento de observações da manipulação digital de opiniões públicas, possibilitando a investigação por observatórios políticos e pesquisadores; c) na área da aplicação de leis, serviços sociais e de pesquisa legais no sentido de haver a reconceituação da culpa individual, das probabilidades e prevenções de crimes. Por fim, d) na esfera estadual, de forma que os estados desenvolvem políticas globais baseada em dados globais e algoritmos ao invés de apoiar-se em opiniões enviesadas de especialistas dos setores.

1.3. Direito a privacidade

1.3.1 História do direito à privacidade da informação e suas legislações

Como vimos, o fenómeno do Big Data e a sua relação com a gestão da informação trouxe consigo inúmeras facilidades e benefícios, mas também algumas questões éticas que precisam ser discutidas e avaliadas. Uma delas é a privacidade da informação, apresentando-se como um enorme desafio a ser superado, gerando preocupações e interesse para a toda a sociedade.

Segundo Smith et al. (2011) a privacidade da informação é uma preocupação constante por parte de vários setores da sociedade, como sejam os empresários, os ativistas, os estudantes, os governantes e, principalmente, os consumidores individuais. Uma pesquisa realizada pela KPMG *Private Entrerprise Beyond*[9] revela que 98% dos consumidores estão preocupados com seus dados pessoais e o que, de facto, acontecem com eles. Certamente, esta preocupação do consumidor está relacionada com os avanços tecnológicos e a revolução da informação.

Gomez et al. (2009), no seuestudo, analisaram as políticas dos 50 sites mais visitados de forma a entender as divulgações sobre os tipos de dados coletados sobre os utilizadores, como essas informações são usadas e com quem são compartilhadas, descobriram que a maioria desses sites usam as informações pessoais dos utilizadores para fins de publicidade personalizada e que, empresas respeitáveis no mercado, como Google, Yahoo, Microsoft e Facebook compartilham os dados coletados de seus utilizadores com várias empresas associadas. Além disso, a maioria desses sites possuíam declarações pouco claras a respeito da privacidade da informação, visto que continham declarações de que as informações não poderiam ser compartilhadas com terceiros, mesmo assim, muitos desses sites, permitiam o rastreamento da informação de seus utilizadores.

Diante deste cenário, muitos estudos, no mundo inteiro, com destaque nos Estados Unidos e Europa, têm sido desenvolvidos no âmbito desta temática. Na Europa, agentes sociais e juristas têm voltado a ssuas atenções para a questão da privacidade. Muitos são os crimes cometidos neste mundo digital que desrespeitam o direito à privacidade, como a falsidade ideológica e perseguição de indivíduos através do roubo de informações pessoais em bases de dados públicas e privadas (Westin, 2003). Diante de

[9] https://itforum.com.br/noticias/estudo-98-dos-consumidores-estao-preocupados-com-dados-pessoais/.

tal fato, conclui-se que há muito ainda a ser feito de forma a garantir a privacidade das informações dos utilizadores numa sociedade da informação. Pode-se começar por entender a história da privacidade. Portanto, de seguida, iremos percorrer a história da privacidade e a sua legislação de forma a apontar um caminho mais adequado no tratamento da questão da privacidade. Sublinhamos que não é nossa intenção construir e propor um amplo modelo teórico sobre a privacidade da informação, mas apontar algumas diretrizes e relatar as dificuldades encontradas neste campo.

Em termos gerais, a privacidade pode ser classificada em: privacidade física e a privacidade informacional. Aquela se refere ao acesso e/ou arredores do espaço privado deste indivíduo. Esta diz respeito ao acesso as informações pessoalmente identificáveis. Salientamos que nesta subseção, iremos tratar diretamente deste último tipo de privacidade. Smith et al. (2011) esclarece que a definição de privacidade informacional surgiu a partir do conceito de privacidade física, uma vez que a Quarta revolução Industrial – caracterizada com a ascensão da internet, da Web 2.0 eInternet das Coisas (IoT), vem colocando a privacidade sobre os indivíduos noutro estágio.

Smith et al. (2011) observa que nos debates públicos e/ou nas pesquisas não há uma distinção entre privacidade física e privacidade informacional. Inclusive, os autores citam um exemplo: "[…] comments about privacy violations in the public media seldom draw a clear distinction between the constructs of physical and information privacy" (Smith et al., 2011, p. 991).

A privacidade é um termo multidisciplinar, visto que diversas disciplinas trazem sua conceção de privacidade, como a Economia, Psicologia, Marketing, Direito, Filosofia, Sociologia, Ciências políticas e Sistemas da informação. No entanto, na maioria dessas disciplinas, a privacidade está relacionada ao direito, quer dizer, o direito de ser deixado só (Borges e Machado, 2019). O primeiro movimento registado nesta perspetiva é datado de 1890, quando dois juristas dos Estados Unidos, Samuel D. Warren e Louis Brandeis escrevem o artigo "O Direito à Privacidade", no qual definiram a privacidade como o "direito de ser deixado em paz "(Smith et al., 2011, p. 994).

Cinquenta e oito anos depois, a definição de privacidade retorna, porém de forma mais relevante sob a égide de salvaguardar a vida familiar e pessoal dos indivíduos. Historicamente, este direito, representa a primeira definição geral de privacidade difundida, outorgado em 1948 pela Declaração Universal dos Direitos do Homem, onde o artigo 12º diz que "ninguém sofrerá intromissões arbitrárias na sua vida privada, na sua família, no seu domicílio ou na sua correspondência, nem ataques a sua honra e reputação. Contra tais intromissões ou ataques toda a pessoa tem direito a proteção da lei" (também consagrado na Convenção Europeia dos Direitos do Homem de 1950 e no Pacto Internacional relativo aos direitos civis e políticos de 1966) (Neiva, 2020).

Westin (2003), postula três fases de desenvolvimento da privacidade contemporânea: 1961-1979, 1980-1989 e 1990-presente. O autor afirma que a política de privacidade defronta-se com três grupos de interesse e orientações ideológicas, são elas: a) posição de alta privacidade – esta atribui valor primário as reivindicações de privacidade, tem alta desconfiança na organização e defende intervenções de privacidade por meio de aplicações de regras legais; b) posição de privacidade limitada – esta, geralmente, considera às reivindicações de privacidade pouco dignas em relação a eficiência dos negócios e dos interesses de proteção social, normalmente, confia nas organizações, mas se opõe a maioria das novas intervenções regulatórias, considerando-as desnecessárias e caras e, por último, c) posição de privacidade equilibrada – esta valoriza fortemente a privacidade por meio de intervenções legais que abordem os abusos certificados em conjunto com iniciativas voluntárias de políticas organizacionais destinadas a promover escolhas individuais de privacidade. Ao longo dos anos, as influências de todas estas posições têm variado. Vejamos, de forma resumida a posição da privacidade em cada um desses períodos.

Conforme Westin (2003) de 1945 a 1960, após a 2ª Guerra Mundial foram momentos de desenvolvimento da tecnologia de informação. Nesta altura, havia uma alta confiança no governo, nas empresas e nos setores sem fins lucrativos e, portanto, bem-estar do público em geral sob a coleta de informações, bem como no uso destas informações por parte dos organismos. Nesta época, embora houvesse importantes

lutas pelas liberdades civis e pelos direitos civis, não foi registado nenhum discurso seja do público em geral ou jurídico em torno do direito à privacidade da informação. Nesta época, a privacidade era reconsiderada como uma questão social de terceiro nível.

De acordo com Westin (2003), as preocupações sobre a privacidade da informação contemporânea, só ocorrerá, de facto, no período de 1961 a 1979. Nesta época, a privacidade era considerada como uma questão social, política e jurídica. Este período é caracterizado, para além do desenvolvimento da tecnologia de ponta, por momentos turbulentos, onde houve lutas pelos direitos civis, anti-guerra e outros protestos e movimentos sociais (por exemplo, a revolução sexual e Watergate). Neste clima de tensões, verificam-se avanços acerca das tecnologias de vigilância física, psicológica e de dados, dos quais foram amplamente adotados por instituições públicas e privadas. Todas essas melhorias no âmbito tecnológico foram aplaudidas, contudo, alguns comentadores / analistas, visualizando os riscos que essas novas tecnologias traziam para a questão da privacidade, anunciaram tais riscos, que foram amplamente divulgados.

Neste periodo, os Estados Unidos e outras nações democráticas reuniram comissões governamentais de maneira a iniciar estudos empíricos no setor privado. Estes estudos tinham como objetivo central entender a natureza, dinâmica e impactos das aplicações de tecnologia sobre a privacidade da informação e, a partir daí, buscar e aplicar novo equilíbrio de privacidade. Como resultado deste estudo, constatou-se que antes da década 70 – devido aos altos custos de processamento e armazenamento de dados, limitações de software e proteção organizacional de bancos de dados como ativos competitivos, as instituições ainda não possuíam o potencial de manipular os dados recolhidos dos indivíduos. No entanto, o estudo conseguiu prever que os custos reduzidos e o progresso dos *softwares* permitiriam que os organismos, no final da década de 70, obtivesem às informações pessoais com muito mais intensidade, transformando, assim, a tomada de decisão. Cabe sublinhar que neste período havia um movimento intenso de rejeição aos padrões de discriminação racial, religiosa, política e de género. Diante desta situação, o estudo também chegou à conclusão que era

essencial a criação de uma legislação que definisse os direitos à privacidade. A partir deste estudo, nos Estados Unidos, as práticas de privacidade tornaram-se dominante.

No âmbito sociocultural, este período presenciou uma mudança fundamental nas relações-públicas e privadas, no qual se verifica limitações de discriminações raciais, género, vida familiar e atividade sexual. Isto representou uma revisão de todos os negócios e sistemas de registo do governo que incorporassem os critérios antigos, transformando as características pessoais em assuntos privados.

No setor privado, a presença de empresas com sistemas de registos fechados e aplicações suspeitas, somado a automação de dados pessoalmente identificados causou muitas preocupações no público em geral. Este alarme, levou à promulgação da primeira lei federal de práticas de informações justas - o *Fair Credit Reporting Act*, em 1970. Esta lei fornecia aviso prévio, acesso do consumidor e direitos de correção, porém não definia padrões de relevância ou privacidade para relatórios de consumidores.

No contexto político, nos finais dos anos 60 e na década de 70, os excessos cometidos pelo FBI e pela CIA, bem como o episódio de Watergate e outras abusos de poder por parte do Presidente Nixon, fez com que a opinião pública mudasse drasticamente, de total confiança nas instituições para desconfiança crítica. Esta situação, possibilitou a promulgação da Lei Federal de Privacidade, em 1974. Outras leis federais se seguiram, como o *Family and Educational Rights and Privacy* (1974) e a Lei do Direito à Privacidade Financeira (1978). Além disso, verificam-se outros avanços no âmbito da privacidade, a título de exemplo, por decisão da Suprema Corte de Warren, as escutas telefónicas passaram a ser permitidas sob ordem judicial com legislação de salvaguardas operacionais. Sem dúvida, foi um período de grandes mudanças, talvez a mais significativa tenha sido – embora não previsto em lei, a instituição de um consenso que determinava que os bancos de dados não devem ser autorizados a consolidar as informações dos cidadãos de arquivos separados de agências governamentais locais ou nacionais, mesmo que isso possa fornecer um quadro mais completo das relações dos cidadãos com o governo.

Em resumo, as décadas de 60 e 70 são caracterizadas por uma época de estudos substanciais acerca da privacidade, onde novas metodologias analíticas de impacto social da tecnologia e leis criativas de privacidade da era da informação foram desenvolvidas. Esta era é denominada por Westin (2003) de primeira era do desenvolvimento da privacidade.

De acordo com Westin (2003), a segunda era do desenvolvimento da privacidade ocorrerá no período de 1980–1989. Esta época é caracterizada pelo desenvolvimento aperfeiçoado de computadores e telecomunicações, contudo, sem mudanças radicais na sociedade da informação, uma vez que não houve grandes desenvolvimentos nas tecnologias de vigilância física ou psicológica. Portanto, não afetou a questão da privacidade, até porque "[…] PCs were unconnected to the larger world" (Westin, 2003, p. 439). No que se refere às atividades empresariais e governamentais, envolvendo a recolha e o uso de informações pessoais não houve grandes mudanças. Podemos dizer que a mudança mais radical neste âmbito foi a Lei Federal de Proteção de Privacidade e Correspondência de Computadores de 1988, a qual criou proteção de arquivos de forma a controlar a troca de dados entre agências federais, por exemplo, para verificar registos de dados públicos de forma a detetar fraudes.

Politicamente, segundo Westin (2003), a privacidade permaneceu uma questão de política social de segundo plano. A legislação do direito à privacidade foi defendida, principalmente, por grupos liberais tradicionais, como o *American Civil Liberties Union* (ACLU), sindicatos e organizações de consumidores, como a Liga Nacional dos Consumidores. Nesta época, dois boletins influentes sobre privacidade – o *Privacy Journa*l de Robert Ellis Smith e o *Privacy Times* de Evan Hendrick — circulavam, fornecendo orientações acerca da defesa do direito à privacidade. Ainda neste período, surge uma pesquisa nacional "The Road to 1984", escrita por Louis Harris and Associates, em 1984, na qual é feita uma análise profunda sobre a privacidade. Neste estudo, é detetado o caráter paradoxal das novas tecnologias de informação e comunicação, trazendo inúmeros benefícios e conveniências para a sociedade, mas preocupaçõescrescentes sobre o uso indevido e abusos. Enfim, nesta época, verifica-se:

"[…] a significant amount of federal legislation adopted to channel new technology applications or new governmental activities into fair information practices or privacy protection frameworks" (Westin, 2003, p. 440). Nestas legislações, podemos encontrar:

- **Lei de Proteção de Privacidade de 1980**, exigia uma base razoável para suspeita de que um crime foi cometido antes que agências governamentais possam fazer buscas não anunciadas em assessorias de imprensa;
- **Cable Communications Policy Act de 1984**, exigia que as empresas de cabo informassem aos assinantes sobre práticas de coleta de informações;
- **Lei de Privacidade de Comunicações Eletrônicas de 1986** (extensão da lei de 1968), que determinava que as escutas telefônicas somente poderiam ser realizadas por ordem judicial, bem como medidas de procedimentos de controle para dados de voz digital e comunicações de vídeo;
- **Video Privacy Protection Act de 1988**, proibia as locadoras de vídeo de divulgar nomes e endereços de seus clientes, bem como os vídeos que os clientes alugam e/ou compravam.

Embora todas estas leis federais não consigam atender todos os casos relacionados a privacidade, nem abrange todos os indivíduos, mesmo assim consegue demonstrar a relevância da privacidade como uma questão política.

O certo é que, enquanto nos EUA permaneciam práticas de informação justa e abordagem regulatória setorial, as nações europeias, no início do ano de 1970, seguiram uma abordagem diferente, ou seja, instituíram leis nacionais de proteção de dados que envolvia todos os setores governamentais e privados, sob agências nacionais independentes de proteção de dados (Flaherty, 1979; Schwartz & Reidenberg, 1996, citados por Westin, 2003). No final dos anos 80, assiste-se, em muitas nações europeias (Áustria, Dinamarca, França, República Federal da Alemanha, Luxemburgo, Noruega e Suécia) a promulgação da lei geral de proteção de dados (Simitis, 1987, citado por Westin, 2003). Na realidade, a promulgação dessas leis é fruto da constante preocupação da "[…] introdução da tecnologia de informação em várias áreas da vida económica e social, e a importância e poder crescentes do processamento automatizado

de dados (OCDE; 2002, p 2). Esta evidência é constatada, quando em 1980, a Organização para a Cooperação e Desenvolvimento Econômicos (OCDE) decidiu publicar Diretrizes relativas à política internacional sobre a proteção da privacidade e dos fluxos transfronteiriços de dados pessoais.

A terceira era do desenvolvimento da privacidade, segundo Smith et al. (2011) corresponde ao ano de 1999 até os tempos atuais. Neste período, nos EUA, a privacidade é considerada como uma questão social e política de primeira ordem. Este facto, deu-se, principalmente, por conta dos ataques terroristas de 11 de setembro e pelo avanço crescente das tecnologias, tais como, a ascensão da internet, o surgimento dos dispositivos sem fio, à interface de privacidade da tecnologia com o desbloqueio do código genético pelo Projeto Genoma Humano, o desenvolvimento de software de mineração de dados baseado em grandes aplicativos de armazenamento de dados, juntamente com mais automação dos sistemas de registos públicos do governo.

Todas essas novas tecnologias permitiram muitos recursos e proteção para a sociedade em geral, contudo, indiciam perigos ao direito de privacidade. Por exemplo, sites da *Web* usam dispositivos de rastreamento, como cookies, para identificar visitantes e documentar seu uso. Dispositivos de comunicação sem fio trouxeram a capacidade de aplicadores da lei governamental ou litigantes privados para localizar utilizadores individuais por tempo e lugar através da tecnologia móvel, e também a potencialidade das empresas enviarem mensagens de marketing para utilizadores sem fio com base no conhecimento de sua localização perto de determinados estabelecimentos comerciais. Um outro exemplo do potencial das tecnologias e a ameaça à privacidade pode ser visualizado por meio do uso global de cartões de crédito juntamente com o uso da Internet, onde as empresas multinacionais conseguem coletar, armazenar, processar e analisar informações sobre os clientes.

Visando estas e outras questões, a União Europeia (UE) implementou a Diretiva de Proteção de Dados da EU (Diretiva da UE 95/46/CE, 1995) de forma a "[…] assegurar o respeito à privacidade e a proteção dos dados pessoais em linha" (OCDE, 2002, p. 2). Esta diretiva tinha como objetivo principal impedir que dados pessoais de consumidores

e funcionários fossem transferidos por empresas multinacionais, a menos que esse país tenha o que a UE considera um regime de proteção de dados adequado ou qualificado para procedimentos especiais. Esta diretiva estabeleceu um sistema abrangente de privacidade de informação, impactando várias outras nações.

Assim sendo, preocupada com a proteção da privacidade e dos fluxos transfronteiriços de dados pessoais, a UE procurou salvaguardar o tratamento de dados nos países em desenvolvimento. Esta influência pode ser atestada na Lei de Alteração de Privacidade da Austrália de 2000 – que foi modelada com base nos princípios europeus – e no acordo Safe Harbor de dados pessoais (2000) entre a UE e os Estados Unidos. O texto ratificado pode ser encontrado em seu Artigo 7º que reza o seguinte: "todas as pessoas têm direito ao respeito pela sua vida privada e familiar, pelo seu domicílio e pelas suas comunicações". De acordo com Neiva (2020) este é um direito sustentado na dignidade da pessoa humana.

Ainda neste período, segundo Pilati e de Olivo (2014) Edward Snowden, analista de sistemas e ex-funcionário da Agência de Segurança Nacional (ASN) nos Estados Unidos da América (EUA), foi o responsável pela descoberta acerca da coleta de dados pessoalmente identificáveis realizada pela ASN. Esta constatação abriu um grande debate público sobre as novas formas de violação ao direito fundamental à privacidade. De acordo com Neiva (2020) esta descoberta sociopolítica causou indignação e ansiedade nas pessoas, comprometendo os direitos, garantias e liberdades dos cidadãos.

Depois deste acontecimento, estudos observacionais das Ciências Sociais e Humanas relatam que os indivíduos passaram a apresentar atitudes paradoxais[10], ou seja, por um

[10] Atitude paradoxais, denominada de *privacy paradox* é quando a violação da privacidade se dá pelo próprio consentimento do usuário.

lado, busca-se, cada vez mais, parâmetros que garantam a não invasão da privacidade dos indivíduos e, por outro, constatam-se pessoas expondo a sua vida privada para o público em geral (Borges & Machado, 2019; Neiva, 2020; Smith et al., 2011).

Em 2016, assiste-se à promulgação da legislação que ratifica a importância da privacidade informacional. O Parlamento Europeu adotou a General Data Protection Regulation (GDPR), ou em português "Regulamento Geral sobre a Proteção de Dados (RGPD)". Este será objeto de nossa apreciação no final desta subsecção, contudo, antes, pretende-se esclarecer a definição de privacidade.

1.3.2 Definição de privacidade

A privacidade pelo facto de ser um termo multidisciplinar, traz um conceito de difícil definição (Borges & Machado, 2019). Para Waren e Brandeis (1890), a privacidade é o direito de ser deixado só. Enquanto Altman (1975, citado por Borges & Machado, 2019, p. 244) afirma que privacidade é o "controle seletivo de acesso a si mesmo". De acordo com Westin (2003, p. 431), a privacidade é "[…] the claim of an individual to determine what information about himself or herself should be known to others". "[…] This, also, involves when such information will be obtained and what uses will be made of it by others".

Conforme Westin (2003), esta dimensão da privacidade reflete as necessidades e desejos particulares de cada indivíduo. Estas necessidades, por sua vez, podem sofrer mutações a depender do progresso do ciclo de vida e de eventuais situações que podem ocorrer ao longo do tempo. Quer dizer, num momento a pessoa pode querer resguardar as suas informações, em outros pode querer abrir suas informações para um completo estranho, que possivelmente nunca mais o encontrará e, portanto, não sofrerá autoridade de julgamento desta pessoa. No ponto de vista do autor, este estado particular de cada indivíduo, confere a privacidade uma condição em termos de complexidade e de escolha pessoal. O direito de escolha – tanto para o autodesenvolvimento do indivíduo, quanto para o exercício de cidadania responsável, outorga o direito à privacidade como parte fundamental da vida civil em sociedade. Por

outro lado, o fato de haver uma proibição legal em aceder a informação pessoal, bem como, propriedade individual, a privacidade está intimamente conectada ao nível político (Neiva, 2020). Sobre outra perspetiva, um indivíduo para obter subsídio do governo, necessita, para não ficar sem o benefício, revelar suas informações. Nesse sentido, a questão da privacidade também está vinculada ao nível social (Westin, 2003). Além do mais, a privacidade está relacionada com o controle, quer dizer, com o direito do indivíduo de controlar suas informações pessoais, bem como o direito de permanecer no anonimato (Smith et al., 2011). Ainda, segundo Smith et al. (2011), a privacidade se conecta a uma mercadoria, tendo em vista que o indivíduo ao fornecer voluntariamente informações *online* (a chamada autovigilância), está cooperando na coleta de dados sobre si mesmos como sujeito económicos. No ponto de vista dos autores, essa participação dos indivíduos na vigilância "[…] is possible because of recent reconceptualization of privacy in the consumer's mind from a right or civil liberty to a *commodity* that can be exchanged for perceived benefits […]" (Smith et al., 2011 p. 994).

Como vimos, não há um campo específico e imutável em que se possa situar a privacidade e, portanto, existirá sempre uma variação material a depender de quem seja o titular do direito e do ambiente em que ele se situa, conferindo à privacidade, contornos próprios. Tendo em conta estas características peculiares da privacidade, somado a realidade extraordinariamente complexa em que estamos vivendo, resultante da revolução da informação e, com esta, a globalização das redes de contatos, a massificação da informação e o crescente acesso aos dados pessoais dos indivíduos, aventuramo-nos numa definição de privacidade. Assim sendo, para esta pesquisa, entende-se que a privacidade é o direito intrínseco ao indivíduo de preservar, controlar os mais diversificados aspetos de sua vida privada – ou seja, a sua intimidade, o seu segredo, suas relações profissionais e sociais, familiar, académica (inclusive as produções científicas com ou sem valor literário), suas convicções, crenças, valores, confidências, hábitos, sua vida afetiva, seus negócios, património, suas comunicações – de outros indivíduos, empresas e organismos, no contexto das novas tecnologias.

1.3.3 Análise crítica do Regulamento Geral sobre a Proteção de Dados (RGPD)

A privacidade informacional como direito individual protege algo valioso como a autodeterminação informativa. Por outro lado, como um direito instrumental, protege outros bens e interesses derivados, como a própria base digital de dados (Filho & Schwartz, 2016). Contudo, estes direitos podem ser lesados quando estas bases digitais de dados são submetidas as soluções de *Big Data*.

Diante desta e de outras realidades, a Comunidade Europeia aprovou, em 27 de abril de 2016, o Regulamento Geral sobre a Proteção de Dados (RGPD, ou em inglês, *GDPR* – *General Data Protection Regulation* – Regulamento 2016/679). Este regulamento estabelece as regras relativas à proteção de dados pessoais de pessoas singulares, sendo extensível diretamente na ordem jurídica de todos os Estados-Membros, impondo uma série de deveres que se destinam, especialmente, a pessoas coletivas públicas. Em Portugal, no sentido de assegurar o RGPD (Regulamento Geral sobre a Proteção de Dados) foi instituída a Lei nº 58/2019.

Contudo, este direito, no contexto do *Big Data* – onde são gerados grandes volumes de dados e informação, estes se apresentam não só como matéria-prima da informação, mas como uma enorme fonte de valor económico e social – é normalmente ponderado. Portanto, a proteção de dados nem sempre se mostra absoluta. Assim sendo, nesta subsecção, pretende-se analisar a definição de dados pessoais elencado no Regulamento Geral sobre a Proteção de Dados (RGPD) com ênfase nas políticas de acesso à informação e a privacidade.

O novo RGPD em seu artigo 4º, define dados pessoais como sendo:

> Qualquer informação relativa a uma pessoa singular identificada ou identificável («titular dos dados»); é considerada identificável uma pessoa singular que possa ser identificada, direta ou indiretamente, em especial por referência a um identificador, como por exemplo um nome, um número de identificação, dados de localização, identificadores por via eletrônica ou a um ou mais elementos

específicos da identidade física, fisiológica, genética, mental, económica, cultural ou social dessa pessoa singular[11].

