



**UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS HUMANAS
DEPARTAMENTO DE FILOSOFIA**

ELENA MARTINS CAVALCANTI GONÇALVES

***Big Data* e a Última Revolução Informacional: uma análise
filosófico-interdisciplinar a respeito dos desafios éticos da coleta massiva de dados**

Brasília

2024

Elena Martins Cavalcanti Gonçalves

***Big Data* e a Última Revolução Informacional: uma análise
filosófico-interdisciplinar a respeito dos desafios éticos da coleta massiva de dados**

Monografia apresentada a título de trabalho de conclusão de curso como requisito parcial para obtenção do grau de Bacharel ao curso de Graduação em Filosofia da Universidade de Brasília.

Orientador: Prof. Dr. André Leclerc.

Brasília

2024

Elena Martins Cavalcanti Gonçalves

***Big Data* e a Última Revolução Informacional: uma análise
filosófico-interdisciplinar a respeito dos desafios éticos da coleta massiva de dados**

Monografia apresentada a título de trabalho de conclusão de curso como requisito parcial para obtenção do grau de Bacharel ao curso de Graduação em Filosofia da Universidade de Brasília.

Brasília, ____ de _____ de _____.

Banca Examinadora

Prof. Dr. André Leclerc
Universidade de Brasília / UnB

Prof.^a Dr.^a Maria Cecília Pedreira de Almeida
Universidade de Brasília / UnB

Prof.^a Dr.^a Mariana Claudia Broens
Universidade Estadual Paulista / Unesp

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

Mb Martins Cavalcanti Gonçalves, Elena
Big Data e a Última Revolução Informacional: uma análise
filosófico-interdisciplinar a respeito dos desafios éticos
da coleta massiva de dados / Elena Martins Cavalcanti
Gonçalves; orientador André Leclerc. -- Brasília, 2024.
62 p.

Monografia (Graduação - Filosofia) -- Universidade de
Brasília, 2024.

1. Filosofia da Informação. 2. Ética. 3. Big data. 4.
Privacidade de dados. 5. Responsabilização. I. Leclerc,
André, orient. II. Título.

*Ao meu pai Antônio (in memoriam).
Obrigada por tudo o que você me ensinou.*

AGRADECIMENTOS

Gostaria de agradecer primeiramente à minha mãezinha que me acompanhou e me manteve em pé nesses momentos tão difíceis que eu vivi durante todo o tempo que eu passei produzindo esse TCC. Eu te amo muito.

Um imenso obrigada ao orientador, professor André Leclerc por toda a paciência que o senhor teve durante esse tempo e pelo apoio dado para que eu pudesse continuar acreditando no meu potencial. Agradeço também pelo senhor ter me introduzido as professoras Mariana Broens e Maria Eunice Gonzalez, duas pensadoras que eu admirei muito desde o início por toda a riqueza do conteúdo que produziram a respeito desse tema tão vasto e importante.

Aos meus amigos loungers que me acompanharam, me escutaram, me fizeram rir e me ajudaram a esquecer de todos os meus problemas por pelo menos um instante. Um abraço imenso em especial para a Bia, Ana e Anabi que sempre estiveram do meu lado durante anos, eu amo vocês demais.

Aos meus amigos virtuais provenientes de lugares que não precisam ser citados. Obrigada por todo apoio que tive em momentos difíceis e as trocas de ideias que vocês me proporcionaram. Mira e Alice, acho que não preciso dizer muito, só que eu amo vocês.

Agradeço também a todos que colaboraram com o meu consumo de café expresso da Biblioteca Central via pix (em ordem alfabética): Bruno “Mamães” Manhães, Felipe Freitas, Pedro Ramalho e Tiago Guedes.

Por fim, agradeço obviamente a mim por não ter desistido de tudo e tacadado fogo nesse trabalho.

RESUMO

A presente monografia tem como objetivo realizar uma análise ética dos usos e consequências da mineração de dados, especialmente no que diz respeito ao *big data*. No primeiro capítulo, são introduzidas reflexões sobre os tipos de ética possíveis no campo da filosofia da informação. Em seguida, no segundo capítulo, discorre-se sobre a discriminação e responsabilização envolvendo inteligência artificial e algoritmos, abordando temas como a impossibilidade de atribuição de responsabilidade à robôs e a moderação de redes sociais. Por último, no terceiro capítulo, é apresentado o tema da privacidade de dados com uma discussão centrada principalmente na coleta massiva de dados, trazendo reflexões sobre o direito à privacidade e as questões de poder e lucro relacionadas ao seu domínio.

Palavras chave: filosofia da informação; ética; *big data*; responsabilização; privacidade de dados.

ABSTRACT

The aim of this work is to carry out an ethical analysis of the uses and consequences of data mining, especially with regard to big data. The first chapter introduces reflections on the types of ethics possible in the field of philosophy of information. The second chapter then discusses discrimination and accountability involving artificial intelligence and algorithms, addressing issues such as the impossibility of assigning responsibility to robots and the moderation of social networks. Finally, in the third chapter, the topic of data privacy is presented with a discussion centered mainly on the massive collection of data, bringing reflections on the right to privacy and the issues of power and profit related to its domain.

Keywords: philosophy of information; ethics; big data; accountability; data privacy.

SUMÁRIO

INTRODUÇÃO	9
1 QUE TIPO DE ÉTICA?	14
1.1 Introdução à uma tecnoética	14
1.2 A ética computacional e de dados	15
2 DISCRIMINAÇÃO E RESPONSABILIZAÇÃO	19
2.1 Inteligência artificial e algoritmos	19
2.2 O chatbot Tay e atribuição de responsabilidade	22
2.3 Redes sociais e moderação de conteúdo	28
3 PRIVACIDADE DE DADOS	33
3.1 Aprendizado de máquina e big data	33
3.2 O direito à privacidade de dados	38
3.3 O lucro da análise de dados	46
CONSIDERAÇÕES FINAIS	55
REFERÊNCIAS BIBLIOGRÁFICAS	58

INTRODUÇÃO

A filosofia da informação é uma área da filosofia cujo debate tem crescido exponencialmente no último século devido às transformações tecnológicas. Luciano Floridi, filósofo da Universidade de Oxford, é um dos pioneiros no assunto. Ele defende que a filosofia da informação pode ser considerada uma disciplina madura porque, de acordo com ele,

[...] (a) representa um campo autônomo (tópicos exclusivos); (b) oferece uma abordagem inovadora para tópicos filosóficos tradicionais e novos (metodologias originais); e (c) pode estar ao lado de outros ramos da filosofia, oferecendo o tratamento sistemático dos fundamentos conceituais do mundo da informação e da sociedade da informação (novas teorias).¹ (FLORIDI, 2011, p. 2, tradução nossa)

Ao falar de filosofia da informação, é fundamental resgatar o percurso histórico levado pelo campo até chegar aos dias de hoje.

Ao longo do século XX, diversos autores (SIMON, MCCARTHY e HAYES, PAGELS, BURKHOLDER apud FLORIDI, 2011, p. 2) apontavam para uma mudança macroscópica não apenas na ciência, como também na filosofia, transformada através dos avanços notórios das tecnologias e ciências da informação e da computação nos últimos anos. No entanto, de acordo com Floridi (2011, p. 3), muitos desses autores estavam equivocados quanto à natureza dessa evolução, assim como também subestimaram as dificuldades da aceitação do novo paradigma de uma filosofia da informação.

Desde o início das publicações dos seminários de Turing, na década de 30, diversas transformações ocorreram em áreas como a “cibernética, a teoria da informação, a IA [inteligência artificial], a teoria dos sistemas, a ciência da computação, a teoria da complexidade e a TIC [tecnologia da informação e computação]”² (FLORIDI, 2011, p. 3, tradução nossa). Com isso, essas áreas conseguiram atrair uma atenção “significativa, porém

¹ “[...] (a) it represents an autonomous field (unique topics); (b) it provides an innovative approach to both traditional and new philosophical topics (original methodologies); and (c) it can stand beside other branches of philosophy, offering the systematic treatment of the conceptual foundations of the world of information and of the information society (new theories).”

² “[...] cybernetics, information theory, AI, system theory, computer science, complexity theory, and ICT”

esporádica” da comunidade filosófica, especialmente em torno da IA, que criou um terreno frutífero para a exploração da área da filosofia da mente.