Relativamente a este conceito, é importante observar quatro termos – "qualquer informação", "relativa a", "pessoa singular" e "identificada ou identificável" para entender que tipo de informação pode ser considerada como dados pessoais.

O primeiro termo, "qualquer informação", encontrado no conceito de dados pessoais, revela a amplitude, de salientar que não há uma discriminação do tipo de informação, quer dizer, abrange todos os dados, independentemente de sua natureza ou do conteúdo da informação, podendo ser objetivos e/ou subjetivos e, apresentar-se-à por diversos meios técnicos- tecnológicos.

No nosso ponto de vista, este conceito aberto e fluído do termo "qualquer informação", por um lado, apresenta-se como um ponto positivo, já que permite maior flexibilidade na resposta jurídica a ser dada em diferentes circunstâncias e, nesse sentido, à doutrina e a jurisprudência podem "ajustar" o conceito do termo levando em consideração as mudanças tecnológicas da época. Por outro lado, revela alguns riscos, como seja o de adotar visões muito amplas ou, contrariamente, consentir uma exceção inadequada de seu alcance.

Com a disponibilidade da internet pela *web*, tornou-se fácil obter informações e dados dos utilizadores. Os notariados, os hospitais, seguradoras e bancos recolhem informações sobre o histórico familiar, financeiro e de saúde; as empresas de telecomunicações, como por exemplo, a Vodafone, possui informação dos números telefónicos mais marcados pelos seus clientes e da frequência destas ligações; as livrarias têm informações sobre hábitos de leitura, possibilitando elaborar perfis

[11] Conferir em: https://gdpr-text.com/pt/read/article-4/. Acesso em 26/03/2023.

literários e económicos de seus clientes; as operadoras de cartão detêm informações sobre perfis de consumo de seus clientes; diversos departamentos de retalho possuem cadastro de seus clientes e mantêm histórico de consumo de produtos, criando assim um perfil de compra personalizado para cada cliente, podendo encaminhar uma publicidade direcionada; fornecedores de internet mantêm registro de acesso a sites, envio e recebimento de e-mails e preferências de conteúdo acedido. Em suma, a intersecção destas informações, cria uma oportunidade para a obtenção de um banco de dados, que especifica as características, hábitos e as atividades quotidianas do utiizador[12].

O segundo termo, "relativo a", assim como o termo "qualquer informação", apresenta-se amplo, salienta-se que pode gerar dúvidas se as informações são referentes ou "relativo à" pessoa singular. Um exemplo seria a informação sobre algo, sobre um objeto, como por exemplo, o valor de um automóvel, que a princípio não se refere diretamente a pessoa singular, mas que em alguns casos, pode revelar as condições financeiras do indivíduo.

O terceiro termo, "pessoa singular", refere-se a pessoa humana e, portanto, o RGPD é aplicável dentro da União Europeia a qualquer ser humano independentemente da sua etnia e/ou nacionalidade. No entanto, o regulamento exclui a proteção a pessoa coletiva (por exemplo, associações, fundações e organizações).

O quarto e último termo, "identificada ou identificável", refere-se a um dado pessoal, ou seja, qualquer informação relativa "a uma pessoa viva, identificada ou

[12]Disponível em: Privacidade na Internet: maio 2015 (privacidadenainternetrony.blogspot.com), acesso em 30/30/2023.

identificável”[13]. Os dados pessoais são constituídos por duas tipologias, a saber: dado pessoal sensível e dado pseudonimizado.

Dado pessoal refere-se a qualquer informação que permite identificar, direta ou indiretamente um indivíduo. Identificação direta diz respeito as características e os atributos que identificam explicitamente uma pessoa, como exemplo: o nome, número do cartão do cidadão, data e local de nascimento, número do telemóvel, endereço residencial, endereço eletrónico (como e-mail, constituído por nome.apelido@empresa.com). Já a identificação indireta corresponde os dados que os sites de vendas de retalho, as operadoras de cartão de crédito, as operadoras de telecomunicações, imagem fotográfica nas redes sociais, histórico da localização via GPS etc., recolhem de forma a identificar os hábitos, costumes e perfis dos utilizadores.

Já o dado pessoal sensível refere-se sobre a origem racial e étnica do indivíduo, convicção religiosa, filosófica, política, dado referente à saúde ou à vida sexual, dado genético ou biométrico.

Por último, cabe destacar que, de acordo com o Regulamento Geral de Proteção de Dados (RGPD), o dado pseudonimizado é aquele que passou por um tratamento de dados pessoais de forma que deixou de poder ser atribuído a um titular de dados específico sem recorrer a informações suplementares. No entanto, essas informações adicionais devem ser mantidas separadamente e estar sujeitas a medidas técnicas e organizacionais que garantam que os dados pessoais não possam ser associados a uma pessoa singular identificada ou identificável. Por exemplo, imaginemos uma empresa que faz entrega de produtos. Essa empresa processa dados, tais como: a distância

[13] Disponível em: https://commission.europa.eu/law/law-topic/data-protection/reform/what-personal-data_pt#:~:text=Dados%20pessoais%20s%C3%A3o%20informa%C3%A7%C3%A3o%20relativa,identifica%C3%A7%C3%A3o%20de%20uma%20determinada%20pessoa, acessado em 30/03/2022.

percorrida pelos motoristas, a frequência e os tipos de viagens realizadas. Estes dados são dados pessoais, pois são dados que dizem respeito aos motoristas. A empresa utiliza esses dados para calcular as despesas dos gastos com o transporte e, portanto, obter uma base para cobrar os clientes. No entanto, a identificação dos motoristas é imprescindível para a realização deste estudo. Entretanto, se outro departamento desta mesma empresa quiser utilizar esses mesmos dados para otimizar os serviços da frota, a identificação dos motoristas já não se faz necessário. Neste caso, cabe a empresa resguardar os dados dos motoristas. Para tal, as empresas costumam utilizar o método de anonimização para substituir identificadores tais como nome, cargo e histórico de navegação por um identificador artificial, um pseudónimo, como por exemplo um código, que, por si só, não representa nenhuma informação[14].

Ou seja, neste caso, os colaboradores deste departamento, só terão acesso aos dados pseudonimizados. Decorre que, de acordo com a RPGD a partir do momento que o dado não permite mais qualquer identificação do seu titular, caso dos dados anonimizado, esse dado, por não ser considerado dado pessoal, não está protegido pela lei. Vejamos o que prevê no RGPD:

> "(…) Os princípios da proteção de dados não deverão, pois, aplicar-se às informações anónimas, ou seja, às informações que não digam respeito a uma pessoa singular identificada ou identificável nem a dados pessoais tornados de tal modo anónimos que o seu titular não seja ou já não possa ser identificado. O presente regulamento não diz, por isso, respeito ao tratamento dessas informações anónimas, inclusive para fins estatísticos ou de investigação."

[14]Disponível em: https://www.migalhas.com.br/coluna/migalhas-de-protecao-de-dados/332299/o-dado-pseudonimizado-e-um-dado-protegido-pela-lei-geral-de-protecao-de-dados, acesso em 30/03/2023.

Todavia, a empresa, na qualidade de controladora dos dados, tem capacidade de reverter o processo que obteve a anonimização e, desta forma, voltar a reidentificar uma pessoa. Nesse sentido, não estamos diante de um dado verdadeiramente anonimizado, mas de um dado pseudonimizado. O departamento usado como exemplo, se não conhecer o processo de criptografia ou a chave, os dados são ininteligíveis. Mas, o fato de ter sofrido a reversão "continuam a ser dados pessoais e são abrangidos pelo âmbito de aplicação do RGPD"[15]. Logo, nesse sentido, "para que os dados sejam verdadeiramente anonimizados, a anonimização tem de ser irreversível"[16].

Assim sendo, conclui-se que de acordo com a lei, um dado pessoal é todo o dado capaz de tornar o indivíduo identificável, por seu turno, entende-se que dados pseudonimizados são também dados pessoais, visto que possuem o poder de reindentificar a pessoa singular e, portanto, este tipo de dado, também está protegido pelo RGPD.

Tal entendimento vai ao encontro dos objetivos da Carta dos Direitos Fundamentais da União Europeia que, em seus artigos 7º e 8º, reconhece [...] a todas as pessoas singulares [...] o respeito pelos seus direitos e liberdades fundamentais, e especialmente pelo seu direito à vida privada, face ao tratamento automatizado dos dados de caráter pessoal»[17]. Também encontra respaldo nas diretrizes da OCDE (2002) para a proteção da privacidade e dos fluxos transfronteiriços de dados pessoais, que elenca em suas diretrizes: democracia pluralista, respeito aos direitos humanas e também, economias de mercado aberto.

---

[15] Idem.

[16] Disponível em: O que são dados pessoais? (europa.eu), acessado em 30/03/2023.

[17] Disponível em: https://www.europarl.europa.eu/ftu/pdf/pt/FTU_4.2.8.pdf, acesado em 01/04/2023.

As diretrizes sobre a privacidade representam:

> um consenso internacional sobre a orientação geral a respeito da coleta e da gestão da informação pessoal. Os princípios determinados nas Diretrizes sobre a Privacidade são caracterizados pela clareza e flexibilidade de aplicação e pela formulação, suficientemente ampla para possibilitar a adaptação às mudanças tecnológicas. Esses princípios abrangem todos os meios utilizados para o processamento automatizado de dados referentes a indivíduos (do computador local à rede de complexas ramificações nacionais e internacionais), todos os tipos de processamento de dados pessoais (da administração do pessoal ao levantamento de perfis de consumidores) e todas as categorias de dados (da circulação de dados ao seu conteúdo, dos mais comuns ao mais sensíveis). (OCDE, 2002). Sublinhamos que estas diretrizes se aplicam tanto a nível nacional quanto internacional.

As orientações constantes deste documento está dividida em cinco partes: Generalidades (definições e alcance das diretrizes); Princípios básicos de aplicação nacional (Princípio de limitação da coleta, Princípio de qualidade dos dados, Princípio de definição da finalidade, Princípio de limitação de utilização, Princípio do back-up de segurança, Princípio de abertura, Princípio de participação do indivíduo, Princípio de responsabilização); Princípios básicos de aplicação nacional: livre fluxo e restrições legais; Implementação nacional, Cooperação internacional. No nosso ponto de vista estas diretrizes aliadas ao RGPD apresentam-se como um arcabouço para assegurar o respeito à privacidade e a proteção dos dados pessoais numa sociedade global de informação. Contudo, há a necessidade de garantir que os sistemas de informação utilizados e governança das empresas obedeçam aos princípios de proteção de dados e protejam os direitos dos indivíduos detentores desses dados.

2. Metodologia da investigação

2.1 Escolhas metodológicas e técnicas de pesquisa

Ao conceber uma investigação, o investigador, antecipadamente, necessita identificar que caminho a percorrer e de que modo este caminho será percorrido. Quer dizer, precisa definir as suas escolhas metodológicas e as técnicas de pesquisa a serem adotadas na investigação. As escolhas metodológicas estão relacionadas diretamente com os objetivos, objeto e com a naturezada pesquisa, já as técnicas de pesquisa se relacionam com a coletae análise dos dados (Oliveira, 2011).

Assim sendo, relativo aos objetivos da pesquisa, esta se enquadra na categoria dos estudos exploratórios. A pesquisa exploratória não requer a formulação de hipóteses a serem testadas no estudo, ao contrário, com os resultados dos estudos, faz emergir hipóteses significativas para pesquisas posteriores. Desta forma, este tipo de pesquisa tem como objetivo principal familiarizar-se com o fenómeno no sentido de desenvolver, esclarecer e descobrir novas ideias (Oliveira, 2011, p. 20). Posto isto, o trabalho aqui apresentado, para além de ter pesquisado informações acerca do fenómeno estudado, também teve como objetivo diagnosticar a situação sobre a temática da informação e à privacidade no contexto do Big Data.

Em relação à natureza da pesquisa, esta é classificada como pesquisa qualitativa. Este género de abordagem propicia um aprofundamento das questões relacionadas com o fenómeno estudado, valorizando o contato direto com a situação estudada (Gil, 1999, citado por Oliveira, 2011). Assim sendo, a preocupação principal deste estudo foi o processo em vez do produto, quer dizer, interessou-nos diagnosticar, explorar alternativas ou descobrir novas ideias sobre o fenómeno da informação e a privacidade no contexto do Big Data.

Quanto à escolha do objeto de estudo, a nossa pesquisa classifica-se em estudo de caso, uma vez que procurámos aprofundar e detalhar os fatos, objeto do nosso estudo, de forma a permitir um amplo e pormenorizado conhecimento da realidade e do fenómeno estudado. Quer dizer, a partir dos artigos elegíveis para o estudo, tentou-se esclarecer

as decisões que se manifestaram e foram construídas nestes artigos. Estamos cientes que num estudo de caso não é possível fazer generalizações para uma população, porém, este trabalho servirá de base para outros estudos.

De acordo com Oliveira (2011, p.28), “o estudo de caso pode ser restrito a uma ou a várias unidades, caracterizando-o como único ou múltiplo”. No caso específico do nosso estudo, englobámos mais de uma unidade de análise, isto é, para além de termos estudado individualmente cada um dos artigos elegidos para o estudo, fizemos uma análise comparativa entre os casos, ou seja, entre os artigos.

Concluído a caracterização das escolhas metodológicas, passamos a identificar as técnicas de coleta e de análise dos dados. Quanto a técnica de coleta de dados, utilizou-se a pesquisa bibliográfica por meio da Revisão Sistemática da Literatura (RSL). Esta será aprofundada mais adiante.

E, por fim, relativo a técnica de análise dados, adotou-se a meta-síntese, uma vez que o nosso propósito foi fazer uma integração interpretativa dos artigos selecionados (dos Reis et al., 2017). Por meio da meta-síntese, procurou-se desmontar a estrutura e os elementos do conteúdo dos artigos que compuseram a seleção final, com vista a esclarecer o sentido e o significado que cada autor atribuiu para a problemática estudada.

Quanto a operacionalização do método de análise, primeiramente realizamos uma leitura superficial de todos os artigos que compuseram a seleção final da RSL, de modo a recortar o conteúdo destes em fragmentos que traduziam uma ideia particular relacionada com os objetivos da pesquisa. O passo seguinte ao recorte de conteúdo, realizou-se uma leitura aprofundada destes mesmos artigos, procurando agrupar os elementos de conteúdo por parentesco de sentido às categorias analíticas. Há três modos de definição destas categorias, a saber: o modo aberto, o modo fechado e o modo misto. No caso específico deste estudo, usou-se o modo fechado, uma vez que escolhemos previamente as categorias de análise baseadas nos objetivos (modelo teórico) definido neste estudo.

Nestas próximas linhas, discorrer-se-á acerca da nossa opção pela técnica de recolha de dados, Revisão Sistemática da Literatura (RSL).

2.2 Técnica de Recolha de Dados: RSL

De acordo com Mancini & Sampaio (2007), a RSL é um tipo de estudo de revisão, que utiliza a própria literatura como fonte de dados. Este tipo de investigação faz uso de métodos sistemáticos e explícitos para identificar, selecionar e avaliar criticamente estudos relevantes para uma pergunta de pesquisa explícita, e coletar e analisar dados desses estudos para incluir na revisão (The PRISMA Group, 2015). Os investigadores, normalmente, fazem uso desse tipo de método quando objetivam integrar as informações de um conjunto de estudos individuais (primários) relacionados a uma pergunta de investigação específica, que apresentam resultados conflitantes e/ou coincidentes. Estes estudos, ao passarem pelo processo do método de RSL, passam a compor os estudos secundários e, estes, por sua vez, têm a incumbência de identificar temas que necessitam de evidências, apoiando, assim, investigações futuras (Galvão & Ricarte, 2019; Higgins & Green, 2011; Mancini & Sampaio, 2007; The PRISMA Group, 2015).

Segundo as orientações PRISMA (2015), as RSL podem ser analisadas com base na abordagem qualitativa e/ou quantitativa. Naquela, ao contrário desta, ocorre quando o investigador não faz uso de técnicas estatísticas para informar os resultados, exceção da apresentação da amostra. Quando o investigador utiliza técnicas estatísticas numa revisão, estas são denominadas de meta-análise.

Alguns autores sublinham que o método RSL possuem muitas vantagens, por exemplo:

> […] limita o viés dos estudos existentes, e também melhora a confiabilidade e a precisão das recomendações, por meio da combinação de informações de estudos individuais, além de possuir uma dimensão da amostra total que é maior do que a de qualquer um dos estudos sobre o tema específico (Roever, 2017, p. 127)

Em conformidade a Roever (2017), Mancini e Sampaio (2007), realçam que:

> […] as revisões sistemáticas nos permitem incorporar um espectro maior de resultados relevantes, ao invés de limitar as nossas conclusões à leitura de somente alguns artigos. Outras vantagens incluem a possibilidade de avaliação da consistência e generalização dos resultados entre populações ou grupos clínicos, bem como especificidades e variações de protocolos de tratamento (Sampaio & Mancini, 2007, p. 8).

Contudo, Mancini e Sampaio (2007), destacam que este tipo de método só é possível ser for conduzido após a publicação de vários estudos abordando o tema específico que se quer investigar. Por outro lado, afirma que é importante que os estudos selecionados para a RSL sejam de grande qualidade.

Segundo, as orientações The PRISMA Group (2015, p. 336), a realização de uma RSL é um processo interativo, neste sentido, "[…] revisores sistemáticos podem precisar modificar o protocolo de revisão original no decorrer do trabalho", se caso seja necessário, convém ser relatado e justificado.

Mancini e Sampaio (2007, p. 85) recomendam que a condução da RSL envolva, pelo menos, dois investigadores, "[…] que avaliarão, de forma independente, a qualidade metodológica de cada artigo selecionado". Contudo, devido à falta de conhecimento de um profissional/pesquisador nesta temática abordada, optamos por realizar a recolha de forma independente, respeitando cada etapa do processo, relatando e justificando qualquer mudança ocorrida ao longo da realização da RSL, evitando, desta forma, viés para o estudo.

Como qualquer método, antes de iniciá-lo, é necessário o seu planeamento (Mancini &Sampaio, 2007; The PRISMA Group, 2015; Galvão &Ricarte, 2019). Assim sendo, para este estudo, o planeamento foi dividido em três etapas, conforme se ilustra na figura 6 abaixo. Na etapa 1, estruturação da revisão, são apresentadas a pergunta e os objetivos de investigação, bem como o desenvolvimento do design do protocolo de revisão. Este, por sua vez, para além de dar lugar a identificação do objeto e objetivos da investigação,

incluí os termos a adotar na pesquisa automática, a explicação de como foi conduzido o processo de seleção, quais os critérios de seleção (inclui os critérios de inclusão e exclusão), quais os critérios de avaliação da qualidade dos estudos primários e, por fim, como será efetuada a recolha e síntese dos dados (Anthony et al., 2022). Na etapa 2, execução da revisão, foi implementado o protocolo de revisão, ou seja, a realização da estratégia de pesquisa que permitiu a escolha dos estudos relevantes. Esta incluiu a seleção da base de dados, a definição das palavras-chaves, os critérios de inclusão e exclusão previamente definidos, bem como a avaliação da qualidade dos estudos de forma a refinar o processo de seleção e, por último, foi efetivada a recolha definitiva e a síntese dos dados. Na fase 3, exposição dos resultados, que inclui a análise dos dados, procurando dar respostas a pergunta de partida e aos objetivos da investigação.

Cabe referir que os caminhos percorridos por todas estas etapas são importantes de forma a adequar a pergunta de partida da RSL com base nas informações disponíveis no âmbito do tema de interesse/pesquisado (Mancini &Sampaio, 2007).

A seguir será feita uma descrição das etapas que constituíram o processo de planeamento da RSL deste trabalho. Sublinhamos que, em cada passo, será feita uma ilustração do conteúdo apresentado, tornando acessível a sua compreensão a uma diversidade de leitores.

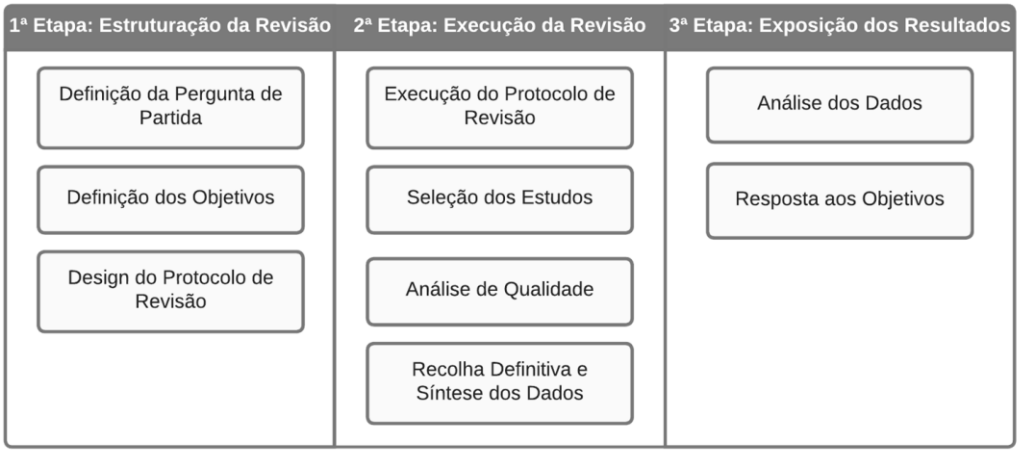

Figura 6- Etapas do planeamento da revisão sistemática da literatura

Fonte: Autor

2.2.1. Estruturação da RSL

A estruturação da revisão é uma fase que deve ser realizada com bastante atenção, uma vez que qualquer erro poderá comprometer o estudo[18]. Assim, ao realizar a estruturação da RSL começamos por definir a pergunta de partida.

Definição da pergunta de partida e dos objetivos

Uma boa RSL depende de uma pergunta ou questão de investigação bem formulada e clara. Esta deve conter: referência a população e intervenção. Para tal, e, seguindo as orientações PRISMA (2015), utilizou-se a técnica PICOS.

Nossa pergunta de partida é: Como a literatura especializada tem tratado a informação e o direito à privacidade no contexto do Big Data?

P: tem tratado a informação e o direito à privacidade no contexto do Big Data;

I: evidências dos estudos (literatura);

C: não há comparação direta neste estudo;

O: quais os estudos que estão citados e qual é o seu nível de qualidade.

Em seguida, traçados alguns objetivos para o estudo:

- Pesquisar nos principais periódicos internacionais estudos que relatem sobre a problemática da informação e a privacidade no contexto do Big Data;
- Identificar os métodos de pesquisa, países, contexto e ano de publicação dos estudos selecionados;
- Identificar sugestões de diretrizes e/ou ações que possam harmonizar interesses pessoais de usuários e interesses económicos no contexto do Big Data;

[18] https://blog.fastformat.co/revisao-da-literatura-o-que-e-como-fazer/#Conduzindo_a_Revisao_da_literatura

- Identificar e relatar as lacunas ainda existentes na literatura no que concerne ao acesso à informação e a privacidade no contexto do Big Data e propor novos direções para os futuros pesquisadores.

Após a formulação da pergunta de partida e a identificação dos objetivos, o próximo passo é a construção do design do protocolo de revisão.

De acordo com Donato & Donato (2019) faz-se necessário, assim como num ensaio clínico, estabelecer um protocolo para as revisões sistemáticas. As autoras sublinham que o protocolo "[...] é um componente essencial no processo de RS e ajuda a garantir a consistência, transparência e a integridade. Este protocolo deverá ser publicado num registo prospetivo na base de dados [...]" (Donato & Donato, 2019, p. 228).

Embora não tenhamos efetuado o registo do nosso protocolo, tivemos como preocupação em validar o mesmo junto aos orientadores antes de sua execução.

D*esign* do protocolo de revisão

Como Donato & Donato (2019) testemunharam, a construção de um design do protocolo de RSL é um passo muito importante para estudos que utilizam este tipo de método, de salientar que este tem a incumbência de definir e limitar o tipo de pesquisa que será útil à aplicação da RSL. A figura 7 ilustra o processo de busca na literatura. Cabe referir que a sua descrição seguiu as recomendações PRISMA (The PRISMA Group, 2015) para o processo de inclusão dos estudos.

Este protocolo é constituído, para além do objetivo geral e os específicos, pela identificação da base de dados, das palavras-chaves, dos critérios de inclusão, exclusão e qualidade.

Por fim, apresentamos o software escolhido para fazermos a sistematização da informação coletada.

Base de dados

Tendo em consideração que os repositórios têm uma ampla divulgação, elevado nível de aceitação e livre acesso à comunidade científica, pretende-se utilizar como fonte de

pesquisa a base de dados da Universidade Nova de Lisboa, a NOVA Discovery[19]. Este repositório científico abrangeum número expressivo de fontes, provendo da EBSCO e, desta forma, aumentar o universo da pesquisa.

Palavras-chaves

Em relação a definição das palavras-chaves será considerada que a informação ocupa um espaço fulcral na pesquisa, para além da informação, efetua intersecção do Big Data e direito à privacidade.

Relativo aos termos da pesquisa, seguiremos as orientações de Donato e Donato (2019), ou seja, pretende-se iniciar a pesquisa utilizando termos com uma linguagem controlada e só depois, iremos utilizara linguagem natural. Assim, utilizando o *thesaurus* da UNESCO (2020) selecionaremos o seguinte termo "*information and privacy in the context of big data"*. Passando para a linguagem não controlada (ou natural) e considerando que a maior parte da literatura relevante se encontra publicada em inglês, os termos da pesquisa a ser utilizados serão: "information and data", "Big Data", "right to privacy".

As palavras-chaves também serão utilizadas em conjunto com operadores booleanos, que serão apresentadas da seguinte forma: "Information" AND "Big Data" AND "Privacy" AND "Rights". A utilização do "AND" admite ao pesquisador combinar os termos da pesquisa de forma que cada resultado da pesquisa contenha todos os termos[20], ou seja, a ideia é que a pesquisa recupere artigos que abrangem as três palavras da pesquisa: "informação, direito à privacidade e Big Data". Por outro lado, como pode ser observado, planeamos fazer uso das aspas para juntar as palavras, contudo e segundo Donato e Donato (2019), esta estratégia é um pouco arriscada, visto que é muito precisa

[19] https://research.ebsco.com

[20] https://connect.ebsco.com/s/article/Pesquisa-com-Operadores-Booleanos?language=en_US.

e, por vezes, o pesquisador pode perder algum artigo relevante. Assim e no sentido de evitar enviesamento para o estudo, paralelamente, será empregue os operadores de proximidade que permitem especificar o número de palavras que podem aparecer entre os 3 termos. Conforme Donato e Donato (2019, p. 231), os operadores de proximidade "[...] recuperam diferentes variações da frase, como a ordem das palavras, o que não é possível com o uso das aspas [...]".

Critérios de inclusão, exclusão e qualidade

Ao realizar uma RSL é importante que o pesquisador defina, de forma explicita, os estudos que serão selecionados e os que irão ser excluídos. Portanto, no sentido de identificar os artigos potencialmente relevantes, pretende-se ter como base os critérios de seleção, exclusão e o critério de qualidade. Relativo aos critérios de seleção, para além da análise das palavras-chaves, teremos como base a leitura dos títulos, dos resumos e, em alguns casos, do artigo na íntegra. Ainda no critério de seleção, temos em vista as publicações que correspondem entre o mês de dezembro de 2012 a dezembro de 2022. Este recuo de dez anos atrás se justifica, uma vez que a diretiva sobre a Proteção de Dados na Aplicação da Lei entrou em vigor 5 de maio de 2016, porém, o prazo máximo para os países da UE transporem para o direito nacional, estendeu-se até o dia 06 de maio de 2018. Ademais, fez-se necessário avançar alguns anos de forma que os Países-Membros pudessem adequarem-se a nova legislação de protecção de dados.

Ainda, no critério de seleção, só serão considerados os artigos escritos na língua portuguesa (Brasil e Portugal) e inglesa. Para mais, apenas serão aceitos estudos que passaram por revisões de pares (*peer review*).

Sublinhamos que os critérios de pesquisa foram atualizados conforme a execução dos primeiros testes para chegar à base final de definição, a qual está descrita abaixo. Este pode ser visualizado na figura 7.

Protocolo de Revisão

Identificação do objetivo da revisão	Analisar e sintetizar a literatura existente sobrea informação e o direito à privacidade no contexto do Big Data.
Delimitação dos objetivos de investigação	I. Pesquisar nos principais periódicos internacionais estudos que relatem sobre a problemática da informação e a privacidade no contexto do Big Data; II. Identificar os métodos de pesquisa, países, contexto e ano de publicação dos estudos selecionados; III. Identificar sugestões de diretrizes e/ou ações que possam harmonizar interesses pessoais de usuários e interesses econômicos no contexto do Big Data; IV. Identificar e relatar as lacunas ainda existente na literatura no que concerne o acesso à informação e a privacidade no contexto do Big Data e propor novos direcionamentos para os futuros pesquisadores.
Seleção das bases de dados	NOVA Discovery (EBSCO)
Palavras-chaves	Informação; Big Data; Direito aPrivacidade; (Levantar possíveis sinônimos e termos em língua inglesa, fazer uso de operadores booleanos, tais como AND (e), OR (ou), AND NOT (e não))
Critérios de inclusão, exclusão e qualidade (informações detalhadas nas figuras8 e 9)	Serão analisados os seguintes aspetos: título, resumo, palavras-chaves, conteúdo, garantia de qualidade, idioma e data. Relativamente ao conteúdo só serão incluídos os estudos que explicitamente relacione a informação e o direito à privacidade no contexto do Big Data. Quanto ao idioma, só serão elegíveis artigos escritos na língua inglesa e portuguesa. Finalmente, no que diz respeito à data, serão considerados os estudos publicados de dezembro de 2012 a dezembro de 2022. Para a garantia da qualidade dos artigos, apenas estudos oriundos de revisão por pares serão considerados; estudos publicados em periódicos/revistas internacionais; estudos com 10 ou mais referenciações; estudos que respondam os objetivos definidos.
Sistematização da informação coletada	Pretende-se incluir os documentos encontrados na base de dados no software “StArt” (State of the Art through Systematic Review) para a gestão de referências bibliográficas, facilitando, desta forma, a gestão dos documentos revistos em cada etapa do processo, assim como o suporte de software de planilhas para extração e análise dos dados

Figura 7- Etapas do planeamento da revisão sistemática da literatura.

Fonte: Autor

A figura 8 apresenta, em detalhe, os critérios de inclusão que sustentam a seleção dos documentos potencialmente relevantes e indica também, os critérios de exclusão sob os quais os artigos serão rejeitados, não sendo mais considerados para efeito do presente estudo.

Critérios de Inclusão (CI)	Sim	Não
CI1- O estudo possui como tema principal a informação e o direito a privacidade no contexto do Big Data		
CI 2- O estudo é escrito em inglês e/ou português		
CI 3- O estudo foi publicado entre dezembro de 2012 e dezembro de 2022		
Critérios de Exclusão (CE)	**Sim**	**Não**
CE 1- O estudo não possui como tema principal a informação e o direito a privacidade no contexto do Big Data		
CE 2- O estudo não aborda legislações existentes que respondem ao direito a privacidade		
CE 3- O estudo recorre a métodos de pesquisa empírica		
CE 4-O estudo é uma publicação do género: ensaio, editorial, material de conferência, livro, relatório, dissertações e teses		
CE 5- O estudo não foi submetido a revisão por pares		

Figura 8- Critérios de elegibilidade.

Fonte: Autor

Critério de qualidade

Em complementaridade aos critérios de seleção, o critério de qualidade é imprescindível para o processo de investigação. De acordo com Mancini e Sampaio (2007, p. 87), a "[…] qualidade de uma revisão sistemática depende da validade dos estudos incluídos nela."

Na perspetiva de Donato e Donato (2019), o pesquisador, ao partir para a leitura completa dos artigos identificadosvpara inclusão na revisão como parte do processo de extração de dados, deve aplicar uma escala de avaliação da qualidade de cada estudo

selecionado. Há várias ferramentas para auxiliar nesta tarefa e, a escolha desta, vai depender da pergunta de partida da revisão (Donato& Donato, 2019). Posto isto, estabelecemos uma *checklist* dos elementos necessários para um estudo de qualidade tendo por base as recomendações PRISMA. Os princípios aos quais obedecem e os critérios de qualidade a serem utilizados no presente protocolo de revisão encontram-se indicados na figura 9.

Ao atender todos os critérios de qualidade um estudo primário pode alcançar uma pontuação máxima de 6 pontos. Isto equivale a receber 1 ponto (indicado como "Sim") para cada critério que seja completamente satisfeito. Nos casos em que um critério não é completamente cumprido, mas ainda é parcialmente atendido, uma pontuação de 0,5 (referente a "Parcialmente") será concedida. Se um critério não for cumprido, total ou parcialmente, a pontuação atribuída será 0 (indicado como "Não"). Qualquer omissão de informação também resultará numa pontuação de 0. Portanto, é fundamental avaliar rigorosamente cada critério de qualidade conforme essas diretrizes para determinar a pontuação precisa de um estudo primário. A pontuação mínima definida para o estudo ser aceite e passar para a fase final é de 4,5 ou mais, ou seja, 75% do valor total da análise de qualidade. Desta forma garantimos que somente os artigos que atingiram uma pontuação considerável com base nos critérios de qualidade, serão selecionados.

Critérios de Qualidade (CQ)	Verificação	Pontuação
CQ1- O estudo é referenciado em outros estudos, contendo pelo menos 10 ou mais referenciações	Não/Parcialmente/Sim	0-1
CQ 2- Os objetivos do estudo estão claramente identificados	Não/Parcialmente/Sim	0-1
CQ 3- Os procedimentos metodológicos são suficientes para a análise	Não/Parcialmente/Sim	0-1
CQ 4- O estudo apresenta os resultados e implicações para futuras pesquisas	Não/Parcialmente/Sim	0-1
CQ 5- O estudo identifica sugestões de diretrizes e/ou ações que possam harmonizar interesses pessoais de usuários e interesses econômicos no contexto do Big Data	Não/Parcialmente/Sim	0-1

CQ 6- O estudo identifica e relata as lacunas ainda existente na literatura no que concerne o acesso à informação e a privacidade no contexto do Big Data	Não/Parcialmente/Sim	0-1

Figura 9- Critérios de qualidade.

Fonte: Autor

Sistematização da informação coletada

Pretendem-se incluir os artigos encontrados na base de dados no software "StArt"[21] (State of the Art through Systematic Review). Este software foi desenvolvido pelo Laboratório de Pesquisa em Engenharia de Software (LaPES) da Universidade Federal de São Carlos para a gestão de referências bibliográficas[22], facilitando, desta forma, a gestão dos documentos revistos em cada etapa do processo. Adicionalmente, irá recorrer-se a utilização de softwares de planilhas (Excel) em conjunto ao software supracitado, para facilitar a extração, tratamento e análise dos dados.

Além da tabela da análise de qualidade, irá-se desenvolver uma tabela para a recolha dos dados dos estudos que passarem para a seleção final. O objetivo é coletar as informações essenciais para se poder analisar os estudos com base nos objetivos definidos e integrar os mesmos a meta-síntese.

A recolha de dados será apresentada conforme categorias (ou temas) que incidiam nos objetivos da RSL, a saber: autor, ano, países de publicação e/ou países estudados, contexto (que tipo de informação é tratada), método de pesquisa, resultados e implicações para futuras pesquisas, sugestões de diretrizes e lacunas existentes.

[21] StArt — Laboratório de Pesquisa em Engenharia de Software (ufscar.br)

[22] http://fernandoscpimentel.blogspot.com/2022/05/revisao-sistematica-da-literatura-o.html

2.3. Execução da revisão sistemática da literatura

No decurso da realização da revisão sistemática da literatura, o protocolo previamente estipulado será aplicado com o intuito de assegurar a completa documentação de todo o procedimento. Isso, por conseguinte, tem como finalidade garantir e evidenciar o patamar elevado de precisão e integridade subjacente a essa metodologia, enquanto possibilita a replicação deste processo em oportunidades futuras (Donato & Donato, 2019).

No que diz respeito a RSL adotada neste estudo, foram utilizados processos manuais de recolha e tratamento de dados, juntamente com ferramentas especificas para automatizar cada etapa da pesquisa. A ferramenta principal utilizada para a revisão sistemática da literatura foi o software “StArt” (State of the Art through Systematic Review), assim como softwares de planilhas para o apoio em todo o processo da RSL.

Esta etapa exige vários passos importantes de forma a conduzir uma boa RSL, a saber: execução do protocolo de RSL – é uma etapa muito importante neste método, pois, é, exatamente, nesta etapa, que identificamos os primeiros estudos que serão incluídos na revisão – seleção dos estudos, a análise de qualidade, recolha definitiva e a síntese dos dados recolhidos.

2.3.1 Identificação e seleção

Com base no protocolo de revisão, em primeiro lugar foi realizada uma pesquisa simples na base de dados definida, “NOVA Discovery”. Nesta primeira pesquisa, apenas, fornecemos para a base de dados, as palavras-chaves/termos de pesquisa e os booleanos estabelecidos. Assim, o seguinte modelo de pesquisa foi aplicado: "Information" AND "Big Data" AND "Privacy" AND "Rights". O objetivo era obter um primeiro resultado de teste para se ter uma visualização da quantidade de estudos que estariam disponíveis pela utilização de uma busca base. Esta pesquisa inicial foi realizada no dia 24 de abril de 2023, tendo como resultado 582 estudos.

Após isso, recorreu-se a utilização das ferramentas de busca avançada da base de dados de forma a realizar a pesquisa utilizando filtros específicos. Nesta etapa, para além de

ter sido usado novamente o modelo de pesquisa aplicada inicialmente, fornecemos novos filtros, a saber: período dos estudos que interessava para nossa busca, ou seja, entre dezembro de 2012 a dezembro de 2022. Ainda foi inserido o tipo de publicação no que se refere a sua qualidade, desta forma, para a nossa pesquisa, importava-nos apenas as publicações revistas por pares e, que estas publicações teriam de estar escritas na língua portuguesa e/ou inglesa. Ademais, foi fornecido algum critério de exclusão para o tipo de fonte, excluindo assim, editoriais, material de conferência, livro, e-book e relatório. Esta estratégia utilizada para a pesquisa avançada pode ser visualizada na figura 10. Esta segunda pesquisa, com a utilização de filtros avançados, foi realizada no mesmo dia da pesquisa anterior, resultando, desta forma, um total de 275 estudos.

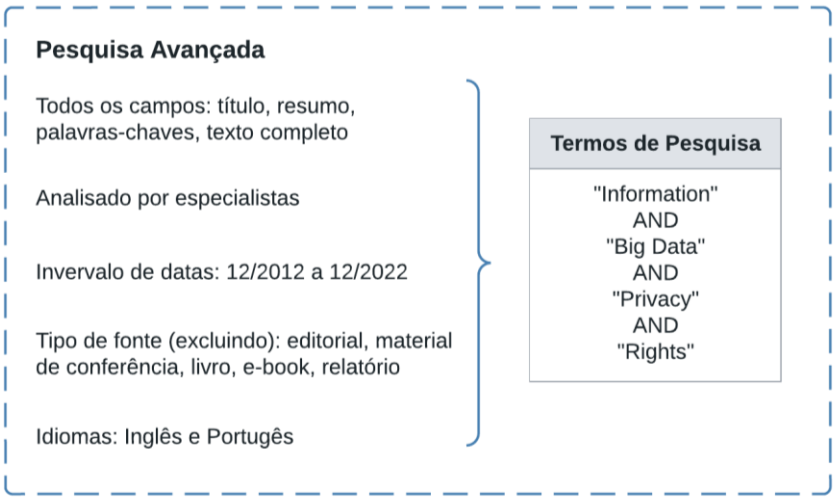

Figura 10- Estratégia de pesquisa.

Fonte: Autor

Concluída a identificação dos estudos, procedeu-se a exportação deles para o formato "bibtex", a fim de facilitar a migração dos dados para o software StArt. Contudo, verificou-se que dos 275 estudos exportados, 26 não foram identificados corretamente no software, devido ao conflito de modelos de referências dos artigos. Desta forma, foi preciso recorrer a introdução manual destes estudos no software "Mendeley" para poder registar as referências e conseguir transformar novamente no formato bibtex.

Após a limpeza e transformação dos dados acima, foram inseridos novamente os estudos e definido o mapeamento da revisão sistemática dentro do software StArt.

Após a importação dos dados para o software supracitado, passou-se logo para a identificação dos estudos duplicados. O software reconheceu automaticamente 34 estudos duplicados, entrando assim para a fase de seleção 241 estudos.

Podemos visualizar a distribuição dos 241 estudos por anos de publicação na figura 11, para uma primeira noção geral da amostra antes da fase de seleção. Apesar da busca ter incidido entre os períodos de dezembro 2012 a dezembro 2022, não foi encontrado na base de dados publicações relacionadas ao final do ano de 2012, começando somente em 2013 e, mostrando um considerável crescimento de estudos nos anos seguintes.

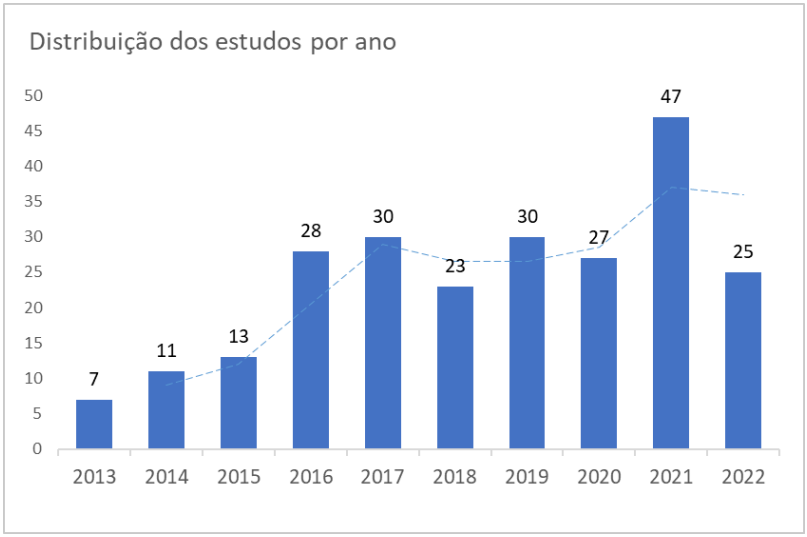

Figura 11- Gráfico de distribuição dos estudos iniciais por ano de publicação.

Fonte: Autor

2.3.2 Elegibilidade

Análise de inclusão e exclusão

Nesta etapa foram aplicados os critérios de inclusão e exclusão para selecionar os estudos que seriam mais relevantes, seguindo o protocolo definido para a RSL. Esses critérios incidiram sobre a leitura do título, resumo e das palavras-chaves dos estudos.

Finalizados os procedimentos aplicados para a seleção dos estudos para a RSL, restaram 82 artigos em texto completo para a próxima etapa, como podemos ver na figura 12. Foram selecionados 74 estudos aceites para a próxima fase e 8 estudos adicionais que ficaram por ser classificados (aceite ou rejeitado), por terem gerado dúvidas quanto a sua inclusão após a realização da leitura do título, resumo e palavras-chaves.

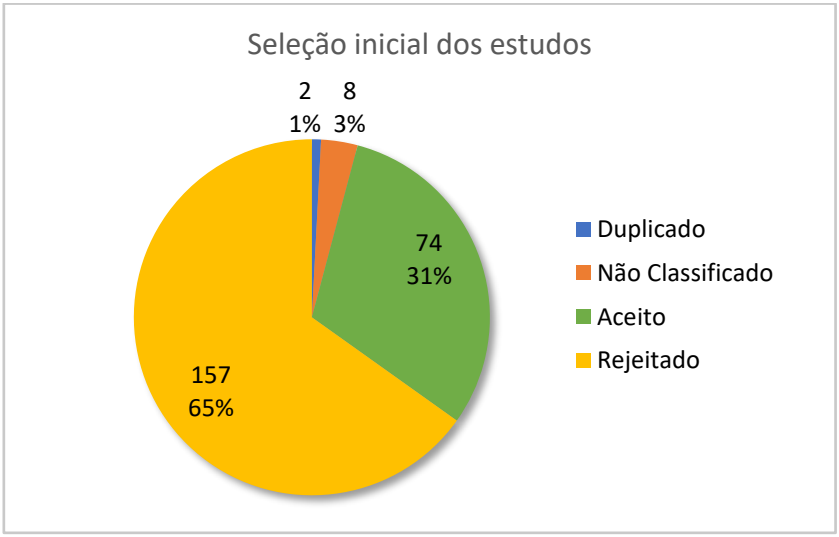

Figura 12- Gráfico da seleção inicial dos estudos.

Fonte: Autor

Após a leitura completa dos 8 estudos não classificados, 6 foram aceites e 2 rejeitados. Passando 80 estudos para a análise de qualidade.

Análise de qualidade

A avaliação de qualidade dos estudos corresponde a uma análise mais detalhada na aplicação dos critérios de seleção, permitindo, desta forma, aferir a relevância dos estudos primários na fase da síntese dos resultados (Kitchenham & Charters, 2007).

O processo de avaliação da qualidade dos estudos, como já foi descrito acima, foi realizado por meio de uma checklist. Esta tinha como baseas orientações Prisma (The PRISMA Group, 2015), sendo adaptada para a avaliação de pesquisa qualitativa. O Instrumento foi composto por seis questões, onde na primeira rastreamos a qualidade científica dos artigos e na segunda, terceira e quarta abordamos o desenho de estudo: os objetivos, os procedimentos metodológicos, os resultados e as implicações dos estudos qualitativos. Já as duas últimas relacionam-se diretamente com os objetivos da RSL. Sublinhamos que o pesquisador realizou leitura completa dos 80 estudos de forma a determinar a elegibilidade final dos artigos. Após esta etapa, excluíram-se 52 artigos. Dos artigos excluídos, 23 foram rejeitados por não atender aos critérios de seleção após a leitura completa, 26 por não atingirem a pontuação necessária dos critérios de qualidade e 1 artigo adicional, por se tratar de uma publicação retratada. As especificações de artigos retratados não foram previstas durante o processo de definição dos critérios de seleção, porém, decidiu-se excluir o mesmo nesta etapa por poder prejudicar a validade dos resultados. O resultado da seleção final dos estudos rejeitados e aceitos podem ser consultados respetivamente no Apêndice A e Apêndice B, incluindo a informação da avaliação de qualidade dos mesmos e justificativa de rejeição.

As informações detalhadas da rejeição foram as seguintes (ver figura 14):

- CE 1) O estudo não possuía como tema principal a informação e o direito a privacidade no contexto do Big Data (N=6).
- CE 2) O artigo não abordava a legislaçõe existente que responde ao direito à privacidade (N=1).
- CE 3) O artigo recorreu ao método de pesquisa empírica (N=11).
- CE 4) O estudo era uma publicação do género: ensaio, editorial, material de conferência, livro, relatório, dissertações e teses (N=7).
- Critério de qualidade) O artigo não atendeu a pontuação de qualidade mínima exigida (N=26).
- Artigo retratado) O artigo apresentou-se como uma publicação retratada (N=1).

Feito isto, resultaram 28 artigos, compondo, assim, nossa amostra final.

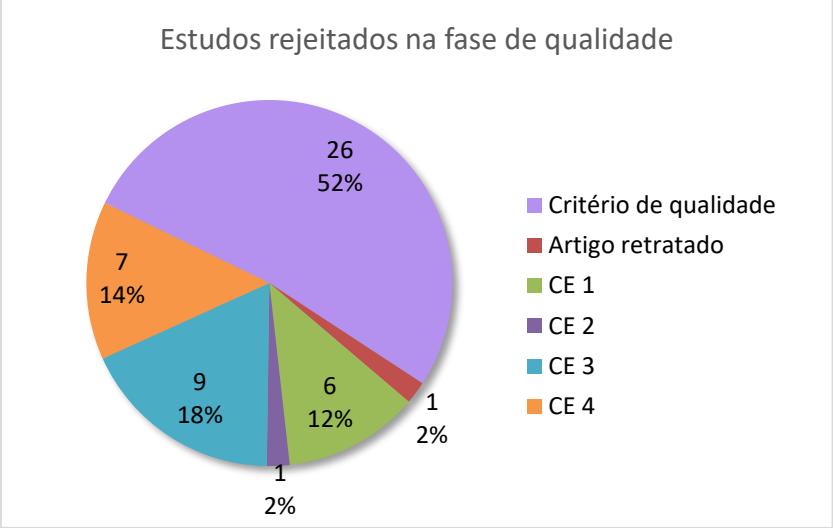

Figura 13- Gráfico de estudos rejeitados após leitura completa.

Fonte: Autor

2.3.3 Inclusão definitiva

Para recapitular, o processo da Revisão Sistemática da Literatura (RSL) foi conduzido em etapas distintas. Inicialmente, ocorreu a identificação dos estudos pertinentes no âmbito da pesquisa. Posteriormente, uma seleção preliminar foi realizada, envolvendo a coleta dos estudos presentes na base de dados e a subsequente exclusão dos duplicados. Prosseguindo, efetuou-se a análise de elegibilidade dos estudos, a qual compreendeu tanto uma avaliação preliminar, conforme os critérios de inclusão e exclusão, baseada nos títulos, resumos e palavras-chave, quanto uma análise completa dos estudos selecionados a fim de aferir sua qualidade e pertinência para a RSL.

Chegamos, então, à fase culminante do processo, caracterizada pela inclusão definitiva dos estudos. Nesta etapa, foram criteriosamente escolhidos 28 estudos que atenderam aos requisitos estabelecidos e demonstraram contribuições substanciais para a investigação em questão. Baseado no fluxograma PRISMA (2015), um registo detalhado das etapas percorridas, juntamente com a enumeração precisa dos estudos, pode ser

consultado na Figura 14, proporcionando uma visão abrangente e estruturada do desenvolvimento da RSL.