Floridi (2011, p. 3-4) considera que a IA “agiu como um cavalo de Troia, introduzindo um paradigma computacional/informacional mais abrangente na cidadela filosófica”³ (tradução nossa). No entanto, até meados dos anos 80, a filosofia da informação ainda era prematura, e percebida como “transdisciplinar em vez de interdisciplinar”⁴ (tradução nossa). Essa inevitável interdisciplinaridade é, até hoje, um dos maiores dificultadores para o reconhecimento da sua importância. De acordo com o autor, muitos filósofos consideram que grande parte dos tópicos discutidos pela filosofia da informação são importantes apenas para pesquisadores de outras áreas, como “mídias de massa, estudos culturais, ciência da computação ou sociologia, para citar alguns exemplos”⁵ (tradução nossa). Até o recente desenvolvimento de uma sociedade da informação, a proposta de uma filosofia da informação era vista como estando em uma “encruzilhada de assuntos técnicos, questões teóricas, problemas aplicados e análises conceituais para ser a área de especialização de alguém”⁶ (tradução nossa).

Além disso, a filosofia dominante/convenção (*mainstream*) é tradicionalmente conservadora (FLORIDI, 2011, p. 4, tradução nossa), uma vez que os padrões e valores da filosofia, por serem em grande parte menos firmes e claros que na ciência, são mais difíceis de serem desafiados. Não obstante, esse é o contexto em que “uma posição culturalmente dominante é frequentemente obtida às custas de abordagens inovadoras ou não convencionais” (FLORIDI, 2011, p. 4, tradução nossa), o que resulta em pensadores de áreas inovadoras serem deixados à margem do grande cânone filosófico.

“É certo que a virada computacional afetou a ciência muito mais rapidamente. Isso explica por que alguns cientistas com mentalidade filosófica estavam entre os primeiros a perceber o surgimento de um novo paradigma”⁷ (FLORIDI, 2011, p. 4, tradução nossa). Só a

³ “[...] *AI acted as a Trojan horse, introducing a more encompassing computational/ informational paradigm into the philosophical citadel.*”

⁴ “[...] *transdisciplinary rather than interdisciplinary.*”

⁵ “[...] *mass media, cultural studies, computer science or sociology departments, to mention a few examples.*”

⁶ “[...] *crossroads of technical matters, theoretical issues, applied problems, and conceptual analyses to be anyone's own area of specialization.*”

⁷ “*Admittedly, the computational turn affected science much more rapidly. This explains why some philosophically minded scientists were among the first to perceive the emergence of a new paradigm.*”

partir dos anos 80 a filosofia da informação pôde se tornar um fenômeno mais massificadamente amplo em diversas ciências e contextos sociais, permitindo sua evolução.

Hoje, a sociedade humana alcançou um estágio de seu desenvolvimento onde as questões relativas à “criação, dinâmica, gestão e utilização da informação e dos recursos computacionais são absolutamente vitais”⁸ (FLORIDI, 2011, p. 4, tradução nossa). A sociedade da informação surge como resultado do avanço tecnológico mais rápido que já ocorreu em toda história. O papel cultural do computador como símbolo tecnológico definidor do novo milênio é muito mais influente do que “os moinhos na Idade Média, os relógios mecânicos no século XVII, o tear ou a máquina a vapor na era da Revolução Industrial”⁹ (FLORIDI, 2011, p. 5, tradução nossa).

Foi necessário uma terceira revolução relacionada ao computador (a Internet), uma geração totalmente nova de estudantes, professores e pesquisadores alfabetizados em informática, uma mudança substancial na estrutura da sociedade, uma transformação radical na sensibilidade cultural e intelectual e um senso de crise generalizado nos círculos filosóficos de várias orientações para que o novo paradigma informacional emergisse.¹⁰ (FLORIDI, 2011, p. 5, tradução nossa)

Floridi é um grande defensor da teoria de que a sociedade atualmente está vivendo a “quarta revolução da autocompreensão humana”, em que ocorre uma reavaliação do papel da humanidade no universo e da natureza do ser humano. Ao falar em quarta, ele pressupõe outras três revoluções que já ocorreram: as copernicana, darwiniana e freudiana. As quatro revoluções podem ser mais bem compreendidas no seguinte fragmento:

Não somos imóveis, no centro do universo (Copérnico); não somos naturalmente distintos e diferentes do resto do mundo animal (Darwin); estamos longe de sermos totalmente transparentes para nós mesmos (Freud). Agora, aos poucos, estamos aceitando a ideia de que podemos ser organismos informacionais [inforgs] entre