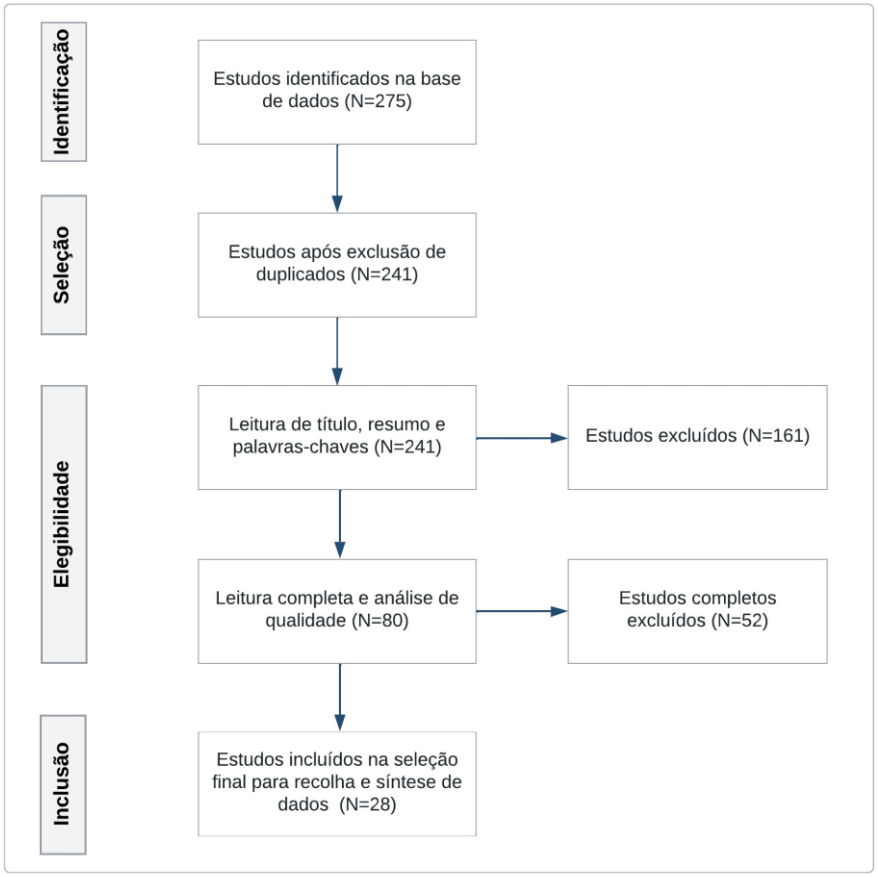

Figura 14- Fluxograma PRISMA com as fases da revisão sistemática da literatura.

Fonte: Autor, adaptado PRISMA (2015)

Os 28 estudos elegíveis passaram, primeiramente, por uma leitura superficial de modo a recortar o conteúdo destes em fragmentos que traduziam uma ideia particular relacionada com a teoria levantada neste estudo. No passo seguinte ao recorte de conteúdo, realizou-se uma leitura aprofundada de todos os estudos, buscando agrupar os elementos de conteúdo por parentesco de sentido às categorias analíticas. As categorias, ou seja, os temas abordados, foram alinhadas com os objetivos preestabelecidos da RSL. Esses objetivos abarcaram aspetos como identificação do autor e ano de publicação, países de origem da pesquisa e/ou países centrais do estudo,

contexto temático (abrangendo o tipo de informação abordada), metodologia empregada, resultados obtidos e implicações que refletirão em futuras investigações. Além disso, considerou-se também a formulação de diretrizes sugeridas e as eventuais lacunas. Assim se estabeleceu uma tabela de recolha dos dados que poderá ser consultada adiante, compilando todos esses pontos para a análise final.

3. Apresentação e análise dos resultados

A informação completa da recolha de dados pode ser encontrada na figura 19, onde se construiu uma tabela para fazer esta recolha e posterior análise.

Anos de publicação dos estudos

Ao concluir a recolha dos dados, os resultados apontam que os anos de 2017 e 2018 foram os anos de maiores publicações acerca da temática estudada, ambos com seis (n=6) publicações, seguido dos anos de 2016, 2020 e 2021, todos estes três com quatro (n=4) publicações. Os anos com menores números de publicações foram 2013, 2014, 2015 e 2022, todos com uma (n=1) publicação. Os anos de 2012 e 2019 não foram registados nenhuma publicação. Na figura 15 pode-se visualizar a distribuição dos anos de publicações dos estudos

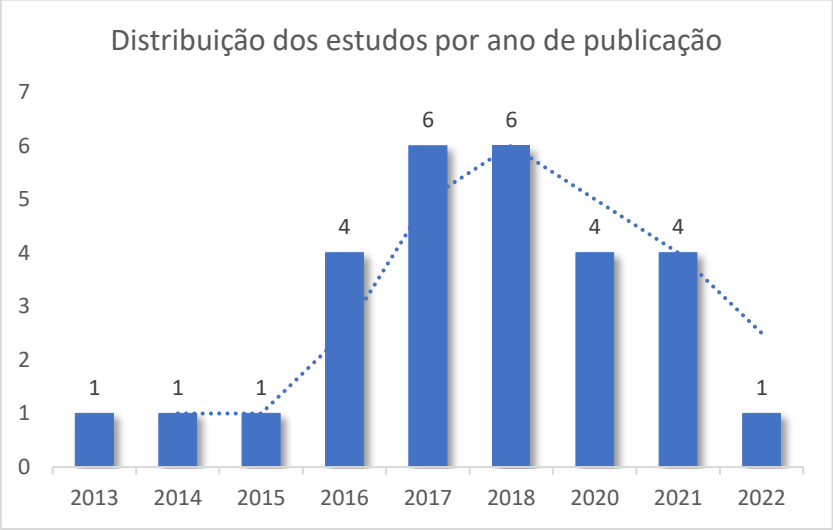

Figura 15- Gráfico de distribuição dos estudos por ano de publicação.

Fonte: Autor

Regiões e/ ou países de publicação dos estudos

De acordo com a figura 16, os vinte oito artigos incluíram um total de seis regiões/países, Canadá (n=1), Rússia (n=1), UAE - Emirados Árabes Unidos (n=1), Ásia (n=2), EUA – Estados Unidos da América (n=6) e UE - União Europeia (n=13). Relativo aos estudos com foco dentro dos países e/ou legislações da UE foram todos agrupados num mesmo bloco, devido as regulamentações que abrangem esses países. Observámos que alguns estudos tiveram o foco em mais de uma região, no caso os EUA e a UE em conjunto, resultando um total de quatro (n=4) estudos.

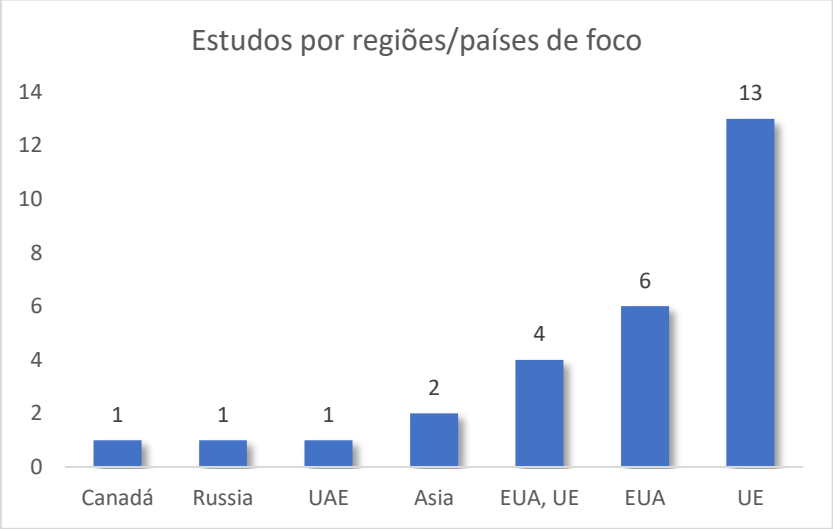

Figura 16- Gráfico de distribuição de países de foco.

Fonte: Autor

Método utilizado nos estudos

Conforme a figura 17, os resultados apontam que todos os estudos usaram algum tipo de revisão qualitativa, dentre estas, encontramos a revisão narrativa (n=1), revisão sistemática da literatura (n=1), revisão documental (n=5), revisão da literatura em conjunto com a revisão documental (n=7) e revisão da literatura (n=14).

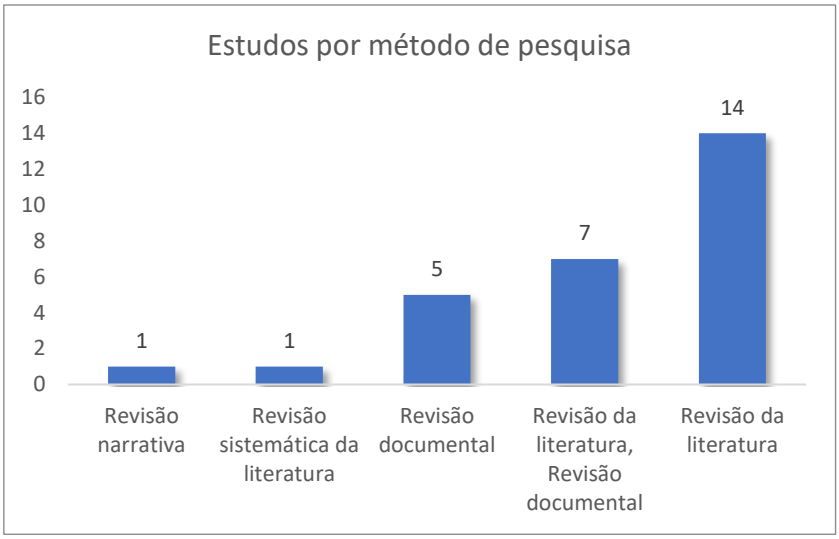

Figura 17- Gráfico de distribuição de método de pesquisa.

Fonte: Autor

3.1. Contexto dos estudos

Relativo ao contexto, os resultados apontam que os estudos abordaram seis temáticas diferentes. A temática mais focalizada foi a regulamentação, apresentando dezasseis (n=16) artigos, seguida da saúde com cinco (n=5) estudos. Os contextos que apresentaram menor enfoque foram o direito da criança (n=2), educação (n=2), ética (n=2) e política criminal (n=1). Na figura 18 pode-se visualizar a distribuição dos estudos por temáticas/setor de foco.

Como podemos observar na figura 18, a maioria dos estudos teve como fulcro a regulamentação artigos 2, 6, 7, 12, 13, 14, 15, 17, 18, 20, 21, 22, 24, 26, 27 e 28. Dentro desta temática encontramos diversos eixos de estudo, tais como: regras de privacidade no contexto dos EUA artigo 2 , perfilamento de dados corporativos e os chilling effects na área de comunicações eletrônicas e sobre o regulamento ePrivacy artigo 6, responsabilidade social dos ativos digitais após a morte artigo 7 , privacidade de dados na UE artigo 12, 24, 25 , reutilização de big data na lei da UE artigo 13 , privacidade e jurisprudência do Tribunal Europeu dos Direitos Humanos (TEDH) artigo 14, 15, legislações sobre a privacidade e os desafios do big data artigo 17 , aplicações viáveis do conceito tradicional ao novo fenómeno

do big data [artigo 18] , atual status de privacidade da informação nos Emirados Árabes Unidos [artigo 20] , big data nos direitos constitucionais dos cidadãos [artigo 21] , desafios jurídicos da utilização da análise de big data nas mídias sociais [artigo 22] , responsabilidade de grupo no contexto do big data [artigo 27] e proteção de dados no contexto de processos de tomada de decisões [artigo 28].

Uma pequena parcela dos estudos abordou o tema da saúde e privacidade artigos 1, 3, 5, 8 e 19. De entre esta temática foram encontradas abordagens voltadas para a utilização do big data e das tecnologias digitais para facilitar medidas preventivas de doenças artigo 1 , implementação de tecnologias médicas que dependem de transferência de dados transfronteiriços artigo 3 , direito à privacidade e o direito a proteção dos dados no contexto da saúde artigo 5 , direitos humanos do cuidado a saúde em pessoas com deficiência artigo 8 , dados de participantes na utilização de pesquisa científica artigo 19 .

Observa-se que poucos estudos tiveram como foco a ética, artigos 11 e 16, focando nos seguintes pontos: Clonagem digital artigo 11 e big data e direitos humanos na investigação científica artigo 16.

Igualmente a temática da ética, ocorrerá com a educação artigos 9 e 10, que versaram sobre: uso de big data para melhorias na educação e tensões de privacidade artigo 9, pesquisa de análise de aprendizagem e a privacidade individual artigo 10.

Os mesmos resultados anteriores, também podem ser visualizados sobre a temática do direito da criança artigos 4 e 23, que discorre sobre: privacidade em relação a publicidade em mídias sociais, aplicativos móveis e jogos artigo 4 e direito e proteção de dados pessoais de crianças na lei RGPD artigo 23.

Por fim, a temática política criminal artigo 21, que foca na proteção dos direitos humanos fundamentais contra invasões terroristas, tendo apenas um artigo.

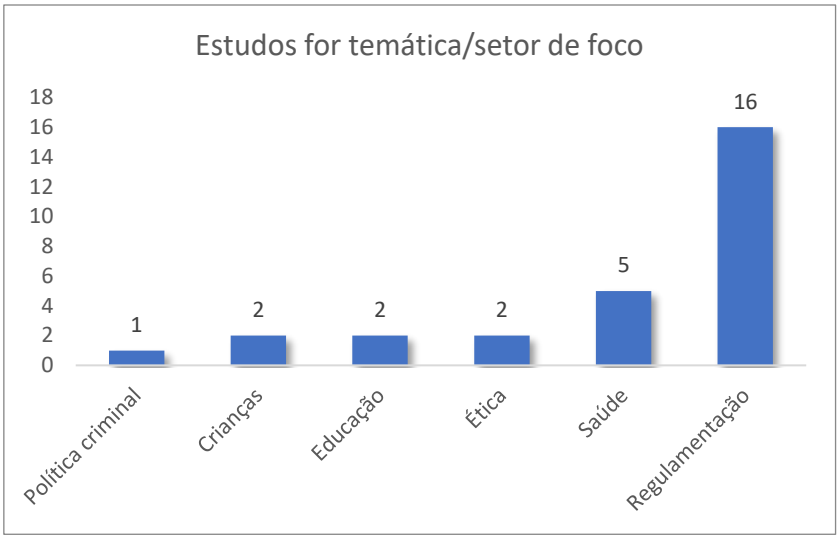

Figura 18- Gráfico de distribuição de temática/setor de foco.

Fonte: Autor

3.1.2. Diretrizes e/ou ações e lacunas existentes na literatura

Os dados sobre as diretrizes e/ou ações que possam harmonizar interesses pessoais de utilizadores e interesse económico, bem como as lacunas existentes na literatura no que concerne ao acesso à informação e a privacidade no contexto do big data foram explorados e analisados conforme o contexto/tema de cada estudo. Os resultados e análise destes serão apresentados de seguida por meio de uma síntese do conhecimento.

3.1.3. Regulamentação

No que concerne a temática da regulamentação Strahilevitz (2013) ressalta os desafios da era do big data e como a indústria e o governo aplicam regras de privacidade no contexto dos EUA, nomeadamente sobre os testes de personalidade refinados aos consumidores. Na perspetiva desse autor, a discriminação de personalidade em preços e serviços de mercado em massa para prever as reações legais, faz-se necessário, primeiramente, compreender seus beneficiários e lesados. Assim sendo, o autor sugere o desenvolvimento de formas mais amplas de impacto na proteção de dados: reconsiderar a natureza da avaliação social − uma vez que se torna excessivamente

onerosa e complexa para as empresas, recomendam um modelo voluntário, que mantém liberdade de decisão dos controladores de dados, tornando esta avaliação uma solução mais aceitável do que as disposições obrigatórias. Para este autor, uma abordagem voluntária é mais consistente com o quadro jurídico existente, que parece ter dificuldades em ir além da mera proteção de dados no uso da informação. Neste sentido, o RGPD – que fornece um dos exemplos mais avançados de regulação nesta área – centra-se no risco.

Para Papakonstantinou & de Hert, (2020) no que se refere ao perfilamento de dados corporativos e os "chilling effects" na área de comunicações eletrónicas e sobre o regulamento ePrivacy, há dois principais problemas: necessidade social de ampliar as operações de análise de big data e diferenças entre abordagens de análise de big data entre o RGPD e as leis de ePrivacy. Ainda, para estes autores (Papakonstantinou & De Hert, 2020), o projeto de Regulamento de ePrivacy não atende às expectativas legislativas pelo fato de não reconhecer a necessidade de regular a análise de big data e por não fornecer especificidade na diferenciação entre operadores de telecomunicações e empresas de internet. Sendo assim, os autores sugerem que o atual projeto de Regulamento de ePrivacy seja retirado e revisado de forma abrangente, especialmente em relação à proteção de dados. Enquanto isso, o RGPD poderia servir como referência para o campo de comunicações eletrônicas, combinado com a Diretiva de ePrivacy conforme recomendado pelo Código de Comunicações Eletrônicas. Assim, estas abordagens proporcionariam aos legisladores da UE tempo para desenvolver um novo texto de ePrivacy mais detalhado e adequado às circunstâncias sociais e de mercado em linha com os objetivos de proteção de dados.

Park et al. (2020) ao analisar sobre a responsabilidade social dos ativos digitais após a morte salientam a responsabilidade social de definir direitos de controle sobre esses ativos, abordando tanto a privacidade quanto os direitos de propriedade. Os autores acrescentam que é fundamental codificar de maneira explícita o princípio ético do tratamento dos ativos digitais após a morte com respeito, estabelecendo um amplo quadro regulatório para limitar a exploração comercial. Para o autor, isto requer a

criação de normas e requisitos padronizados, com poder regulatório, em harmonia com os esforços. Além disso, deve haver uma cláusula de proteção que preserve os direitos de privacidade e propriedade dos ativos digitais após a morte, proibindo usos comerciais não autorizados e alterações. O objetivo ético é garantir que esses ativos sejam tratados com o mesmo respeito que durante a vida da pessoa, preservando a integridade de seus propósitos originais.

Tzanou (2017) ao analisar os desafios da privacidade de dado na UE, chega à conclusão de que o julgamento Schrems é um marco constitucional na proteção judicial dos direitos fundamentais, pois reforça o direito à privacidade contra a vigilância eletrônica moderna e destaca que o acesso a dados pessoais por autoridades públicas não se justifica com base na mera disponibilidade desses dados por empresas privadas e, também, amplia a interpretação das leis de proteção de dados da UE para transferências de dados transfronteiriços e reafirma a importância da ativação da privacidade. Contudo, no ponto de vista deste autor, a análise dos direitos fundamentais pelo tribunal ainda apresenta lacunas e omissões. O recém-adotado Privacy Shield, que substitui o sistema Safe Harbour, não aborda de maneira satisfatória as preocupações com os direitos fundamentais levantadas pelo Tribunal de Justiça da União Europeia (TJUE). Suas salvaguardas para direitos fundamentais são limitadas e mistura o regime de transferência de dados transatlânticos com a regulamentação de operações de contraterrorismo. Desta forma, o autor sugere a criação de um Mecanismo de Ombudsman, contudo, não garante total reparação para indivíduos.

Em relação as barreiras e facilitadores legais para a reutilização de big data na lei da UE, Ursic & Custers (2016) afirmam que lei de proteção de dados da UE atua principalmente como uma barreira, limitando o processamento de dados e impondo obrigações adicionais aos reutilizadores de dados. No ponto de vista destes autores, os direitos humanos também podem ser uma barreira, pois questionam a reutilização de dados que não considera os direitos e interesses dos titulares. Já, as leis de retenção de dados e segurança cibernética, podem ser tanto barreiras quanto facilitadores, dependendo das exigências impostas. Da mesma forma, as leis de propriedade intelectual e concorrência

podem atuar como barreiras ou facilitadores para a reutilização de dados, dependendo de como afetam a exploração de conjuntos de dados e a concorrência justa. A lei de proteção do consumidor pode bloquear práticas comerciais injustas ou promover publicidade mais precisa e direcionada, dependendo do contexto da reutilização destes dados. Os autores realçam que o cenário regulatório da UE é altamente complexo quando se trata de reutilização de dados, assim sendo, identificar as barreiras e facilitadores mais importantes para a reutilização de dados pode ser útil para regular ainda mais a reutilização de dados de forma a facilitar um ambiente digital sustentável e dinâmico.

No que concerne às violações de privacidade e jurisprudência do Tribunal Europeu dos Direitos Humanos (TEDH), van der Sloot (2015) realça que, em contraste com outros direitos qualificados protegidos pelo TEDH, o direito à privacidade é frequentemente limitado por objetivos sociais, como segurança, proteção de valores sociais e bem-estar económico do país. O tribunal parece adotar diferentes teorias, como a teoria unitária para segurança, a teoria da preponderância para valores sociais e a teoria do interesse geral para questões econóômicas. Assim sendo, o interesse privado e o interesse comum são avaliados de acordo com essas teorias em diferentes contextos e, por consequência, contrariando a Convenção Europeia dos Direitos Humanos, que enfatiza que a privacidade é um direito voltado principalmente para o interesse privado. Para este autor, é incerto se as reivindicações sobre o Big Data serão declaradas admissíveis sob a Convenção Europeia dos Direitos Humanos. Além disso, mesmo se forem consideradas admissíveis, não está claro como o TEDH poderia abordar esses dilemas de maneira satisfatória. Os testes existentes de interesse público, interesse individual e equilíbrio de interesses não se aplicam bem ao Big Data devido à complexidade dos interesses em jogo. Portanto, um quarto teste precisa ser desenvolvido, considerando desenvolvimentos estruturais e sociais, para que o TEDH possa abordar eficazmente questões de privacidade relacionadas ao Big Data e manter a relevância do Artigo 8 da ECHR.

O mesmo autor (Van de Sloot), em um outro estudo realizado em 2018 reafirma que o direito à privacidade está vinculado aos interesses e direitos individuais, assim sendo, há a necessidade de demonstração de interferência específica e dano real para ser considerado admissível sob a Convenção Europeia dos Direitos Humanos. No entendimento de van der Sloot (2018), o surgimento de tecnologias modernas dificulta a identificação de violações concretas e danos pessoais, visto que o TEDH avalia alegações em abstrato, analisando a qualidade das leis em casos de vigilância em massa. Para minimizar este problema, o autor propõe uma abordagem influenciada pela discussão filosófica sobre liberdade, com a atenção voltando para o princípio de 'não-dominação', ou seja, não fazer uso abusivo do poder e controle. Em casos de violações tradicionais de privacidade, o princípio de não interferência é aplicado, restringindo reclamações apenas aos diretamente afetados. No entanto, em situações envolvendo novas tecnologias de dados, como vigilância em massa por serviços de inteligência, o TEDH adota o princípio de não-dominação, permitindo ações mesmo sem danos individuais claros. Esta abordagem amplia a proteção legal e impõe limites ao poder governamental, mas é criticada por desviar-se dos procedimentos legais nacionais e por conceder ao TEDH um poder não estabelecido na Convenção Europeia dos Direitos Humanos, possivelmente afetando a legitimidade democrática e o ideal do Estado de Direito.

Crawford & Schultz (2014), ao analisar as ferramentas analíticas do Big Data, concluem que os problemas de privacidade dos algoritmos preditivos muitas vezes são imprevisíveis por si só, e seus efeitos podem nem mesmo ser totalmente compreendidos por seus programadores. Em muitos contextos, é impossível garantir a privacidade diferencial ao usar um algoritmo de aprendizado que retira dados de uma distribuição contínua, tornando difícil prever quando um algoritmo de aprendizado preditivo revelará informações de identificação pessoal sobre um indivíduo. Neste sentido, torna-se desafiador definir proteções de privacidade em torno desses dados. No ponto de vista desses autores, o uso do Big Data para além de contornar regulamentos antidiscriminatório existentes, pode levar a violações de privacidade na área da saúde e

da aplicação da lei. Para evitar tal situação, os autores propõem um processo de dados procedimentais em vez de tentar regular a coleta, uso ou divulgação de dados pessoais. Isso permitiria aos indivíduos exercerem seus direitos de devido processo de dados em relação a essas determinações.

Wang (2017) afirma que, sem dúvida, o big data transformou a vida das pessoas na sua maneira de viver, trabalhar e pensar. Por outro lado, também tem aumentado a eficiência da produtividade, da segurança, da conveniência, das oportunidades e lucros nos negócios e na vida diária. Contudo, o fenómeno do big data ainda enfrenta incertezas legais, já que não existe uma legislação única que aborde especificamente questões legais do big data. Posto isto, o autor propõe que os conceitos e princípios legais tradicionais do direito de banco de dados, do direito de proteção de dados, do direito de propriedade intelectual, do direito contratual e do direito internacional privado sejam interpretados e aplicados às especificidades do mercado de big data. Além disso, as melhores práticas e diretrizes também podem ser úteis para contribuir com o bem-estar dos cidadãos, bem como para o progresso socioeconômico.

Xanthidis et al. (2020), indo na mesma direção do estudo anterior, acreditam que o crescente interesse e uso de plataformas e sistemas que utilizam tecnologias emergentes como big data, serviços de computação em nuvem e Internet das Coisas (IoT) trazem muitos benefícios que pode afetar enormemente a atividade corporativa e pessoal e remodelar as sociedades globais. os negócios e a vida humana em geral. Contudo, estes benefícios podem afetar a privacidade da informação das pessoas. Assim sendo, há necessidade de encontrar um equilíbrio entre os múltiplos benefícios dos desenvolvimentos tecnológicos e a proteção dos direitos humanos básicos, como a privacidade das informações pessoais. Neste contexto, os autores aconselham estudar o âmbito escopo e o nível de penetração da IoT no país, e analisar os elementos qualitativos de esta penetração e as suas perspetivas nos próximos anos. Isto é especialmente interessante em termos de perspetiva da população local em relação ao impacto de tais desenvolvimentos na privacidade da informação. Sugere-se que novos

estudos nesta direção possam focar em aspetos específicos da economia emergente, dos usos da IoT que podem ter um impacto direto na sociedade e nas empresas locais.

Ainda na mesma esteira dos estudos anteriores, Zharova & Elin (2017) enfatizam que as leis, nomeadamente, as leis russas, que abrangem privacidade e proteção de dados, revelam-se inadequadas para proteger os direitos dos cidadãos em face ao uso cada vez maior de conjuntos massivos de dados e à sua análise por ferramentas de Big Data. Além disso, as sanções existentes para o uso indevido de dados pessoais são insignificantes e muitas vezes, não funcionam como impedidores quando os benefícios comerciais da exploração dos dados dos utilizadores (por exemplo, através de publicidade direcionada) são muito maiores. Deve haver responsabilidade mais clara e sanções mais rigorosas no caso de infrações. Deste modo, sublinham a necessidade de se desenvolver uma lei sobre "Big Data", na qual possamos identificar e definir riscos e ameaças. O nível desejado de proteção exigido no processamento de informações pode ser solicitado pela tecnologia Big Data. Nesta lei devemos definir os algoritmos de separação, tipos de dados, no processo de seu tratamento pelo Big Data e nos procedimentos de processamento dessas informações. Ademais, deveríamos definir o termo "operador nacional de 'Big Data'", que tem de especificar o tipo de público. Os autores vão além e sugerem a modificação do conceito de "dados pessoais" de forma que as leis que regem o tratamento de dados pessoais, também se apliquem aos operadores.

Outro estudo, que de certa forma, vai na mesma direção do estudo anterior é o dos autores Andrew & Baker (2021), ao abordarem sobre os desafios jurídicos da utilização de análises de big data nas mídias sociais, concluem que os esforços do RGPD relativo à ética da codificação de dados, especialmente em relação aos dados comportamentais, são limitados, visto que suas derrogações criam uma passagem através da qual as empresas conseguem escapar às restrições da lei. Com efeito, a lei cria espaço para um mercado de dados comportamentais no qual o interesse comercial florescerá em detrimento do interesse privado. Neste sentido, fazem duas recomendações: à necessidade de uma consideração jurídica adequada dos direitos de propriedade que envolvem a propriedade. Nesta perspetiva, é importante garantir que todos os

indivíduos sabem se seus dados podem ser negociados, mesmo que os dados são completamente anonimizados e conceder-lhes o direito de não disponibilizar seus dados anóônimos para esse fim de dados comportamentais. A outra recomendação é que os dados coletados com a intenção de serem anonimizados, devem estar sujeitos ao consentimento e acordado com o indivíduo.

Custers et al. (2018), ao realizarem uma comparação entre oito estados-membros da UE relativo a proteção da privacidade e dos dados pessoais (privacidade informacional), chegaram à conclusão de que há diferenças significativas entre os níveis de aplicação pelas diferentes autoridades de proteção de dados. Isto se dá devido a diferentes competências jurídicas, orçamentárias e de pessoal disponíveis, políticas e fatores culturais. Assim sendo, afirmam que sempre haverá diferenças nas leis e práticas nacionais e recomendam replicar esta pesquisa após o RGPD está em vigor há alguns anos.

Georgiadis & Poels (2022), ao pesquisarem sobre metodologias de Avaliação de Impacto à Privacidade (PIA) para lidar com riscos de privacidade e proteção de dados em Análise de Big Data, encontraram nove "Privacy Touch Points" (PTPs) como categorias de riscos específicos. Três metodologias se destacaram: a CNIL cobre PTPs de privacidade, o ICO do Reino Unido os de proteção de dados, e a abordagem LINDDUN abrange ambos. Outras sete metodologias, incluindo a ISO/IEC 29134:2017, possuem pouca cobertura para riscos específicos de privacidade e proteção de dados em Análise de Big Data. Os autores sublinham que é importante criar orientações específicas para Avaliações de Impacto à Proteção de Dados (DPIA) no contexto da Análise de Big Data. Assim sendo, considerando as categorias de riscos de privacidade e proteção de dados definidas neste estudo, planejam um estudo adicional usando a método Delphi para validar e aprimorar essas categorias, com a colaboração de especialistas em privacidade e big data. O objetivo é adequar as DPIAs do RGPD para o cenário da análise de Big Data.

Helm (2016) ao fazer uma pesquisa sobre privacidade de grupo no contexto do big data, conclui que os direitos de grupo em nome de valores fundamentais, como a justiça e a liberdade, poderiam ser alcançados por um discurso que trata da dimensão social da

privacidade e os valores que estão em questão. O desenvolvimento de tal direito exige um trabalho de equipa interdisciplinar, em primeiro lugar porque precisa ser pensado como uma reação às lacunas de proteção relacionadas a novos tipos de grupos. Tais lacunas devem ser identificadas de acordo com valores fundamentais, que foram violados devido à falta de regulamentação. Ou seja, é necessária uma nova forma de pesquisa interdisciplinar, que não envolve apenas a partilha de informações e perspetivas, mas também requer trabalho conjunto em equipes. Estas equipes precisam incluir, pelo menos, três perspetivas disciplinares: a perspetiva de disciplinas empiricamente especializadas, como antropologia empírica, informação ou ciência da comunicação, a perspetiva das disciplinas teoricamente especializadas como a teoria política, cultural ou social, bem como, a perspetiva normativa de juristas e especialistas em ética.

E, por fim, o estudo de Mantelero (2018), ao estudar a regulamentação de proteção de dados no contexto de aplicações com uso intensivo de dados para processos de tomada de decisão, aponta que este processo afeta a proteção de dados. Assim sendo, o autor sugere o desenvolvimento de formas mais amplas no impacto na proteção de dados, quer dizer, reconsiderar a natureza da avaliação social − uma vez que esta se torna excessivamente onerosa e complexa para as empresas, recomenda modelo voluntário, que mantém liberdade de decisão dos controladores de dados, tornando esta avaliação uma solução mais aceitável do que as disposições obrigatórias. Uma abordagem voluntária é mais consistente com o quadro jurídico existente, que parece ter dificuldades em ir além da mera proteção de dados no uso da informação. Neste sentido, o RGPD – que fornece um dos exemplos mais avançados de regulação nesta área, centra-se no risco.

3.1.4. Saúde

Nageshwaran et al. (2021), concluíram em seu estudo que as tecnologias digitais e big data fortaleceram medidas de saúde pública contra o SARS-CoV-2 durante a pandemia. Neste sentido, os autores sugerem consulta eficaz às partes interessadas acerca das questões de privacidade no sentido de maximizar a eficiência na implementação de

estratégias digitais. Assim, governos, empresas de tecnologia e organizações de saúde devem colaborar na criação de diretrizes para aplicativos digitais no combate à COVID-19, garantindo privacidade e foco na contenção da epidemia, visto que a cooperação estratégica se mostra essencial para evitar uso indevido de dados pessoais.

Minssen et al. (2020), ao analisarem os desafios legais associados e como podem afetar as organizações que implementam tecnologias médicas baseadas em nuvem que dependem de transferências de dados transfronteiriças de titulares de dados da EU, verificaram que litígios recente, avanços e desafios técnicos geraram incerteza legal significativa sobre a solidez e validade dos quadros legais internacionais para transferências de dados entre EUA/UE, preocupando desta forma algumas empresas inovadoras e stakeholders na área de desenvolvimento de medicamentos e saúde, que dependem de transferências internacionais eficazes e computação em nuvem. Assim sendo, apontam que é importante estabelecer salvaguardas adequadas de proteção de dados e incorporar novas tecnologias, melhorando a transparência e fornecendo remediações e compensações em caso de violações. Mecanismos legalmente sólidos são necessários para garantir operações de aplicativos de dados e proteção de direitos das partes envolvidas. Considerar soluções técnicas, códigos de conduta e mecanismos de certificação do RGPD é vital para enfrentar desafios e melhorar sistemas de transferência de dados.

Mostert et al. (2018), ao relatarem um estudo sobre a diferença do direto a privacidade e o direto à proteção dos dados no contexto da saúde dentro da UE, concluíram que um sistema abrangente de proteção de dados na pesquisa em saúde intensiva deve garantir duas funções-chave: salvaguardas gerais para proteger os direitos individuais, independentemente da base legal de processamento, e salvaguardas específicas quando necessário para permitir a reutilização de dados pessoais. Isso inclui prestações de contas, transparência, direitos dos sujeitos de dados e segurança, sugerindo, estudos adicionais de forma a obter a definição precisa dessas salvaguardas.

Petersen (2017), realça que o big data pode avançar a pesquisa médica, melhorar os resultados dos cuidados de saúde, tornar o atendimento médico mais acessível e

capacitar os indivíduos a terem maior controle sobre sua própria saúde. Contudo, o big data também apresenta ameaças significativas ao direito à privacidade e igualdade, especialmente quando atores privados discriminam com base em dados de saúde. Neste contexto, os governos devem adotar uma abordagem proativa para proteger indivíduos contra discriminação baseada em dados relacionados à saúde. Além de rever o alcance da legislação antidiscriminação, os governos podem considerar a implementação de leis para proibir o uso de mineração de dados de saúde em processos de tomada de decisão, relacionados a emprego, educação ou acesso a serviços financeiros. Pelo menos, os empregadores e outros atores privados devem ser obrigados a divulgar quando estão minerando dados e realizando processos de reidentificação.

Wolf (2018) ao analisar os dados de participantes na utilização de pesquisa genómica nos Estados Unidos, conclui que apesar das normas gerais relativas ao consentimento, a lei de investigação dos Estados Unidos permite a investigação utilizar registos médicos e bioespécimes sem consentimento. Além disso, apesar das normas gerais de confidencialidade do paciente, as leis de cuidados de saúde dos Estados Unidos geralmente permitem a utilização de registos para investigação e uma variedade de atividades de melhoria da qualidade. A informação médica individual pode ser utilizada para inúmeras funções governamentais legítimas, tais como atividades de saúde pública, supervisão da qualidade e auditorias dos gastos do governo. Contudo, muitos dos requisitos de consentimento que a pesquisa encontrou, dizem respeito a contextos específicos – investigação entre populações vulneráveis e testes genéticos, que podem não ser generalizáveis para informação médica de forma mais geral. Desta forma, o estudo identificou três riscos e ou implicações: uso de informações sem consentimento, acesso inadequado à informação e uso inadequado de informações. Para este autor, o fato deste procedimento não solicitar o consentimento informado, faz-se necessário uma melhor comunicação sobre como a informação individual sobre saúde é realmente tratada, é necessário considerar quais proteções são adequadas para proteger as informações em uso. Por outro lado, é preciso algum nível de proteção de confidencialidade. No mínimo, as proteções devem ser consistentes com as proteções

para informações médicas em geral. Contudo, vale a pena considerar se são necessárias proteções adicionais porque a utilização não beneficiará diretamente os indivíduos. Dadas as críticas ao “excecionalíssimo genético”, vale a pena discutir se a informação genética deve receber proteções adicionais, como vemos em algumas das leis dos EUA relativas a testes e informações genéticas. Outra sugestão é que haja alguma reflexão sobre o contexto jurídico e cultural em que estas proteções legais foram e serão adotadas. Se forem concedidas proteções, deverá haver algum mecanismo de aplicação, proporcionar uma reparação significativa aos indivíduos que sofrem danos.

3.1.5. Ética

Truby & Brown (2021) ao estudarem as questões legais e éticas levantadas pela clonagem digital e clones de "pensamentos digitais", observaram que o uso da clonagem digital é inevitável devido à digitalização dos processos e atividades humanas. Com o avanço da IA, alimentada pelo processamento de dados, é natural que a usemos para interpretar grandes quantidades de dados sobre nós mesmos. No entanto, a clonagem digital de pensamentos também traz riscos éticos e negativos, permitindo a previsão e manipulação de comportamentos pessoais com base em dados pessoais. Isso pode ter efeitos tanto positivos quanto negativos na sociedade, incluindo decisões políticas e pessoais. Assim, é importante exigir transparência sobre o uso de clones de pensamentos digitais, bem como responsabilidade e aplicabilidade dos algoritmos de IA que têm acesso a dados pessoais. Os autores também destacam a importância de dar às pessoas o direito de fazer uma escolha informada sobre seus dados antes de criar um clone de pensamento digital. Propõe considerar amplamente os aspetos legais e éticos levantados pela clonagem digital e clones de pensamentos digitais ao elaborar regulamentações.

Vayena & Tasioulas (2016) ao pesquisar a relação entre big data e direitos humanos na investigação científica, principalmente na área da saúde, concluíram que o big data tem um inexplorado potencial como abordagem à investigação científica. Contudo, paralelamente a concretização deste potencial, surge um grande desafio. Assim sendo, será necessário o desenvolvimento de uma ética adaptada às novas realidades, as novas

capacidades e riscos do ambiente digital em rápida evolução. Os direitos humanos, por exemplo, precisarão informar a ética do big data na pesquisa científica. Contudo, estes princípios não podem ser modelos fixos e pré-existentes que podem ser simplesmente impostas mecanicamente ao novo ambiente de uma forma padronizada. Em vez disso, os princípios destes devem passar por um processo dinâmico de evolução à medida que interagem com as mudanças ambiental, social e tecnológico, e como interagem entre si em conjunto.

3.1.6. Educação

Reidenberg & Schaub (2018), afirmam que uso crescente do Big Data na educação oferece benefícios, mas também riscos de privacidade. A mineração de dados educacionais e análises de aprendizado podem prejudicar a privacidade e equidade dos alunos. Assim, para equilibrar essas questões, os autores propõem que as salvaguardas legais para a privacidade na educação reflitam a realidade da educação orientada por dados, expandindo as proteções de privacidade para abranger claramente as análises de aprendizado. É essencial tornar obrigatórias as avaliações de impacto de privacidade e precaução como parte do desenvolvimento e implantação de sistemas em contextos educacionais. As regras de aquisição pública devem ser usadas como um meio importante para garantir a privacidade desde o início em projetos de Big Data para educação.

Steiner et al. (2016) buscam levantar em seu estudo, as considerações de privacidade e proteção de dados e a política formulada no projeto “LEA's BOX” (projeto financiado pela Comissão Europeia, o qual está desenvolvendo uma "toolbox" de análise de aprendizado) de forma a encontrar um equilíbrio entre a pesquisa de análise de aprendizado e a privacidade individual. O framework desenvolvido de proteção e privacidade dos dados estabelece a base para um código de conduta apropriado, exigindo que as tecnologias e ferramentas desenvolvidas no projeto estejam alinhadas a essas bases. As diretrizes definidas transformam-se em requisitos para o LEA’s BOX, implementados com uma abordagem de "ethics by design". Para efeito, os autores destacam as seguintes recomendações: 1) usar padrões de segurança atualizados e

estratégias adequadas de criptografia e anonimização de dados, 2) definir regras claras e bem documentadas de propriedade e direitos de acesso aos dados e exibi-las de forma adequada para todos os grupos de usuários.

3.1.7. Crianças

Montgomery et al. (2017), ao analisar as Implicações de privacidade em relação a publicidade em mídias sociais, aplicativos móveis e jogos direcionados a crianças, constatam que as Pesquisas que exploram as perceções de privacidade de crianças e adolescentes online são limitadas, ignorando rastreamento invisíveis de Jovens e, portanto, não compreendem completamente as práticas de marketing e os riscos. Na opinião dos autores, a regulamentação de privacidade infantil não considera adequadamente adolescentes e simplifica processos persuasivos dos anúncios. Neste sentido, os autores sugerem que profissionais de saúde devem atualizar declarações de políticas sobre mídia e propaganda infantil para abordar práticas contemporâneas de mídia digital, riscos à privacidade e vulnerabilidades de crianças mais velhas e adolescentes. Os formuladores de políticas devem ampliar as salvaguardas de privacidade infantil para abranger a coleta de dados e práticas de marketing em plataformas digitais, incluindo dispositivos de Internet das Coisas (IoT).

Caglar (2021) ao abordar o direito à privacidade e proteção de dados pessoais das crianças incorporado recentemente na lei RGPD, conclui que, embora a lei represente um avanço significativo relativo à proteção e privacidade dos dados das crianças, precisa de melhorias para encarar os desafios enfrentados na prática, especialmente quando implementar as leis relativas à obtenção do consentimento. Como sugestões, o autor recomenda que a idade de consentimento deve ser justificada com base em evidências, incluir as opiniões das crianças ao projetar um produto, uma vez que isso ajuda a garantir que os riscos sejam identificados e, por último, avaliar o nível de compreensão e, se necessário, alterar as políticas de privacidade. Explorar como se dá a obtenção do consentimento informado de titulares de dados que não são utilizadores e o equilíbrio entre as oportunidades apresentadas pela tecnologia e a crescente preocupações éticas, de privacidade e de segurança.

3.1.8. Política criminal

Drewer & Miladinova (2017) ao analisar a proteção dos direitos humanos fundamentais contra invasões terroristas externas e/ou da cibercriminalidade e o potencial uso do big data para a Europol, concluíram que a Europol está na posição ideal não só para abordar os mais importantes desafios futuros colocados por grandes quantidades de informação, mas também para explorar todas as oportunidades, a fim progredir ainda mais e produzir resultados tangíveis e eficazes para apoiar os objetivos de prevenção e combate à criminalidade enquadrado no âmbito dos seus objetivos. As capacidades únicas da Europol proporcionam a oportunidade de construir uma plataforma de sistema de informação capaz de facilitar uma gestão mais eficaz e eficiente, operacional e estratégica aos principais problemas policriminosos e as ameaças à segurança transfronteiriças. Contudo, este sistema em conjunto com as aplicações de big data esbarra na questão dos direitos humanos. Assim, os autores sugerem que a proteção de dados deverá conter, pelo menos, uma descrição geral das operações de tratamento previstas, uma avaliação dos riscos para a proteção de dados, medidas propostas para mitigar os riscos, garantias processuais, medidas de segurança e mecanismos para garantir a proteção dos dados pessoais e para

demonstrar conformidade com o Regulamento da Europol.

ID	Autor	Ano	País(es)	Método de pesquisa	Temática/Setor	Contexto	Resultados e Implicações	Sugestões e Lacunas
1	Nageshwaran, Gopinath ; Harris, Rebecca C ; Guerche-Seblain, Clotilde El	2021	Asia	Revisão da literatura	Saúde	Big data e tecnologias digitais para facilitar as medidas preventivas de controle da COVID-19 em quatro países asiáticos	As tecnologias digitais e big data fortaleceram medidas de saúde pública contra o SARS-CoV-2 durante a pandemia. Consulta eficaz às partes interessadas é vital para abordar questões de privacidade e maximizar a eficiência na implementação de estratégias digitais.	Governos, empresas de tecnologia e organizações de saúde devem colaborar na criação de diretrizes para aplicativos digitais no combate à COVID-19, garantindo privacidade e foco na contenção da epidemia. Cooperação estratégica é essencial para evitar uso indevido de dados pessoais.
2	Strahilevitz, Lior Jacob	2013	EUA	Revisão da literatura, Revisão documental	Regulamentação	As sutilezas da regulamentação de privacidade, com foco particular nas consequências distributivas das regras de privacidade no contexto dos EUA	Os desafios da era do Big Data, e como a indústria e o governo aplicam testes de personalidade refinados aos consumidores. O impacto da discriminação de personalidade em preços e serviços de mercado em massa, ressaltando a importância de compreender seus beneficiários e prejudicados para prever as reações legais.	Falta de uma disposição abrangente na lei de privacidade dos EUA, levando a regulamentações reativas e sugere soluções estruturais, como fortalecimento da proteção de privacidade ou reforço da Federal Trade Commission, para abordar as preocupações de privacidade percebidas nos EUA.
3	Minssen, Timo ; Seitz, Claudia ; Aboy, Mateo ; Corrales Compagnucci, Marcelo	2020	EUA, UE	Revisão da literatura, Revisão documental	Saúde	Privacy Shield Framework UE-EUA, os desafios legais associados e como podem afetar as organizações que implementam tecnologias médicas baseadas em nuvem que dependem de transferências de dados transfronteiriças de titulares de dados da UE	Litígios recente, avanços e desafios técnicos geraram incerteza legal significativa sobre a solidez e validade dos quadros legais internacionais para transferências de dados entre EUA/UE. Isso preocupa empresas inovadoras e stakeholders na área de desenvolvimento de medicamentos e saúde, que dependem de transferências internacionais eficazes e computação em nuvem.	É crucial estabelecer salvaguardas adequadas de proteção de dados e incorporar novas tecnologias, melhorando a transparência e fornecendo remediações e compensações em caso de violações. Mecanismos legalmente sólidos são necessários para garantir operações de aplicativos de dados e proteção de direitos das partes envolvidas. Considerar soluções técnicas, códigos de conduta e mecanismos de certificação do RGPD é vital para enfrentar desafios e melhorar sistemas de transferência de dados.
4	Montgomery, Kathryn C ; Chester, Jeff ; Milosevic, Tijana	2017	EUA	Revisão da literatura, Revisão documental	Crianças	Implicações de privacidade em relação a publicidade em mídias sociais, aplicativos móveis e jogos direcionados a crianças.	Pesquisas limitadas exploram perceções de privacidade de crianças e adolescentes online, frequentemente ignorando rastreamento invisível. Jovens não compreendem completamente práticas de marketing e riscos. Regulamentação de privacidade infantil não considera adequadamente adolescentes e simplifica processos persuasivos dos anúncios.	Profissionais de saúde devem atualizar declarações de políticas sobre mídia e propaganda infantil para abordar práticas contemporâneas de mídia digital, riscos à privacidade e vulnerabilidades de crianças mais velhas e adolescentes. Os formuladores de políticas devem ampliar as salvaguardas de privacidade infantil para abranger a coleta de dados e práticas de marketing em

								plataformas digitais, incluindo dispositivos de Internet das Coisas (IoT). Escolas, em colaboração com organizações sem fins lucrativos, deve criar programas de educação midiática e alfabetização digital para capacitar os jovens na cultura digital atual sem comprometer seu direito fundamental à privacidade.
5	Mostert, Menno ; Annelien L. Bredenoord ; Bart van der Slootb ; Johannes J.M. van Delden	2018	UE	Revisão da literatura, Revisão documental	Saúde	A diferença do direto a privacidade e o direto à proteção dos dados no contexto da saúde dentro da UE	Um sistema abrangente de proteção de dados na pesquisa em saúde intensiva deve garantir duas funções-chave: salvaguardas gerais para proteger os direitos individuais, independentemente da base legal de processamento, e salvaguardas específicas quando necessário para permitir a reutilização de dados pessoais. Isso inclui prestações de contas, transparência, direitos dos sujeitos de dados e segurança. A definição precisa dessas salvaguardas requer estudos adicionais.	É importante reconhecer os limites da lei de proteção de dados. A dependência na distinção entre dados pessoais e não pessoais para proteger privacidade e outros direitos relevantes pode ser insuficiente. Leis rígidas ou estáticas de proteção de dados podem prejudicar o desenvolvimento de estruturas de governança de informações adequadas em escala nacional ou internacional, onde diversas normas éticas, legais, sociais e profissionais precisam ser conciliadas.
6	Papakonstantinou, Vagelis ; de Hert, Paul	2020	EUA, UE	Revisão da literatura, Revisão documental	Regulamentação	Compreensão das implicações do perfilamento de dados corporativos e os "chilling effects" na área de comunicaçõeseletrônicas e sobre o regulamento ePrivacy	Dois principais problemas de regulamentação de políticas foram identificados na pesquisa acadêmica. Primeiro, a necessidade social de ampliar as operações de análise de big data, enquanto o segundo refere-se às diferenças entre abordagens de análise de big data entre o RGPD e as leis de ePrivacy. O projeto de Regulamento de ePrivacy não atendeu às expectativas legislativas ao não reconhecer a necessidade de regular a análise de big data e ao não fornecer especificidade na diferenciação entre operadores de telecomunicações e empresas de internet.	Sugerem que o atual projeto de Regulamento de ePrivacy seja retirado e revisado de forma abrangente, especialmente em relação à proteção de dados. Enquanto isso, o RGPD poderia servir como referência para o campo de comunicações eletrônicas, combinado com a Diretiva de ePrivacy conforme emendada pelo Código de Comunicações Eletrônicas. Essas abordagens proporcionariam aos legisladores da UE tempo para desenvolver um novo texto de ePrivacy mais detalhado e adequado às circunstâncias sociais e de mercado, em linha com os objetivos de proteção de dados.
7	Park, Yong Jin ; Sang, Yoonmo ; Lee, Hoon ; Jones-Jang, S Mo	2020	EUA	Revisão da literatura	Regulamentação	O que acontece com os ativos digitais deixados após a morte. O foco está na responsabilidade social de definir direitos de controle	Primeiramente, abordagens "únicas" não são eficazes para abranger as diversas dimensões dos ativos digitais, que apresentam problemas dimensionais. Além disso, a ontologia dos ativos digitais após a morte vai além da simples transferência de	É fundamental codificar de maneira explícita o princípio ético de tratar os ativos digitais após a morte com respeito, estabelecendo um amplo quadro regulatório para limitar a exploração comercial. Isso requer a criação de normas e requisitos padronizados,

						sobre esses ativos após a morte, abordando tanto a privacidade quanto os direitos de propriedade.	propriedade entre o falecido e os familiares. Os direitos sobre esses ativos permanecem em um contexto de autorregulação, muitas vezes determinado por processos judiciais civis, enquanto as forças de mercado frequentemente marginalizam o status do usuário digital após a morte. A ausência de integridade decorre da falta de voz dos principais interessados, ou seja, dos usuários falecidos.	com poder regulatório, em esforços internacionais harmonizados. Além disso, deve haver uma cláusula de proteção que preserve os direitos de privacidade e propriedade dos ativos digitais após a morte, proibindo usos comerciais não autorizados e alterações. O objetivo ético é garantir que esses ativos sejam tratados com o mesmo respeito que durante a vida da pessoa, preservando a integridade de seus propósitos originais.
8	Petersen, Carole J	2017	Asia	Revisão da literatura	Saúde	As implicações dos direitos humanos do cuidado de saúde orientado por dados, concentrando-se nos direitos das pessoas que vivem com deficiências ou que possam ser percebidas como tendo um risco elevado de desenvolver uma deficiência no futuro.	O big data pode avançar a pesquisa médica, melhorar os resultados dos cuidados de saúde, tornar o atendimento médico mais acessível e capacitar os indivíduos a terem maior controle sobre sua própria saúde. No entanto, o big data também apresenta ameaças significativas ao direito à privacidade e igualdade, especialmente quando atores privados discriminam com base em dados de saúde.	Os governos devem adotar uma abordagem proativa para proteger indivíduos contra discriminação baseada em dados relacionados à saúde. Além de rever o alcance da legislação antidiscriminação, os governos podem considerar a implementação de leis para proibir o uso de mineração de dados de saúde em processos de tomada de decisão relacionados a emprego, educação ou acesso a serviços financeiros. Pelo menos, os empregadores e outros atores privados devem ser obrigados a divulgar quando estão minerando dados e realizando processos de reidentificação.
9	Reidenberg, Joel R ; Schaub, Florian	2018	EUA	Revisão da literatura, Revisão documental	Educação	Como as tecnologias de aprendizado também criam tensões éticas entre privacidade e o uso de Big Data para melhorias na educação	O uso crescente do Big Data na educação oferece benefícios e riscos de privacidade. A mineração de dados educacionais e análises de aprendizado podem prejudicar a privacidade e equidade dos alunos. Para equilibrar essas questões, é necessário combinar mecanismos tecnológicos, salvaguardas organizacionais e proteções legais, além de realizar avaliações do impacto educacional e garantir transparência, responsabilidade e segurança nas tecnologias de análise de aprendizado.	É necessário que as salvaguardas legais para a privacidade na educação reflitam a realidade da educação orientada por dados, expandindo as proteções de privacidade para abranger claramente as análises de aprendizado. É essencial tornar obrigatórias as avaliações de impacto de privacidade e precaução como parte do desenvolvimento e implantação de sistemas em contextos educacionais. As regras de aquisição pública devem ser usadas como um meio importante para garantir a privacidade desde o início em projetos de Big Data para educação.