⁸ “[...] *creation, dynamics, management and utilization of information and computational resources are absolutely vital.*”

⁹ “[...] *mills in the Middle Ages, mechanical clocks in the seventeenth century, and the loom or the steam engine in the age of the Industrial Revolution.*”

¹⁰ “*It has required a third computer-related revolution (the Internet), a whole new generation of computer-literate students, teachers, and researchers, a substantial change in the fabric of society, a radical transformation in cultural and intellectual sensibility, and a widespread sense of crisis in philosophical circles of various orientations, for the new informational paradigm to emerge.*”

muitos agentes (Turing), inforgs que não são tão drasticamente diferentes de artefatos inteligentes e projetados, mas que compartilham com eles um ambiente global que, em última análise, é feito de informações, a infosfera. A revolução da informação não se trata de nos estendermos, mas de reinterpretarmos quem somos.¹¹ (FLORIDI, 2011, p. 5, tradução nossa)

Principalmente ao longo da última década, a capacidade das TICs em criar novas realidades e promover uma interpretação informacional de cada aspecto do mundo e daqueles que o vivem tem transformado e modificado cada vez mais a vivência humana. O desenvolvimento de uma interface progressivamente menos visível torna cada vez mais difícil distinguir o limite entre “o aqui (analógico, baseado em carbono, off-line) e o lá (digital, baseado em silício, on-line)”¹² (FLORIDI, 2014, p. 43). Curiosamente, isso favorece muito mais o “lá”, uma vez que o mundo digital tem se espalhado cada vez mais no mundo analógico e se fundido com ele. Esse fenômeno, comumente chamado na computação de “internet das coisas”, está se tornando, ou deverá se tornar, o “próximo estágio no desenvolvimento da era da informação”¹³, que Floridi se refere como vida “onlife”¹⁴ (FLORIDI, 2014, p. 43, tradução nossa). No futuro não muito distante, a distinção entre a vida *online* e *offline* deverá ficar cada vez mais desfocada, até chegar ao ponto de desaparecer.

A “infosfera” se tornará cada vez menos um ambiente virtual apoiado por um mundo material. Aliás, com o passar dos anos, será o mundo em si que se compreenderá de uma maneira progressivamente mais informacionalizada, até se tornar uma forma de expressão da infosfera. “No final dessa mudança, a infosfera terá deixado de ser uma forma de se referir ao espaço de informações para se tornar sinônimo da própria realidade”¹⁵ (FLORIDI, 2014, p. 50, tradução nossa).

Proponho pensar neste trabalho seguindo o contexto colocado até aqui. Pretendo realizar reflexões éticas a respeito dos usos e consequências relacionados aos avanços das citadas TICs na sociedade informacional. Para isso, procuro primeiramente discutir sobre os

¹¹ “*We are not immobile, at the centre of the universe (Copernicus); we are not unnaturally distinct and different from the rest of the animal world (Darwin); and we are far from being entirely transparent to ourselves (Freud). We are now slowly accepting the idea that we might be informational organisms among many agents (Turing), inforgs not so dramatically different from clever, engineered artefacts, but sharing with them a global environment that is ultimately made of information, the infosphere. The information revolution is not about extending ourselves, but about re-interpreting who we are.*”

¹² “[...] *here (analogue, carbon-based, offline) and there (digital, silicon-based, online).*”

¹³ “[...] *the next stage in the development of the information age.*”

¹⁴ Este nome se dá pela junção entre o termo *online* e *life*, que significa “vida” em inglês.

¹⁵ “*At the end of this shift, the infosphere will have moved from being a way to refer to the space of information to being synonymous with reality itself.*”

tipos de éticas possíveis de serem estudadas em uma filosofia da informação, tratando do pensamento de autores como Rocci Luppigini para embarcar na construção de uma ideia de tecnoética, e depois, aprofundarei dentro dessa ideia para pensar em uma ética computacional através de James Moor para, em seguida, utilizar a ideia de Moor para explorar a ética de dados proposta por Luciano Floridi.