10	Steiner, Christina M ; Michael D. Kickmeier-Rust ; Dietrich Albert	2016	UE	Revisão da literatura	Educação	LEA's BOX é um projeto de pesquisa e desenvolvimento financiado pela Comissão Europeia que está desenvolvendo uma "toolbox" de análise de aprendizado. Neste artigo, as considerações de privacidade e proteção de dados e a política formulada para o projeto são delineadas para encontrar um equilíbrio entre a pesquisa de análise de aprendizado e a privacidade individual.	O framework desenvolvido de proteção e privacidade dos dados estabelece a base para um código de conduta apropriado, exigindo que as tecnologias e ferramentas desenvolvidas no projeto estejam alinhadas a essas bases. As diretrizes definidas se transformam em requisitos para o LEA's BOX, implementados com uma abordagem de "ethics by design".	Destacam as seguintes recomendações: 1) usar padrões de segurança atualizados e estratégias adequadas de criptografia e anonimização de dados, 2) definir regras claras e bem documentadas de propriedade e direitos de acesso aos dados e exibi-las de forma adequada para todos os grupos de usuários. Além disso, 3) permitir que os usuários, especialmente os aprendizes, influenciem e intervenham no sistema e suas análises, e 4) projetar todos os algoritmos, análises e visualizações com a consciência de que podem estar equivocados. No final, essas medidas devem ser comunicadas para otimizar a confiança e a credibilidade.
11	Truby, Jon ; Brown, Rafael	2021	EUA, UE	Revisão da literatura	Ética	As questões legais e éticas levantadas pela clonagem digital e clones de "pensamentos digitais", e a necessidade de re-conceitualizar as atuais noções teóricas sobre privacidade de dados	O uso da clonagem digital é inevitável devido à digitalização dos processos e atividades humanas. Com o avanço da IA, alimentada pelo processamento de dados, é natural que a usemos para interpretar grandes quantidades de dados sobre nós mesmos. No entanto, a clonagem digital de pensamentos também traz riscos éticos e negativos, permitindo a previsão e manipulação de comportamentos pessoais com base em dados pessoais. Isso pode ter efeitos tanto positivos quanto negativos na sociedade, incluindo decisões políticas e pessoais.	Exigir transparência sobre o uso de clones de pensamentos digitais e responsabilidade e explicabilidade dos algoritmos de IA que têm acesso a dados pessoais. Também é destacada a importância de dar às pessoas o direito de fazer uma escolha informada sobre seus dados antes de criar um clone de pensamento digital. Propõe considerar amplamente os aspetos legais e éticos levantados pela clonagem digital e clones de pensamentos digitais ao elaborar regulamentações.
12	Tzanou, Maria	2017	EUA, UE	Revisão da literatura	Regulamentação	Desafios da Privacidade de Dados na UE: Admissibilidade, Fluxos Transfronteiriços e Vigilância Eletrônica nos EUA	O julgamento Schrems é um marco constitucional na proteção judicial dos direitos fundamentais no âmbito do contraterrorismo. Ele reforça o direito à privacidade contra a vigilância eletrônica moderna e destaca que o acesso a dados pessoais por autoridades públicas não se justifica com base na mera disponibilidade desses dados por empresas privadas. O julgamento também amplia a	O recém-adotado Privacy Shield, que substitui o sistema Safe Harbour invalidado, não aborda de maneira satisfatória as preocupações com os direitos fundamentais levantadas pelo TJUE. Suas salvaguardas para direitos fundamentais são limitadas e mistura o regime de transferência de dados transatlânticos com a regulamentação de operações de contraterrorismo. A criação de um Mecanismo de

							interpretação das leis de proteção de dados da UE para transferências de dados transfronteiriços e reafirma a importância da ativação da privacidade. No entanto, a análise dos direitos fundamentais pelo tribunal ainda apresenta lacunas e omissões.	Ombudsman é a única novidade, porém não garante total reparação para indivíduos
13	Ursic, Helena and Custers, Bart	2016	UE	Revisão documental	Regulamentação	As barreiras e facilitadores legais existentes para a reutilização de big data na lei da UE, com base na abordagem que segue a dicotomia tradicional das leis.	A análise revela que a lei de proteção de dados atua principalmente como uma barreira, limitando o processamento de dados e impondo obrigações adicionais aos reutilizadores de dados. Direitos humanos também podem ser uma barreira, pois questionam a reutilização de dados que não considera os direitos e interesses dos titulares. Leis de retenção de dados e segurança cibernética podem ser tanto barreiras quanto facilitadores, dependendo das exigências impostas. As leis de propriedade intelectual e concorrência podem atuar como barreiras ou facilitadores para a reutilização de dados, dependendo de como afetam a exploração de conjuntos de dados e a concorrência justa. A lei de proteção do consumidor pode bloquear práticas comerciais injustas ou promover publicidade mais precisa e direcionada, dependendo do contexto da reutilização de dados.	O cenário regulatório da UE é altamente complexo quando se trata de reutilização de dados. Identificar as barreiras e facilitadores mais importantes para a reutilização de dados pode ser útil para regular ainda mais a reutilização de dados, a fim de facilitar um ambiente digital sustentável e dinâmico.
14	Van der Sloot, Bart	2015	UE	Revisão documental	Regulamentação	As violações da privacidade e jurisprudência do Tribunal Europeu dos Direitos Humanos (TEDH)	A Convenção Europeia dos Direitos Humanos enfatiza que a privacidade é um direito voltado principalmente para o interesse privado. Em contraste com outros direitos qualificados protegidos pelo TEDH, o direito à privacidade é frequentemente limitado por objetivos sociais, como segurança, proteção de valores sociais e bem-estar econômico do país. O tribunal parece adotar diferentes teorias, como a teoria unitária para segurança, a teoria da preponderância para valores sociais e a teoria do interesse geral para questões econômicas. Em geral, o equilíbrio entre o interesse	É incerto se as reivindicações sobre o Big Data serão declaradas admissíveis sob a Convenção Europeia dos Direitos Humanos. Além disso, mesmo se forem consideradas admissíveis, não está claro como o TDHE poderia abordar esses dilemas de maneira satisfatória. Os testes existentes de interesse público, interesse individual e equilíbrio de interesses não se aplicam bem ao Big Data devido à complexidade dos interesses em jogo. Portanto, um quarto teste precisa ser desenvolvido, considerando desenvolvimentos estruturais e sociais, para que o TDHE possa abordar

							privado e o interesse comum é avaliado de acordo com essas teorias em diferentes contextos.	eficazmente questões de privacidade relacionadas ao Big Data e manter a relevância do Artigo 8 da ECHR.
15	Van der Sloot, Bart	2018	UE	Revisão da literatura, Revisão documental	Regulamentação	Violação do direito a privacidade, analisando a jurisprudência do Tribunal Europeu dos Direitos Humanos (TEDH) e interpretando a jurisprudência sob uma perspetiva filosófica.	O direito à privacidade está vinculado aos interesses e direitos individuais, exigindo demonstração de interferência específica e dano real para ser considerado admissível sob a Convenção Europeia dos Direitos Humanos. O surgimento de tecnologias modernas torna difícil identificar violações concretas e danos pessoais. O TEDH avalia alegações em abstrato, analisando a qualidade das leis em casos de vigilância em massa. A abordagem pode ser influenciada pela discussão filosófica sobre liberdade, com a atenção se voltando para o princípio de 'não-dominação', que considera relações de poder e potencial de uso abusivo do poder.	Em casos de violações tradicionais de privacidade, o princípio de não interferência é aplicado, restringindo reclamações apenas aos diretamente afetados. No entanto, em situações envolvendo novas tecnologias de dados, como vigilância em massa por serviços de inteligência, o TEDH adota o princípio de não-dominação, permitindo ações mesmo sem danos individuais claros. Essa abordagem amplia a proteção legal e impõe limites ao poder governamental, mas é criticada por desviar-se dos procedimentos legais nacionais e por conceder ao TEDH um poder não estabelecido na Convenção Europeia dos Direitos Humanos, possivelmente afetando a legitimidade democrática e o ideal do Estado de Direito.
16	Vayena, Effy ; Tasioulas, John	2016	UE	Revisão da literatura	Ética	A relação entre big data e direitos humanos na investigação científica, principalmente na área da saúde e, versa sobre dois outros direitos: direito à privacidade familiar e direito a ciência.	O big data tem um inexplorado potencial como abordagem à investigação científica. Contudo, paralelamente a concretização deste potencial, surge um grande desafio. Assim sendo, será necessário o desenvolvimento de uma ética adaptada às novas realidades, as novas capacidades e riscos do ambiente digital em rápida evolução.	Será necessário o desenvolvimento de uma ética adaptada às novas realidades, às novas capacidades e riscos de um ambiente digital em ampla evolução. Os direitos humanos, por exemplo, precisarão informar a ética do big data na pesquisa científica. Contudo, estes princípios não podem ser modelos fixos e pré-existentes que podem ser simplesmente impostas mecanicamente ao novo ambiente de uma forma padronizada. Em vez disso, os princípios eles próprios estão passando por um processo dinâmico de evolução à medida que interagem com as mudanças ambiental, social e tecnológico, e como interagem entre si em conjunto. Este artigo não apresenta lacunas.
17	Crawford, Kate ; Jason Schultz	2014	EUA	Revisão da literatura	Regulamentação	Para responder às práticas em evolução do Big Data, este artigo examina diversos regimes de privacidade	A natureza das ferramentas analíticas dinâmicas do Big Data é tal que os problemas de privacidade dos algoritmos preditivos muitas vezes são imprevisíveis por si próprios, e seus efeitos podem nem mesmo ser	Este artigo propõe o devido processo de dados procedimentais. Em vez de tentar regular a coleta, uso ou divulgação de dados pessoais “ex ante”, o devido processo de dados procedimentais regularia a

						existentes e explica por que essas abordagens não abordam adequadamente os desafios atuais do Big Data.	totalmente compreendidos por seus programadores. Como os cientistas da computação demonstraram, em muitos contextos, é impossível garantir a privacidade diferencial ao usar um algoritmo de aprendizado que retira dados de uma distribuição contínua. Isso torna difícil prever quando um algoritmo de aprendizado preditivo revelará informações de identificação pessoal sobre um indivíduo, tornando desafiador definir proteções de privacidade em torno desses dados. O uso do Big Data não apenas pode contornar regulamentos antidiscriminação existentes, mas também pode levar a violações de privacidade na área da saúde e da aplicação da lei.	equidade dos processos analíticos do Big Data em relação à forma como eles usam dados pessoais (ou metadados derivados ou associados a dados pessoais) em qualquer processo adjudicativo, incluindo processos nos quais o Big Data é usado para determinar atributos ou categorias para um indivíduo. Isso permitiria aos indivíduos exercerem seus direitos de devido processo de dados em relação a essas determinações.
18	Wang, Faye Fangfei	2017	UE	Revisão da literatura	Regulamentação	O artigo visa eliminar obstáculos legais na era do big data ao definir o big data de uma perspetiva legal e buscar soluções para os principais desafios legais do big data, como privacidade e segurança de dados, proteção de propriedade intelectual e questões jurisdicionais. Propõe uma interpretação e aplicação viáveis dos conceitos tradicionais ao novo fenómeno do big data.	O big data parece transformar a maneira como as pessoas vivem, trabalham e pensam. Também pode aumentar a eficiência, produtividade, segurança, conveniência, oportunidades e lucros nos negócios e na vida diária. Conforme apontado neste artigo, o fenómeno do big data ainda enfrenta incertezas legais, já que não existe uma legislação única que aborde especificamente questões legais do big data.	Os conceitos e princípios legais tradicionais do direito de banco de dados, do direito de proteção de dados, do direito de propriedade intelectual, do direito contratual e do direito internacional privado precisam ser interpretados e aplicados às especificidades do mercado de big data. Além disso, as melhores práticas e diretrizes também podem ser úteis para contribuir com o bem-estar dos cidadãos, bem como para o progresso socioeconômico.
19	Wolf, Leslie E	2018	EUA	Revisão documental	Saúde	Os dados de participantes na utilização de pesquisa genômica nos Estados Unidos.	Apesar das normas gerais relativas ao consentimento, a lei de investigação dos Estados Unidos permite a investigação utilizar registos médicos e bioespécimes sem consentimento. Além disso, apesar das normas gerais de confidencialidade do paciente, as leis de cuidados de saúde dos Estados Unidos geralmente	Ao não solicitar o consentimento informado, é necessária uma melhor comunicação sobre como a informação individual sobre saúde é realmente. É necessário considerar quais proteções são adequadas para proteger as informações em uso. Por outro lado, é preciso algum nível de proteção de

							permitem a utilização de registos para investigação e uma variedade de atividades de melhoria da qualidade. A informação médica individual pode ser utilizada para inúmeras funções governamentais legítimas, tais como atividades de saúde pública, supervisão da qualidade e auditorias dos gastos do governo. Contudo, muitos dos requisitos de consentimento que a pesquisa encontrou, dizem respeito a contextos específicos – investigação entre populações vulneráveis e testes genéticos, que podem não ser generalizáveis para informação médica de forma mais geral. Desta forma o artigo identificou três riscos e ou implicações: uso de informações sem consentimento, acesso inadequado à informação e uso inadequado de informações.	confidencialidade. No mínimo, as proteções devem ser consistentes com as proteções para informações médicas em geral. Contudo, vale a pena considerar se são necessárias proteções adicionais porque a utilização não beneficiará diretamente os indivíduos. Dadas as críticas ao “excepcionalismo genético”, vale a pena discutir se a informação genética deve receber proteções adicionais, como vemos em algumas das leis dos EUA relativas a testes e informações genéticas. Outrasugestão é que haja alguma reflexão sobre o contexto jurídico e cultural em que estas proteções legais foram e serão adotadas.Se forem concedidas proteções, deverá haver algum mecanismo de aplicação, proporcionar uma reparação significativa aos indivíduos que sofrem danos como resultado da utilização das suas informações para investigação sobre a causa dos cuidados de saúde baseados em dados.
20	Xanthidis, Dimitrios ; Manolas, Christos ; Xanthidou, Ourania Koutzampasopoulou ; Wang, Han-I	2020	UAE	Revisão da literatura	Regulamentação	A privacidade da informação e às tecnologias de informação emergentes de Big Data (analítica), Computação e Serviços em Nuvem e Internet das Coisas, e sobre o status atual de privacidade de informações nos Emirados Árabes Unidos.	O crescente interesse e uso de plataformas e sistemas que utilizam tecnologias emergentes como big data (análises), serviços de computação em nuvem e Internet das Coisas (IoT) trazem muitos benefícios que pode afetar enormemente a atividade corporativa e pessoal e remodelar as sociedades globais, os negócios paisagem e da vida humana em geral. No entanto, estes benefícios podem afetar a privacidade da informação das pessoas. Há necessidade de encontrar um equilíbrio entre os múltiplos benefícios dos desenvolvimentos tecnológicos e a proteção dos direitos humanos básicos, como a privacidade das informações pessoais.	Neste contexto, é interessante e relevante estudar o escopo e o nível de penetração da IoT no país, e analisar os elementos qualitativos de esta penetração e as suas perspetivas nos próximos anos. Isto é especialmente interessante em termos de perspetiva da população local em relação ao impacto de tais desenvolvimentos na privacidade da informação. Sugere-se que novos estudos nesta direção possam focar em aspetos específicos da economia emergente, dos usos da IoT que podem ter um impacto direto na sociedade e nas empresas locais.
21	Zharova, Anna Konstantinovna ;	2017	Russia	Revisão da literatura	Regulamentação	O impacto da tecnologia Big Data nos direitos constitucionais dos cidadãos	Apesar de existir diversas leis na Federação Russa que abrangem privacidade e proteção de dados, estas se revelam	Deve haver responsabilidade mais clara e sanções mais rigorosas no caso de infrações. É necessário desenvolver uma lei sobre “Big Data”, na qual

	Vladimir Mikhailovich Elin					russos relativos a privacidade da informação.	inadequadas para proteger os direitos dos cidadãos em face ao uso cada vez maior de conjuntos massivos de dados e à sua análise por ferramentas de Big Data. Além disso, as sanções existentes para o uso indevido de dados pessoais os dados são insignificantes e muitas vezes não funcionam como dissuasores quando os benefícios comerciais da exploração dos dados dos utilizadores (por exemplo, através de publicidade direcionada) são muito maiores.	devemos identificar e definir riscos e ameaças. O nível desejado de proteção exigido no processamento de informações pode ser solicitado pela tecnologia Big Data. Nesta lei devemos definir os algoritmos de separação, tipos de dados, no processo de seu tratamento pelo Big Data e nos procedimentos de processamento dessas informações. Nesta lei, deveríamos definir o termo "operador nacional de 'Big Data'", que tem de trabalhar o tipo de público. É necessário modificar o conceito de "dados pessoais" para que as leis que regem o tratamento de dados pessoais também se apliquem aos operadores.
22	Andrew, Jane ; Max Baker	2021	UE	Revisão documental	Regulamentação	Aborda sobre os desafios jurídicos da utilização de análises de big data nas mídias sociais. nos logs de software e no rastreamento de transações financeiras.	Atualmente o esforço do RGPD relativo a ética da codificação de dados, especialmente em relação aos dados comportamentais, é limitada. Suas derrogações criam uma passagem através da qual as empresas conseguem escapar às restrições da lei. Com efeito, a lei cria espaço para um mercado de dados comportamentais no qual o interesse próprio comercial provavelmente florescerá.	Há duas recomendações que emergem da análise deste artigo: à necessidade de uma consideração jurídica adequada dos direitos de propriedade que envolvem a propriedade. Nestaperspetiva, é importante garantir que todos os indivíduos sabem se seus dados podem ser negociados, mesmo que os dados são completamente anonimizados e conceder-lhes o direito de não disponibilizar seus dados anônimos para esse fim.de dados comportamentais. A outra recomendação é que os dados coletados com a intenção de serem anonimizados, devem estar sujeitos ao consentimento e acordo do indivíduo.
23	Caglar, Cansu	2021	UE	Revisão da literatura	Crianças	Aborda o direito a privacidade e proteção de dados pessoais das crianças incorporados recentemente na lei RGPD.	Embora a lei RGPD represente um avanço significativo relativo à proteção e privacidade dos dados das crianças, precisa de melhorias para encarar os desafios enfrentados na prática, especialmente quando implementar as leis relativas à obtenção do consentimento.	Como sugestões, o autor recomenda que a idade de consentimento deve ser justificada com base em evidências, incluir as opiniões das crianças ao projetar um produto, uma vez que isso ajuda a garantir que os riscos sejam identificados e, por último, avaliar o nível de compreensão e, se necessário, alterar as políticas de privacidade. Explorar como se dá a obtenção do consentimento informado de titulares de dados que não são utilizadores e o equilíbrio entre as

								oportunidades apresentadas pela tecnologia e a crescente preocupações éticas, de privacidade e de segurança.
24	Bart Custers ; Francien Dechesne ; Alan M. Sears ; Tommaso Tani ; Simone van der Hof	2018	UE	Revisão documental	Regulamentação	Faz uma comparação entre oito estados-membros da UE relativo a proteção da privacidade e dos dados pessoais (privacidade informacional). A comparação centra-se em cinco temas principais: sensibilização e confiança, políticas governamentais para proteção de dados pessoais, as leis aplicáveis e regulamentos, implementação dessas leis e regulamentos, e supervisão e aplicação.	Diferenças significantes existem nos (níveis de) aplicação pelas diferentes autoridades de proteção de dados, devido a diferentes competências jurídicas, orçamentos e pessoal disponíveis, políticas e fatores culturais.	Fornece sugestões sobre como um país poderia mover-se numa direção específica em relação a determinados aspetos do seu quadro de proteção de dados. pode-se esperar que continuarão a existir diferenças nas leis e práticas nacionais. Portanto, recomendamos replicar esta pesquisa após o RGPD está em vigor há alguns anos.
25	Drewer, Daniel ; Miladinova, Vesela	2017	UE	Revisão narrativa	Política criminal	A proteção dos direitos humanos fundamentais contra invasões terroristas externas e/ou da cibercriminalidade e o potencial uso do big data para a Europol.	Europol está na posição ideal não só para abordar os mais importantes desafios futuros colocados por grandes quantidades de informação, mas também para explorar todas as oportunidades, a fim de para progredir ainda mais e produzir resultados tangíveis e eficazes apoiar os objetivos de prevenção e combate à criminalidade enquadrado no âmbito dos seus objetivos. As capacidades únicas da Europol proporcionam a oportunidade de construir um sistema de informação plataforma capaz de facilitar uma gestão mais eficaz e eficiente resposta operacional e estratégica aos principais problemas policriminosos e ameaças à segurança transfronteiriças.	A proteção de dados deverá conter, pelo menos, uma descrição geral das operações de tratamento previstas, uma avaliação dos os riscos para a proteção de dados, as medidas propostas para mitigar os riscos, as garantias processuais, as medidas de segurança e mecanismos para garantir a proteção dos dados pessoais e para demonstrar conformidade com o Regulamento Europol.

26	Georgiadis, Georgios ; Poels, Geert	2022	UE	Revisão sistemática da literatura	Regulamentação	Riscos de privacidade e proteção de dados específicos do contexto de Big Data Analytics que poderiam impactar negativamente os direitos e liberdades dos indivíduos	Pesquisas abordaram metodologias de Avaliação de Impacto à Privacidade (PIA) para lidar com riscos de privacidade e proteção de dados em Análise de Big Data. Encontraram-se nove "Privacy Touch Points" (PTPs) como categorias de riscos específicos. Três metodologias se destacaram: a CNIL cobre PTPs de privacidade, o ICO do Reino Unido os de proteção de dados, e a abordagem LINDDUN abrange ambos. Outras sete metodologias, incluindo a ISO/IEC 29134:2017, possuem pouca cobertura para riscos específicos de privacidade e proteção de dados em Análise de Big Data.	Uma abordagem preferencial é criar orientações específicas para Avaliações de Impacto à Proteção de Dados (DPIA) no contexto da Análise de Big Data, considerando as categorias de riscos de privacidade e proteção de dados definidas neste estudo. Planeja-se um estudo adicional usando a técnica Delphi para validar e aprimorar essas categorias, com a colaboração de especialistas em privacidade e big data. O objetivo é adequar as DPIAs do RGPD para o cenário da Análise de Big Data.
27	Helm, Paula	2016	Canadá	Revisão da literatura	Regulamentação	A privacidade de grupo no contexto do Big Data	Os direitos de grupo em nome de valores fundamentais, como a justiça e a liberdade, poderiam ser alcançadas por um discurso que trata da dimensão social da privacidade e os valores que estão em questão. O desenvolvimento de tal direito exige um trabalho de equipa interdisciplinar, em primeiro lugar porque precisa ser pensado como uma reação às lacunas de proteção relacionadas a novos tipos de grupos Tais lacunas devem ser identificadas de acordo com valores fundamentais, que foram violados devido à falta de regulamentação.	É necessária uma nova forma de pesquisa interdisciplinar, que não envolve apenas a partilha de informações e perspetivas, mas também requer trabalho conjunto em equipes. Essas equipes precisam incluir, pelo menos, três perspetivas disciplinares: a perspetiva de disciplinas empiricamente especializadas, como antropologia empírica, informação ou ciência da comunicação, a perspetiva das disciplinas teoricamente especializadas como a teoria política, cultural ou social, bem como a perspetiva normativa de juristas e especialistas em ética.
28	Mantelero, Alessandro	2018	UE	Revisão da literatura	Regulamentação	A regulamentação de proteção de dados no contexto de aplicações com uso intensivo de dados para processos de tomada de decisão.	O uso intensivo de análises de Big Data para processos de tomada de decisão afeta a proteção dos dados.	Neste contexto sugere-se o desenvolvimento de formas mais amplas de impacto na proteção de dados: reconsiderar a natureza da avaliação social, uma vez que está se torna excessivamente onerosa e complexa para as empresas, recomenda-se, assim, um modelo voluntário, que mantém liberdade de decisão dos controladores de dados, tornando esta avaliação uma solução mais aceitável do que as disposições obrigatórias. Uma abordagem voluntária é mais consistente com o quadro jurídico existente, que

								parece ter dificuldades em ir além da mera proteção de dados no uso da informação. Neste sentido, o RGPD – que fornece um dos exemplos mais avançados de regulação nesta área – centra-se no risco.

Figura 19- Tabela da recolha dos dados.

Fonte: Autor

4. Discussão

O estudo realizado teve como finalidade entender como a literatura especializada tem tratado a informação e o direito à privacidade no contexto do Big Data. Para isso, procurou-se identificar estudos que relatem a problemática da informação e a privacidade no contexto do Big Data; identificar os métodos de pesquisa, países de foco, contexto e ano de publicação dos estudos selecionados; identificar sugestões de diretrizes e/ou ações que possam harmonizar interesses pessoais de usuários e interesses econômicos no contexto do Big Data; identificar e relatar as lacunas ainda existente na literatura no que concerne o acesso à informação e a privacidade no contexto do Big Data e propor novos direcionamentos para os futuros pesquisadores.

O resultado obtido coloca em evidência que há um número considerável de investigações preocupadas com a privacidade da informação no contexto do big data. Esta constatação vem de encontro aos resultados obtidos neste estudo. De Mauro et al. (2016) que sublinham que as aplicações de big data, tornou-se comum em trabalhos académicos e científicos tanto nas áreas das Tecnologias de Informação e Comunicação (TIC) quanto em outras áreas do conhecimento, como: sociologia, medicina, biologia, economia e gestão.

No âmbito desta pesquisa, foram identificados e analisados diversos estudos que abordam a complexa interação entre informação, privacidade e o uso crescente do Big Data. No que se refere aos métodos de pesquisa, o fato de termos realizado uma metassíntese qualitativa, incluímos apenas estudos com a abordagem qualitativa. Dentre estes estudos qualitativos, o método utilizado foi a revisão da literatura, variaram entre os tipos de pesquisa bibliográfica, pesquisa documental e pesquisa bibliográfica com pesquisa documental.

O estudo observou que a maioria dos estudos analisados tinha como foco países pertencentes à União Europeia. Isso reflete as crescentes preocupações regulatórias e legais que surgiram na região devido à rápida evolução das tecnologias de Big Data e suas implicações para a privacidade das pessoas. A abordagem mais rígida da UE em

relação à proteção de dados, especialmente com a implementação do RGPD, influenciou significativamente a pesquisa nesta área. No entanto, observou-se a carência de estudos publicados no Brasil, talvez esta situação se dá pelo facto do regulamento de proteção de dados no Brasil ter sido promulgado apenas em 2018, posterior a Europa e EUA, ou pela base de dados utilizada não conter publicações desse país.

Referente ao contexto, o maior número de estudos concentrou-se na questão da regulamentação. Este quadro reafirma a preocupação dos investigadores sobre a normatização das análises de big data, especialmente, no âmbito da privacidade da informação do indivíduo. De acordo com os estudos (Zharova & Elin , 2017; Andrew & Baker, 2021; Georgiadis et al, 2022), embora o regulamento de proteção de dados tenha representado um avanço significativo para a proteção da informação das pessoas, ainda requer melhorias de forma a encarar os desafios enfrentados na prática. Neste sentido, Mantelero (2018) propõe o desenvolvimento de um regulameno mais amplo no que tange o impacto da proteção de dados, quer dizer, reconsiderar a natureza da avaliação social. Esta posição está em conformidade com a literatura, nomeadamente com as reflexões de Floridi (2011), antecipando os avanços tecnológicos da sociedade da informação, ele propõe estudar o fenómeno da informação sob uma perspetiva filosófica de forma que as questões de privacidade sejam capazes de acompanhar as novas formas de violação decorrentes das novas configurações sociais. Por outro lado, a concentração desses estudos no âmbito da regulamentação, pode-se configurar uma preocupação, visto que há outros campos de atuação do big data, que representa uma fonte profícua da informação no contexto do big data, como por exemplo, a internet das coisas (IoT). Esta proporciona aos objetos do dia a dia (smartphones, veículos, aspirador de pó, frigorífico, prédios e outros providos de tecnologias com capacidade computacional e de comunicação) conectados à internet uma variedade de tipos e formatos de dados. Esta variedade de tipos e formatos de dado representa um dos principais desafios para as aplicações do *Big Data* e para os legisladores. O uso desses objetos inteligentes é possível controlar espaços, tempos e informações. Contudo, juntamente com todos estes benefícios emergem novos desafio (regulamentação,

segurança e padronização) que merecem uma ampla discussão no meio académico. Portanto, os resultados mostram a necessidade de investigação com foco, principalmente, nas aplicações da internet das coisas e sua conexão com a inteligência artificial.

A revisão da literatura abrangeu um período significativo, que compreendeu os anos de 2012 a 2022. Esse intervalo de dez anos permitiu uma análise detalhada da evolução das preocupações e abordagens relacionadas à informação, privacidade e Big Data. Além disso, é crucial mencionar que esse período também coincide com a aprovação e implementação do RGPD nos países membros da UE. A entrada em vigor do RGPD em 25 de maio de 2018 trouxe mudanças substanciais na maneira como as organizações coletam, tratam e protegem os dados pessoais dos cidadãos da UE. Essa regulamentação teve um impacto profundo na pesquisa e nas práticas de proteção de dados.

Os estudos evidênciam que o projeto de lei de proteção de dados não atende às expetativas legislativas pelo fato de não reconhecer a necessidade de regular a análise de big data (Papakonstantinou & De Hert, 2020). Neste sentido e de forma a harmonizar interesses pessoais de utilizadores com interesses económicos no contexto do big data, alertam para a necessidade de um amplo debate que tenha como foco principal os setores que se beneficiam e aqueles que são prejudicados com as aplicações analíticas do big data. Wang (2017) propõe que os conceitos e princípios legais tradicionais do direito de banco de dados, do direito de proteção de dados, do direito de propriedade intelectual, do direito contratual e do direito internacional privado sejam interpretados e aplicados às especificidades do mercado de big data. Além disso, as melhores práticas e diretrizes também podem ser úteis para contribuir com o bem-estar dos cidadãos, bem como para o progresso socioeconômico. Para Xanthidis et al. (2020), há necessidade de encontrar um equilíbrio entre os múltiplos benefícios do desenvolvimento tecnológico e a proteção dos direitos humanos básicos, como a privacidade das informações pessoais. Neste contexto, os autores aconselham estudar o escopo e o nível de penetração da IoT nos países, e analisar os elementos qualitativos desta penetração e as suas perspetivas para os próximos anos. No ponto de vista Van

der Sloot (2018), o surgimento de tecnologias modernas dificulta a identificação de violações concretas e danos pessoais. Assim, propõe uma abordagem influenciada pela discussão filosófica sobre liberdade, com a atenção voltando para o princípio de 'não-dominação', ou seja, não fazer uso abusivo do poder e controle. Portanto, diante das reflexões feitas pelos autores, acredita-se que há a possibilidade de harmonizar interesses pessoais com interesses económicos quando realizadas às análises de big data, porém há um grande caminho a ser percorrido e, este, poderá começar por discutir um amplo regulatório para a proteção dos dados.

Van der Sloot (2015) enfatiza que é incerto se as reivindicações sobre o Big Data serão declaradas admissíveis sob a Convenção Europeia dos Direitos Humanos (ECHR) e, por outro lado, não está claro como o Tribunal Europeu dos Direitos Humanos (TEDH) poderia abordar esses dilemas de maneira satisfatória. Para este autor, os testes existentes de interesse público, interesse individual e equilíbrio de interesses não se aplicam bem ao Big Data devido à complexidade dos interesses em jogo. Assim sendo, propõe novos direcionamentos para os investigadores e cientistas da área, criar um quarto teste que inclua desenvolvimentos estruturais e sociais de forma que o TEDH possa abordar eficazmente questões de privacidade relacionadas ao Big Data e manter a relevância do Artigo 8 da ECHR. Xanthidis et al. (2020) aconselham investigadores estudarem o desenvolvimento da IoT e seu impacto na privacidade da informação, focando, especialmente, em aspetos específicos da economia emergente, que podem ter um impacto direto na sociedade e nas empresas locais. Os autores (Custers et al., 2018) propõem replicar o estudo feito por eles, visto que encontraram diferenças significativas entre os níveis de aplicação pelas diferentes autoridades de proteção de dados ao realizarem uma comparação entre oito estados-membros da UE relativo a proteção da privacidade e dos dados pessoais (privacidade informacional).

4.1. Limitações

Os resultados deste RSL devem ser considerados no contexto de suas limitações, dado que a amostra composta basicamente por países membros da UE pode ter restringido a amplitude do fenómeno estudado. Por exemplo, os estudos incluídos [artigos 3, 5, 6, 10, 11, 12,

[13, 14, 15, 16, 18, 22, 23, 24, 25, 26, e 28] retratam a questão da informação e a privacidade no contexto de big data como foco principal nos países membros da UE, o que não reflete a complexidade da problemática da informação e a privacidade no contexto do Big Data, dado seu alcance global. Apesar de a revisão ter incluído artigos de diferentes países/continentes, como Ásia, Estados Unidos da América, Emirados Árabes Unidos, Rússia e Canadá, ainda não representa uma perspetiva global sobre o fenómeno estudado. Podendo ser uma causa dessa limitação a definição da língua de pesquisa, que incidiu sobre o inglês e o português.

Os reduzidos estudos incluídos após a vigorarão da lei de RGPD não permitiram um detalhamento sobre como a literatura tem tratado a questão da informação e o direito à privacidade no contexto do Big Data. Entretanto, o processo de interpretação dos resultados, relativo as temáticas/setores com base maioritariamente na legislação, ampliou e fortaleceu a explicação compreensiva para os temas elaborados que representam a lei sobre os direitos humanos, nomeadamente a privacidade e a proteção dos dados no contexto de big data.

E, por fim, apresentamos lacunas na qualidade da apresentação dos estudos, alertando os pesquisadores, para além de identificar de forma clara o contexto, o objeto, os objetivos, a metodologia, os resultados e conclusão, adicionalmente as implicações.

4.2. Implicações

Este estudo permite apresentar sugestões para possíveis caminhos de investigações futuras. Com efeito, começam-se por salientar algumas referentes a novos estudos sobre os países. Como já relatado, foi detetado seis países. Assim, sugere-se a continuação da investigação de forma incluir novos países/continentes. Como podemos ver, o ocidente tem feito um grande progresso na definição e defesa da privacidade humana, dado as questões culturais, políticas e históricas. Mas ao mesmo tempo a privacidade é um direito universal e deve incidir sobre todos os seres humanos, por isso seria interessante a continuação de uma pesquisa focada em outras línguas e países, como no Médio-oriente, Ásia e América do Sul. Há uma necessidade de investigar como

as culturas, normas sociais e valores afetam a perceção da privacidade no contexto do Big Data.

Outro aspeto a ser explorado relaciona-se com estudos posteriormente a vigoração da Lei do RGPD de forma a ter uma visão recente sobre a aplicação da lei relativa a informação e a proteção da privacidade.

Igualmente, pode ainda ser alvo de investigação a ser considerado a inclusão de estudos empíricos qualitativos, visto que este tipo de estudo pode oferecer resultados e implicações futuras mais realista, já que a investigação se dá no contexto estudado.

E por fim, sugere-se a exploração mais aprofundada dos desafios práticos na implementação de diretrizes de privacidade, bem como a investigação das implicações das tecnologias emergentes na interseção entre informação e privacidade. A colaboração entre pesquisadores multidisciplinares é crucial para abordar essas lacunas e fornecer insights abrangentes sobre esse tópico em constante evolução.

Conclusão

É necessário refletir sobre a informação e a proteção da privacidade no contexto de big data. Os estudos analisados permitem tecer alguns recortes conclusivos acerca dos desafios que encontramos, no que toca a informação e à privacidade no contexto do Big Data e baseados em conceitos como regulamentação, saúde, ética, educação, direito da criança e política criminal. Em geral, os estudos apontam que as soluções de big data trazem benefícios que podem remodelar as sociedades globais. os negócios e a vida humana em geral (Xanthidis, Manolas & Xanthidou; Wang, 2017; Zharove & Vladimir, 2017). No entanto, todos os estudos mostram preocupações com a informação e o direito à privacidade e a possibilidade de Big Data lesar liberdades e direitos humanos, bem como garantias individuais. Em concreto, os estudos rogam para uma compreensão dos limites da lei de proteção de dados, atentando às novas realidades, às novas capacidades e riscos de um ambiente digital em ampla evolução (Vayena & Tasioulas, 2016), evidenciando que os princípios legais tradicionais do direito de banco de dados, do direito de proteção de dados, do direito de propriedade intelectual, do direito contratual e do direito internacional privado precisam ser interpretados e aplicados às especificidades do mercado de big data (Wang, 2017). Estes discursos evidenciam a preocupação com o fenómeno da regulamentação no sentido em que as soluções realizadas pela tecnologia de big data provocam reações indesejáveis (Neiva, 2019). Assim sendo, sublinham a necessidade de redefinir, adequar e reforçar a legislação de forma a delimitar o caráter sensível e peculiar dos dados individuais a serem coletados, armazenados e partilhados, nomeandomente as melhores práticas e diretrizes para contribuir com o bem-estar.

À medida que a tecnologia continua a evoluir e novos desafios emergem, há uma necessidade crescente de pesquisas interdisciplinares que explorem os limites éticos, legais, sociais e técnicos da interação entre informação, privacidade e Big Data. Além disso, considerando o escopo global da tecnologia e dos dados, seria benéfico expandir o foco da pesquisa para além da UE, Estados Unidos e outros países ocidentais,

incorporando perspetivas de outras regiões e suas abordagens únicas à privacidade e proteção de dados.

Enfrentamos desafios atuais urgentes relacionados à proteção de nossa privacidade em um cenário de constante evolução tecnológica. No entanto, também devemos nos preparar para um futuro em que a preservação da privacidade se tornará ainda mais complexa e vital. Isso não se trata apenas de buscar respostas práticas para as questões atuais do big data, mas também de explorar as implicações filosóficas mais profundas da privacidade da informação.

A busca por um equilíbrio entre a transparência e a proteção da privacidade nos levará a questionar não apenas as práticas atuais, mas também os valores fundamentais que sustentam nossa compreensão da liberdade e da autonomia individual numa sociedade digital. Em última análise, a proteção da privacidade da informação é uma jornada que vai além das medidas técnicas; é uma exploração filosófica das fronteiras da nossa autonomia em um mundo cada vez mais interconectado e monitorado. É nosso desafio e responsabilidade garantir que, à medida que avançamos para o futuro, a privacidade continue sendo um pilar central de nossa sociedade e consigamos estabelecer uma cooperação saudável entre todos as partes envolventes. Esse campo continuará sendo dinâmico e repleto de oportunidades para contribuições significativas.

REFERÊNCIAS BIBLIOGRÁFICAS

Andrew, J., & Baker, M. (2021). The general data protection regulation in the age of surveillance capitalism. *Journal of Business Ethics*, *168*(3), 565–578. https://discovery.ebsco.com/linkprocessor/plink?id=2d89a52a-c0af-3ac8-a2b3-45a6748daaa5

Anthony, B., Kamaludin, A., Romli, A., Raffei, A. F. M., Phon, D. N. A. L. E., Abdullah, A., & Ming, G. L. (2022). Blended Learning Adoption and Implementation in Higher Education: A Theoretical and Systematic Review. *Technology, Knowledge and Learning*, *27*(2), 531–578. https://doi.org/10.1007/s10758-020-09477-z

Araújo, C. (2003). Ciência da informação como ciência social. *Ciência Da Informação*, *32*(3), 21–27. https://doi.org/10.1590/s0100-19652003000300003

Araújo, C. (2009). Correntes teóricas da ciência da informação. *Ciência Da Informação*, *38*(3), 192–204.

Baumgaertner, B., & Floridi, L. (2016). Introduction: The Philosophy of Information. *Topoi*, *35*(1), 157–159. https://doi.org/10.1007/s11245-016-9370-7

Berger, P., & Luckmann, T. (2004). A CONSTRUÇÃO SOCIAL DA REALIDADE Tratado de Sociologia do Conhecimento. In *Doubleday & Company, Inc.* (Vol. 12, Issue 2). Editora Vozes Ltda. https://doi.org/10.5433/1980-511x.2017v12n2p316

Borges, L. C., & Machado, D. D. Q. (2019). O Círculo: um estudo observacional dos conflitos entre o desenvolvimento tecnológico e os limites da privacidade. *Revista Tecnologia e Sociedade*, *15*(38), 242–258. https://doi.org/10.3895/rts.v15n38.8500

Caglar, C. (2021). Children's Right to Privacy and Data Protection: Does the Article on Conditions Applicable to Child's Consent Under the GDPR Tackle the Challenges of the Digital Era or Create Further Confusion? *European Journal of Law & Technology*, *12*(2), 1–31. https://discovery.ebsco.com/linkprocessor/plink?id=12a21650-e6df-351a-94e8-

3bfed995c791

Capurro, R., & Hjorland, B. (2007). O conceito de informação. *Perspectivas Em Ciencia Da Informacao*, *12*(1), 148–207.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(2), 171–209. https://doi.org/10.1007/s11036-013-0489-0

Crawford, K., & Schultz, J. (2014). Big data and due process: Toward a framework to redress predictive privacy harms. *BCL Rev.*, *55*, 93.

Custers, B., Dechesne, F., Sears, A. M., Tani, T., & van der Hof, S. (2018). A comparison of data protection legislation and policies across the EU. *Computer Law & Security Review*, *34*(2), 234–243. https://discovery.ebsco.com/linkprocessor/plink?id=ee4597fb-35f7-3b83-9410-9ae07940ccde

De Magalhães, L. H., & Souza, R. R. (2019). SISTEMA DE RECUPERAÇÃO DA INFORMAÇÃO: uma abordagem baseada em ontologias. *PontodeAcesso*, *13*(2), 63. https://doi.org/10.9771/rpa.v13i2.28452

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, *65*(3), 122–135. https://doi.org/10.1108/LR-06-2015-0061

Donato, H., & Donato, M. (2019). Stages for undertaking a systematic review. *Acta Medica Portuguesa*, *32*(3), 227–235. https://doi.org/10.20344/amp.11923

dos Reis, S. M. G., Leite, A. C. A. B., Alvarenga, W. de A., Araújo, J. S., Zago, M. M. F., & Nascimento, L. C. (2017). Metassíntese sobre o homem como pai e cuidador de um filho hospitalizado. *Revista Latino-Americana de Enfermagem*, *25*. https://doi.org/10.1590/1518-8345.1850.2922

Drewer, D., & Miladinova, V. (2017). The BIG DATA Challenge: Impact and opportunity of large quantities of information under the Europol Regulation. *Computer Law & Security Review*, *33*(3), 298–308.

https://discovery.ebsco.com/linkprocessor/plink?id=4424f074-088d-315e-b3c9-68d90986606f

Erl, T., Khattak, W., & Buhler, P. (2016). Big Data Fundamentals Concepts, Drivers & Techniques. *Prentice Hall Service Technology Series from Thomas Erl*.

Filho, A., & Schwartz, G. (2016). "BIG DATA BIG PROBLEMA! PARADOXO ENTRE O DIREITO À PRIVACIDADE E O CRESCIMENTO SUSTENTÁVEL. *CONPEDI LAW REVIEW*, *2*(3), 311–331. https://doi.org/10.21902/clr.v2i3.314

Floridi, L. (1999). Information ethics: On the philosophical foundation of computer ethics. *Ethics and Information Technology 1:*, 37–56. https://doi.org/10.2139/ssrn.3844352

Floridi, L. (2002). What is the philosophy of information? *Metaphilosophy*, *33*(1–2), 123–145. https://doi.org/10.1111/1467-9973.00221

Floridi, L. (2003). From data to semantic information. *Entropy*, *5*(2), 125–145. https://doi.org/10.3390/e5020125

Floridi, L. (2005). Is Semantic Information Meaningful Data? *SSRN Electronic Journal*, *LXX*(2), 351–370. https://doi.org/10.2139/ssrn.3845324

Floridi, L. (2011). *The Philosophy of Information*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199232383.001.0001

Floridi, L. (2012). Big Data and Their Epistemological Challenge. *Philosophy & Technology*, *25*(4), 435–437. https://doi.org/10.1007/s13347-012-0093-4

Floridi, L. (2020). *Research: The tetralogy project*. https://www.philosophyofinformation.net/research/

Galvão, M. C. B., & Ricarte, I. L. M. (2019). Systematic literature review: concept, production and publication. *Logeion: Filosofia Da Informação*, *6*(1), 57–73.

Georgiadis, G., & Poels, G. (2022). Towards a privacy impact assessment methodology to support the requirements of the general data protection regulation in a big data analytics context: A systematic literature review. *Computer Law & Security*

Review, *44*(null), N.PAG-N.PAG. https://discovery.ebsco.com/linkprocessor/plink?id=31b4527b-4faf-3ad7-9f0e-43c60ee6c552

Gomez, J., Pinnick, T., & Soltani, A. (2009). *Know Privacy*. *March*, 44.

Helm, P. (2016). Group Privacy in Times of Big Data. A Literature Review. *Digital Culture & Society*, *2*(2), 137–152. https://discovery.ebsco.com/linkprocessor/plink?id=c1a313ae-327d-3b14-bab9-29a1ad4d3972

Higgins, J., & Green, S. (2011). Cochrane Handbook for Systematic Reviews of Interventions. *The Cochrane Collaboration*. www.cochrane-handbook.org

Kitchenham, B. A., & Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering (Software Engineering Group, Department of Computer Science, Keele …. *Technical Report EBSE 2007- 001. Keele University and Durham University Joint Report*, *January*.

Laney, D. (2001). 3-D Data Management: Controlling Data Volume, Velocity and Variety. *Application Delivery Strategies by META Group Inc.*, *949*.

Loh, S. (2019). *Volume , velocidade , variedade , veracidade e valor : como os 5 Vs do big data estão impactando as organizações e a sociedade*.

Mancini, M., & Sampaio, R. (2007). ESTUDOS DE REVISÃO SISTEMÁTICA: UM GUIA PARA SÍNTESE CRITERIOSA DA EVIDÊNCIA CIENTÍFICA. *Rev. Bras. Fisioter., São Carlos,11*(1), 83–89.

Mantelero, A. (2018). *AI and Big Data: a blueprint for a human rights, social and ethical impact assessment*. https://discovery.ebsco.com/linkprocessor/plink?id=2a066cf5-a4b8-306f-b41e-951024066593

Matsuno, K. (1996). Internalist stance and the physics of information. *BioSystems*, *38*(2–3), 111–118. https://doi.org/10.1016/0303-2647(95)01580-9

Minssen, T., Seitz, C., Aboy, M., & Corrales Compagnucci, M. (2020). The EU-US Privacy Shield Regime for Cross-Border Transfers of Personal Data under the GDPR: What are the legal challenges and how might these affect cloud-based technologies, big data, and AI in the medical sector? *European Pharmaceutical Law Review*, *4*(1), 34–50. https://discovery.ebsco.com/linkprocessor/plink?id=4224ce5c-aced-3f17-8fb9-6751b453c664

Montgomery, K. C., Chester, J., & Milosevic, T. (2017). Children's privacy in the big data era: Research opportunities. *Pediatrics*, *140*(5, Supp 2), S117–S121. https://discovery.ebsco.com/linkprocessor/plink?id=db1c75cd-1a4a-323f-b8eb-13444ee3e052

Mostert, M., Bredenoord, A. L., van der Sloot, B., & van Delden, J. J. M. (2018). *From privacy to data protection in the EU: implications for big data health research*. https://discovery.ebsco.com/linkprocessor/plink?id=23acec52-ecde-370f-8773-9160c6c3dd57

Nageshwaran, G., Harris, R. C., & Guerche-Seblain, C. El. (2021). *Review of the role of big data and digital technologies in controlling COVID-19 in Asia: Public health interest vs. privacy.* https://discovery.ebsco.com/linkprocessor/plink?id=a4f5420f-43d9-305a-b12d-b1baa45c1fcb

Neiva, L. (2020). O direito à privacidade no tempo do big data: narrativas profissionais na União Europeia. *Revista Tecnologia e Sociedade*, *16*(45), 1–20. https://doi.org/10.3895/rts.v16n45.11439

OCDE. (2002). Síntese Diretrizes da OCDE para a Proteção da Privacidade e dos Fluxos Transfronteiriços de Dados Pessoais. *ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÓMICOS*. www.oecd.org/bookshop/

Oliveira, M. F. de. (2011). Metodologia científica: um manual para a realização de pesquisas em Administração. *Metodologia Científica*, 1–73. http://medcontent.metapress.com/index/A65RM03P4874243N.pdf%5Cnhttps://books.google.com/books?id=zUDsAQAAQBAJ&pgis=1%5Cnhttp://materiaprima.p

ro.br/extensao/pesquisa/metodologia_pesquisa_cientifica.pdf

Papakonstantinou, V., & de Hert, P. (2020). *Big data analytics in electronic communications: a reality in need of granular regulation (even if this includes an interim period of no regulation at all)*. https://discovery.ebsco.com/linkprocessor/plink?id=2be9716b-054f-3d17-a145-169deafa008e

Park, Y. J., Sang, Y., Lee, H., & Jones - Jang, M. (2020). The ontology of digital asset after death: policy complexities, suggestions and critique of digital platforms. *Digital Policy, Regulation and Governance*, *ahead-of-p*. https://doi.org/10.1108/DPRG-04-2019-0030

Parlamento Europeu. (2016). Regulamento (UE) 2016/679. In *Jornal Oficial da União Europeia* (Vol. 2014, Issue 3, pp. 1–119).

Petersen, C. J. (2017). Big Data, Health Care, and International Human Rights Norms. *Asia Pacific Journal of Health Law & Ethics*, *11*(1), 1–22. https://discovery.ebsco.com/linkprocessor/plink?id=62407d25-c7d7-3f00-8eb7-2fefbe2fa1c8

Pilati, J. I., & de Olivo, M. V. C. (2014). Um novo olhar sobre o Direito à privacidade: Caso Snowden e Pós-modernidade jurídica. *Seqüência: Estudos Jurídicos e Políticos*, *35*(69), 281–300. https://doi.org/10.5007/2177-7055.2014v35n69p281

Polanyi, M. (1958). PERSONAL KNOWLEDGE Towards a Post-Critical Philosophy. In *Routledge & Kegan Paul Ltd*. Taylor & Francis Group.

Reidenberg, J. R., & Schaub, F. (2018). Achieving Big Data Privacy in Education. *Theory and Research in Education*, *16*(3), 263–279. https://discovery.ebsco.com/linkprocessor/plink?id=ca7d263c-c0d1-3495-b6c8-716d27cc7980

Rifkin, J. (2003). *The age of access* (pp. 1–4).

Ripoll, L., & Matos, J. C. M. (2021). Desinformação e informação semântica: a Filosofia

da Informação e o pensamento de Luciano Floridi na contribuição à confiabilidade informacional. *Em Questão*, *26*(2), 211–232. https://doi.org/10.19132/1808-5245262.211-232

Roever, L. (2017). Understanding systematic review studies. *Rev Soc Bras Clin Med*.

Rossetti, R., & Angeluci, A. (2021). Ética Algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação. *Galáxia (São Paulo)*, *46*, 1–18. https://doi.org/10.1590/1982-2553202150301

Roy, S. B., & Basak, M. (2013). Journal of documentation: A bibliometric study. *Library Philosophy and Practice*, *2013*(August 2013).

Saias, J. M. (2003). *Uma Metodologia para a construçao automática de Ontologias e sua aplicação em Sistemas de Recuperação de Informação*.

Saracevic, T. (2009). Information science. *Encyclopedia of Library and Information Sciences*, 2570–2586. https://doi.org/10.1081/E-ELIS3-120043704

Schrader, A. M. (1986). The domain of information science: Problems in conceptualization and in consensus-building. *Information Services and Use*, *6*(5), 169–205. https://doi.org/10.3233/ISU-1986-65-601

Schwab, K. (2017). *A quarta revolução industrial*. Levoir.

Shannon, C. E., & Weaver, W. (1949). The Mathematical Theory of Communication. *THE UNIVERSITY OF ILLINOIS PRESS*.

Shapiro, F. R. (1995). Brief Communication: Coinage of the Term Information Science. *Journal of the American Society for Information Science*, *46*(5), 321–397. http://doi.wiley.com/10.1002/%28SICI%291097-4571%28199506%2946%3A5%3C384%3A%3AAID-ASI8%3E3.0.CO%3B2-3%0Ahttps://search-proquest-com.dbgateway.nysed.gov/docview/216899838/146E9EF0544D45A3PQ/12?accountid=8012

Silva, S. P. da. (2019). Comunicação digital, economia de dados e a racionalização do

tempo: algoritmos, mercado e controle na era dos bits. *Revista Contracampo*, *38*(1), 157–169. https://doi.org/10.22409/contracampo.v0i0.27138

Smith, J. H., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly: Management Information Systems*, *35*(4), 989–1015. https://doi.org/10.2307/41409970

Steiner, C. M., Kickmeier-Rust, M. D., & Albert, D. (2016). LEA in Private: A Privacy and Data Protection Framework for a Learning Analytics Toolbox. *Journal of Learning Analytics*, *3*(1), 66–90. https://discovery.ebsco.com/linkprocessor/plink?id=212fb576-c770-3fcc-b8d2-c2e497759fe7

Strahilevitz, L. J. (2013). Toward a Positive Theory of Privacy Law. *Harvard Law Review*, *126*(7), 2010–2139. https://litigation-essentials.lexisnexis.com/webcd/app?action=DocumentDisplay&crawlid=1&doctype=cite&docid=126+Harv.+L.+Rev.+2010&srctype=smi&srcid=3B15&key=93e5586f13ebfb3c100f1e2e857c248f%5Cnhttp://www.harvardlawreview.org/media/pdf/vol126_strahilevitz.pd

The PRISMA Group. (2015). Principais itens para relatar Revisões sistemáticas e Meta-análises: A recomendação PRISMA. *Epidemiologia e Serviços de Saúde*, *24*(2), 335–342. https://doi.org/10.5123/s1679-49742015000200017

Truby, J., & Brown, R. (2021). *Human digital thought clones: the Holy Grail of artificial intelligence for big data*. https://discovery.ebsco.com/linkprocessor/plink?id=6bcbdd8a-d0e7-3c44-a89b-9689ccc251d3

Tzanou, M. (2017). *European Union regulation of transatlantic data transfers and online surveillance*. https://discovery.ebsco.com/linkprocessor/plink?id=19004459-755b-3185-8910-c109f389de81

Ursic, H., & Custers, B. (2016). Legal Barriers and Enablers to Big Data Reuse-A Critical Assessment of the Challenges for the EU Law. *European Data Protection Law*

Review (Lexxion), *2*(2).

Van der Sloot, B. (2015). *How to assess privacy violations in the age of Big Data? Analysing the three different tests developed by the ECtHR and adding for a fourth one*. https://discovery.ebsco.com/linkprocessor/plink?id=c013dd9e-33cf-3382-997d-2ac08df8576e

Van der Sloot, B. (2018). *A new approach to the right to privacy, or how the European Court of Human Rights embraced the non-domination principle*. https://discovery.ebsco.com/linkprocessor/plink?id=b83eb179-b78c-379a-ac9f-aafa552da07c

Vayena, E., & Tasioulas, J. (2016). *The dynamics of big data and human rights: the case of scientific research.* https://discovery.ebsco.com/linkprocessor/plink?id=a1e2b334-c5cd-3f14-aa5d-7b1a35cb5b70

Veldkamp, L., & Chung, C. (2019). Data and the Aggregate Economy. *Working Paper*, 1–36.

Wang, F. F. (2017). Big Data Regulatory Debates in the EU. *European Business Law Review*, *28*(4), 593–613. https://discovery.ebsco.com/linkprocessor/plink?id=00edadeb-4590-36e5-913b-692c7ed07ab4

Westin, A. F. (2003). Social and Political Dimensions of Forest Certification. *The Society for the Psychological Study of Social Issues*, *59*(2), 431–453.

Wiener, N. (1961). Cybernetics or control and communication in the animal and the machine. In *MIT Press paperback edition*. The Massachusetts Intitute of Technology.

Wolf, L. E. (2018). Risks and Legal Protections in the World of Big-Data. *Asia Pacific Journal of Health Law & Ethics*, *11*(2), 1–15.

Wormell, I. (1998). Informetria: explorando bases de dados como instrumentos de

análise. *Ciência Da Informação*, *27*(2), 210–216. https://doi.org/10.1590/s0100-19651998000200016

Xanthidis, D., Manolas, C., Xanthidou, O. K., & Wang, H.-I. (2020). Information privacy and emerging technologies in the UAE: Current state and research directions. *International Journal of Technoethics*, *11*(2), 1–17. https://discovery.ebsco.com/linkprocessor/plink?id=df165cb1-5780-310e-9142-8fa4db7d8bde

Zharova, A. K., & Elin, V. M. (2017). The use of Big Data: A Russian perspective of personal data security. *Computer Law & Security Review: The International Journal of Technology Law and Practice*, *33*(4), 482–501. https://discovery.ebsco.com/linkprocessor/plink?id=adab7655-f09d-3fae-a98b-79e1fcf6b395

Zwitter, A. (2014). Big Data ethics. *Big Data and Society*, *1*(2), 1–6. https://doi.org/10.1177/2053951714559253

APÊNDICE A: Estudos rejeitados

Artigos	Autor	Ano	CQ 1	CQ 2	CQ 3	CQ 4	CQ 5	CQ 6	Total	Justificação
WHAT PRIVACY IS FOR	Cohen, Julie E	2013							0	CE 1. O estudo não possui como tema principal a informação e o direito a privacidade no contexto do Big Data
Protecting and Utilizing Health and Medical Big Data: Policy Perspectives from Korea.	Lee, Dongjin ; Park, Mijeong ; Chang, Seungwon ; Ko, Haksoo	2019	0	1	1	1	0	1	4	Abaixo da qualidade
BIG DATA AND PUBLIC HEALTH: NAVIGATING PRIVACY LAWS TO MAXIMIZE POTENTIAL	Thorpe, Jane Hyatt ; Gray, Elizabeth Alexandra	2015	1	1	0,5	0,5	0	1	4	Abaixo da qualidade
Algorithms, Future and Digital Rights: Some Reflections	Martínez-Ávila, Daniel	2018							0	CE 2. O estudo não aborda legislações existentes que respondem ao direito a privacidade
Redefining and Renewing Humanism in the Digital Age [Opinion].	Messner, Dirk	2020							0	CE 4. O estudo é uma publicação do género: ensaio, editorial, material de conferência, livro, relatório, dissertações e teses
What are the threats and potentials of big data for qualitative research?	Mills, Kathy A	2017							0	CE 1. O estudo não possui como tema principal a informação e o direito a privacidade no contexto do Big Data
Next generation privacy	Omotubora, Adekemi ; Basu, Subhajit	2020	0	1	0,5	0,5	0,5	0,5	3	Abaixo da qualidade
Big brother' can watch us.	Power, Daniel J	2016	1	1	0,5	0,5	0,5	0,5	4	Abaixo da qualidade
A critique of the regulation of data science in healthcare research in the European Union.	Rumbold, John M M ; Pierscionek, Barbara K	2017	1	0,5	0,5	0,5	0,5	0,5	3,5	Abaixo da qualidade
Big Data and the Illusion of Choice: Comparing the Evolution of India's Aadhaar and China's Social Credit System as Technosocial Discourses	Shahin, Saif ; Zheng, Pei	2020							0	CE 3. O estudo recorre a métodos de pesquisa empírica
Privacy Preserving, Protection of Personal Data, and Big Data: a Review of the Colombia Case.	Silva, Jesus ; Solano, Darwin ; Fernandez, Claudia ; Romero, Ligia ; Villa, Jesus Vargas	2020							0	CE 3. O estudo recorre a métodos de pesquisa empírica
Mismanagement of personally identifiable information and the reaction of interested parties to safeguarding privacy in South Korea.	Song, Dong Hyun ; Son, Chang Yong	2017							0	CE 3. O estudo recorre a métodos de pesquisa empírica

Risk regulation of big data: has the time arrived for a paradigm shift in EU data protection law?	Spina, Alessandro	2014							0	CE 3. O estudo recorre a métodos de pesquisa empírica
Privacy preserving parallel clustering based anonymization for big data using MapReduce framework.	Usha Lawrance, Josephine ; Jesu Vedha Nayahi Jesudhasan	2021							0	CE 3. O estudo recorre a métodos de pesquisa empírica
Procedural law for the data-driven society.	Van der Sloot, Bart ; Van Schendel, Sascha	2021							0	CE 3. O estudo recorre a métodos de pesquisa empírica
Analysis of Legal Issues of Personal Information Protection in the Field of Big Data.	Wang, Chuyun ; Feifei Guo ; Mengxuan Ji	2022							0	Artigo retratado
A COVID-19 Auxiliary Diagnosis Based on Federated Learning and Blockchain.	Wang, Ziyu ; Cai, Lei ; Zhang, Xuewu ; Choi, Chang ; Su, Xin	2022							0	CE 1. O estudo não possui como tema principal a informação e o direito a privacidade no contexto do Big Data
A Protection Model of Citizen Personal Information Administrative Law Based on BD Analysis and Edge Computing	Yang, Fang	2022	0	0,5	0,5	1	0,5	0,5	3	Abaixo da qualidade
EDUCATIONAL PRIVACY IN THE ONLINE CLASSROOM: FERPA, MOOCS, AND THE BIG DATA CONUNDRUM.	Young, Elise	2014							0	CE 3. O estudo recorre a métodos de pesquisa empírica
International Law Protection of Cross-Border Transmission of Personal Information Based on Cloud Computing and Big Data	Ziyi, Xu	2022	0	0,5	0,5	0,5	0,5	0,5	2,5	Abaixo da qualidade
Legal Governance of Brain Data Derived from Artificial Intelligence	Ahluwalia, Mahika	2021	0	0	0	0	0,5	0	0,5	Abaixo da qualidade
GINA, Big Data, and the Future of Employee Privacy	Areheart, Bradley A ; Roberts, Jessica L	2019							0	CE 4. O estudo é uma publicação do género: ensaio, editorial, material de conferência, livro, relatório, dissertações e teses
I Agree to Disagree": Comparative Ethical and Legal Analysis of Big Data and Genomics for Privacy, Consent, and Ownership	Belani, Seema ; Tiarks, Georgina C ; Mookerjee, Neil ; Rajput, Vijay	2021	0	0,5	0	0,5	1	0	2	Abaixo da qualidade
Solutions to Big Data Privacy and Security Challenges Associated With COVID-19 Surveillance Systems	Bentotahewa, Vibhushinie ; Hewage, Chaminda ; Williams, Jason	2021	1	0,5	0,5	0,5	1	0	3,5	Abaixo da qualidade
Quis custodiet ipsos custodes? Security, big data and secrecy	Broeders, Dennis	2017							0	CE 4. O estudo é uma publicação do género: ensaio, editorial, material de conferência, livro,

										relatório, dissertações e teses
The chilling effects of algorithmic profiling: mapping the issues	Buchi, Moritz ; Fosch-Villaronga, Eduard ; Lutz, Christoph ; Tamo-Larrieux, Aurelia ; Velidi, Shruthi ; Viljoen, Salome	2020							0	CE 1. O estudo não possui como tema principal a informação e o direito a privacidade no contexto do Big Data
Data mining and automated prediction: A pedagogical primer for classroom discussion	Callanan, Gerard A ; David F. Perri ; Sandra M. Tomkowicz	2018	0	0,5	0,5	0	0,5	0,5	2	Abaixo da qualidade
The Economic Impact of the European Reform of Data Protection	Ciriani, Stéphane	2015	1	1	0	0	0,5	0,5	3	Abaixo da qualidade
Data protection in attention markets: protecting privacy through competition?	Colangelo, Giuseppe ; Maggiolino, Mariateresa	2017	1	0,5	0	0	0,5	0,5	2,5	Abaixo da qualidade
Villain or guardian? 'The smart toy is watching you now … .'	Collingwood, Lisa	2021	0	0,5	0	0	0,5	0,5	1,5	Abaixo da qualidade
Corporate Avatars and the Erosion of the Populist Fourth Amendment.	Cover, Avidan Y	2015							0	CE 4. O estudo é uma publicação do género: ensaio, editorial, material de conferência, livro, relatório, dissertações e teses
Access to big data as a remedy in big tech	Dadson, Nick and Snoddy, Iain and White, Joshua	2021	0	0,5	0	0,5	0,5	0,5	2	Abaixo da qualidade
Where Copyright Meets Privacy in the Big Data Era: Access to and Control Over User Data in Agriculture and the Role of Copyright.	Dagne, Tesh W	2022							0	CE 4. O estudo é uma publicação do género: ensaio, editorial, material de conferência, livro, relatório, dissertações e teses
EU merger control and big data	de Peyer, Ben	2017	1	1	0	0,5	0,5	0,5	3,5	Abaixo da qualidade
Surveillance, Big Data Analytics and the Death of Privacy	Doughty, Howard A	2014	0	0,5	0,5	0,5	0	0	1,5	Abaixo da qualidade
Privacy in public spaces: what expectations of privacy do we have in social media intelligence?	Edwards, Lilian ; Urquhart, Lachlan	2016	1	0,5	0,5	0,5	1	0,5	4	Abaixo da qualidade
THE RIGHT TO BENEFIT FROM BIG DATA AS A PUBLIC RESOURCE.	FAN, MARY D	2021							0	CE 4. O estudo é uma publicação do género: ensaio, editorial, material de conferência, livro, relatório, dissertações e teses
Ethics in Health Informatics.	Goodman, Kenneth W	2020							0	CE 4. O estudo é uma publicação do género: ensaio, editorial, material de conferência, livro, relatório, dissertações e teses

Big Data e proteção do direito à privacidade no contexto da sociedade da informação	Guerra Martins, Marcelo ; Ribeiro Gomes Jorgetto, Leonardo Felipe de Melo ; Arantes Sutti, Alessandra Cristina	2019	0	0,5	0,5	0,5	0	0	1,5	Abaixo da qualidade
Towards an understanding of privacy management architecture in big data: An experimental research.	Hajli, Nick ; Shirazi, Farid ; Tajvidi, Mina ; Huda, Nurul	2020							0	CE 3. O estudo recorre a métodos de pesquisa empírica
Ethical AI and Big Data in Times of Pandemic.	Hickok, Merve	2020	0	1	0,5	1	0,5	0,5	3,5	Abaixo da qualidade
Patient Privacy in the Era of Big Data.	Kayaalp, Mehmet	2018	1	0,5	0,5	0	0,5	0,5	3	Abaixo da qualidade
Digital markets, data, and privacy: competition law, consumer law and data protection	Kerber, Wolfgang	2016	1	0,5	0,5	0,5	1	0,5	4	Abaixo da qualidade
Data analytics and consumer profiling: finding appropriate privacy principles for discovered data	King, Nancy J ; Forder, Jay	2016	1	0,5	0,5	0,5	0,5	0,5	3,5	Abaixo da qualidade
A Survey on Big Data Market: Pricing, Trading and Protection	Liang, F ; Yu, W ; An, D ; Yang, Q ; Fu, X ; Zhao, W	2018	1	0,5	0,5	0	0,5	0,5	3	Abaixo da qualidade
Consumer Valuation of Personal Information in the Age of Big Data.	Lim, Sesil ; Woo, JongRoul ; Lee, Jongsu ; Huh, Sung-Yoon	2018							0	CE 3. O estudo recorre a métodos de pesquisa empírica
Impact of Social Media Behavior on Privacy Information Security Based on Analytic Hierarchy Process	Liu, Yuxuan ; Tse, Woon Kwan ; Kwok, Pui Yu ; Chiu, Yu Hin	2022							0	CE 1. O estudo não possui como tema principal a informação e o direito a privacidade no contexto do Big Data
Two Concepts of Group Privacy.	Loi, Michele and Christen, Markus	2020							0	CE 1. O estudo não possui como tema principal a informação e o direito a privacidade no contexto do Big Data
Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection.	Mantelero, Alessandro	2016	1	0,5	0	0	0,5	0,5	2,5	Abaixo da qualidade
Perspectivas jurídicas da relação entre big data e proteção de dados	Marks Szinvelski, MÃ¡rtin and Silva Arceno, Taynara and Baratieri Francisco, Lucas	2019	0	0,5	0,5	0,5	0,5	0,5	2,5	Abaixo da qualidade
Consumers' privacy choices in the era of big data.	Dengler, Sebastian ; Jens Prüfer	2021							0	CE 3. O estudo recorre a métodos de pesquisa empírica
Privacy, security and data protection in smart cities: a critical EU law perspective	Edwards, Lilian	2016							0	CE 3. O estudo recorre a métodos de pesquisa empírica

APÊNDICE B: Estudos aceitos

ID	Artigos	Autor	Ano	CQ 1	CQ 2	CQ 3	CQ 4	CQ 5	CQ 6	Total
1	Review of the role of big data and digital technologies in controlling COVID-19 in Asia: Public health interest vs. privacy.	Nageshwaran, Gopinath ; Harris, Rebecca C ; Guerche-Seblain, Clotilde El	2021	1	1	1	0,5	1	0	4,5
2	TOWARD A POSITIVE THEORY OF PRIVACY LA	Strahilevitz, Lior Jacob	2013	1	1	1	1	0	1	5
3	The EU-US Privacy Shield Regime for Cross-Border Transfers of Personal Data under the GDPR: What are the legal challenges and how might these affect cloud-based technologies, big data, and AI in the medical sector?	Minssen, Timo ; Seitz, Claudia ; Aboy, Mateo ; Corrales Compagnucci, Marcelo	2020	0,5	1	0,5	1	0,5	1	4,5
4	Children's privacy in the big data era: Research opportunities.	Montgomery, Kathryn C ; Chester, Jeff ; Milosevic, Tijana	2017	1	1	0,5	1	1	1	5,5
5	From privacy to data protection in the EU: implications for big data health research	Mostert, Menno ; Annelien L. Bredenoord ; Bart van der Slootb ; Johannes J.M. van Delden	2018	1	1	0,5	0,5	1	1	5
6	Big data analytics in electronic communications: a reality in need of granular regulation (even if this includes an interim period of no regulation at all)	Papakonstantinou, Vagelis ; de Hert, Paul	2020	0,5	0,5	1	1	1	1	5
7	The ontology of digital asset after death: policy complexities, suggestions and critique of digital platforms	Park, Yong Jin ; Sang, Yoonmo ; Lee, Hoon ; Jones-Jang, S Mo	2020	0,5	1	0,5	1	0,5	1	4,5
8	Big Data, Health Care, and International Human Rights Norms.	Petersen, Carole J	2017	0	1	1	1	0,5	1	4,5
9	Achieving Big Data Privacy in Education	Reidenberg, Joel R ; Schaub, Florian	2018	1	1	0,5	0,5	1	0,5	4,5
10	LEA in Private: A Privacy and Data Protection Framework for a Learning Analytics Toolbox	Steiner, Christina M ; Michael D. Kickmeier-Rust ; Dietrich Albert	2016	1	1	1	0,5	1	1	5,5
11	Human digital thought clones: the Holy Grail of artificial intelligence for big data	Truby, Jon ; Brown, Rafael	2021	1	1	1	1	1	0,5	5,5
12	European Union regulation of transatlantic data transfers and online surveillance	Tzanou, Maria	2017	1	1	1	0,5	0,5	1	5
13	Legal barriers and enablers to big data reuse: a critical assessment of the challenges for the EU law	Ursic, Helena and Custers, Bart	2016	0,5	1	1	0,5	0,5	1	4,5
14	How to assess privacy violations in the age of Big Data? Analysing the three different tests developed by the ECtHR and adding for a fourth one.	Van der Sloot, Bart	2015	1	1	1	0,5	1	0,5	5
15	A new approach to the right to privacy, or how the European Court of Human Rights embraced the non-domination principle	Van der Sloot, Bart	2018	1	0,5	1	1	0,5	0,5	4,5

16	The dynamics of big data and human rights: the case of scientific research.	Vayena, Effy ; Tasioulas, John	2016	1	0,5	0,5	1	1	0,5	4,5
17	BIG DATA AND DUE PROCESS: TOWARD A FRAMEWORK TO REDRESS PREDICTIVE PRIVACY HARMS.	Crawford, Kate ; Jason Schultz	2014	1	1	1	1	1	0,5	5,5
18	Big Data Regulatory Debates in the EU.	Wang, Faye Fangfei	2017	0	1	1	1	1	1	5
19	Risks and Legal Protections in the World of Big-Data.	Wolf, Leslie E	2018	0	1	0,5	1	1	1	4,5
20	Information privacy and emerging technologies in the UAE: Current state and research directions.	Xanthidis, Dimitrios ; Manolas, Christos ; Xanthidou, Ourania Koutzampasopoulou ; Wang, Han-I	2020	0	1	0,5	1	1	1	4,5
21	The use of Big Data: A Russian perspective of personal data security	Zharova, Anna Konstantinovna ; Vladimir Mikhailovich Elin	2017	1	1	0,5	1	0,5	1	5
22	The general data protection regulation in the age of surveillance capitalism	Andrew, Jane ; Max Baker	2021	1	0,5	0,5	1	0,5	1	4,5
23	Children's Right to Privacy and Data Protection: Does the Article on Conditions Applicable to Child's Consent Under the GDPR Tackle the Challenges of the Digital Era or Create Further Confusion?	Caglar, Cansu	2021	1	1	1	1	1	1	6
24	A comparison of data protection legislation and policies across the EU.	Bart Custers ; Francien Dechesne ; Alan M. Sears ; Tommaso Tani ; Simone van der Hof	2018	1	1	1	0,5	1	0,5	5
25	The BIG DATA Challenge: Impact and opportunity of large quantities of information under the Europol Regulation	Drewer, Daniel ; Miladinova, Vesela	2017	1	1	1	0,5	1	0	4,5
26	Towards a privacy impact assessment methodology to support the requirements of the general data protection regulation in a big data analytics context: A systematic literature review	Georgiadis, Georgios ; Poels, Geert	2022	1	1	1	1	1	1	6
27	Group Privacy in Times of Big Data. A Literature Review.	Helm, Paula	2016	1	0,5	0,5	0,5	1	1	4,5
28	AI and Big Data: a blueprint for a human rights, social and ethical impact assessment	Mantelero, Alessandro	2018	1	0,5	0,5	0,5	1	1	4,5