Após essa introdução, entrarei nos problemas éticos em si, começando por questões a respeito de discriminação e responsabilização envolvendo algoritmos e inteligência artificial, áreas da computação que serão brevemente explicadas no início do capítulo. Usarei como exemplo o robô Tay, e com o apoio de outros pensadores que estudaram sobre o caso, procuro discutir a quem deve ser atribuído a responsabilidade quando ocorrem casos de discriminação reproduzido por robôs. Logo em seguida, discutirei sobre como agem os moderadores de redes sociais e o tratamento para questões como discurso de ódio.

Em seguida, introduzirei aspectos do aprendizado de máquina, um subcampo da inteligência artificial, para falar a respeito da coleta massiva de dados, conhecida pelo termo *big data*. Nessa abordagem, pretendo apontar os desafios éticos concernentes ao armazenamento, coleta e uso de dados, começando pelo direito à privacidade, onde trarei novamente uma ideia de Luciano Floridi, em que o autor propõe uma interpretação autoconstitutiva do valor da privacidade. Com isso, explicarei com mais detalhe a tese de organismos informacionais (*inforgs*) e a visão da entidade humana como composta identitariamente pela sua informação. Logo após, trarei um foco maior ao domínio de dados para explicar sua relação com o lucro e o poder enquanto considerado uma nova *commodity*.

1 QUE TIPO DE ÉTICA?

1.1 Introdução à uma tecnoética

Em seu artigo de 2009 intitulado “*The Emerging Field of Technoethics*”, Rocci Lupplicini procura definir tecnoética limitando inicialmente suas duas partes. Primeiramente, traz a definição de ética como um ramo da filosofia focado em questões de costumes e valores, preocupando-se com questões como certo e errado, bem e mal, e responsabilizações. “No sentido geral, teorias éticas são teorias e crenças sobre padrões morais e como eles afetam a conduta”¹⁶ (LUPPICINI, 2009, p. 4, tradução nossa). Já o termo “tecnologia” possuiu diversos significados no decorrer dos séculos. Na Grécia antiga, estava relacionado com a palavra *techné* (técnica) referindo-se a arte e ao artesanato, enquanto no contexto atual possui uma maior relação com a própria atividade humana. “Isso ocorre porque muito do que as pessoas fazem e como o fazem depende do uso da tecnologia, particularmente em áreas onde contribui para agregar eficiência e eficácia à prática humana”¹⁷ (LUPPICINI, 2009, p. 4, tradução nossa). Em seu texto, o autor usa o termo “tecnologia” como “um conceito relacional aplicado ao projeto, desenvolvimento ou aplicação de dispositivos, máquinas e técnicas para uso humano”¹⁸ (LUPPICINI, 2009, p. 4, tradução nossa).

Tecnoética seria, para Lupplicini, “[...] um campo interdisciplinar preocupado com todos os aspectos éticos da tecnologia dentro de uma sociedade moldada pela mesma”¹⁹ (LUPPICINI, 2009, p. 4, tradução nossa). Nesse campo, observa-se não apenas uma conexão de práticas e processos humanos com a tecnologia, mas também sua inserção nos aspectos sociais, políticos e humanos da vida. Além disso, examina-se também os dilemas que ocorrem como resposta desta intervenção tecnológica dentro dos aspectos citados como, por exemplo, o debate crítico a respeito do uso responsável da tecnologia para promover os interesses humanos na sociedade. Em suma, a tecnoética procura “fornecer embasamento conceitual para esclarecer o papel da tecnologia em relação aos afetados por ela e ajudar a orientar a

¹⁶ “*In the most general sense, ethical theories are theories and beliefs about moral standards and how they affect conduct.*”

¹⁷ “*This is because a great deal of what people do and how they do it depends on the use of technology, particularly in areas where technology contributes added efficiency and effectiveness to human practice.*”

¹⁸ “*(...) a relational concept applying to the design, development, or application of devices, machines, and techniques for human use.*”

¹⁹ “*(...) an interdisciplinary field concerned with all ethical aspects of technology within a society shaped by technology.*”

resolução ética de problemas e a tomada de decisões em áreas de atividade que dependem da tecnologia”²⁰ (LUPPICINI, 2009, p. 4, tradução nossa).

Lupplicini, ao longo do texto, aponta alguns campos envolvendo a tecnologia que podem trazer diferentes debates na tecnoética, e.g. na área da engenharia, biotecnologia, ambiental, militar, etc. No presente trabalho, o foco estará principalmente em algo próximo daquilo que ele define como ética da internet.

O interesse pela ética da Internet surgiu a partir do final dos anos 80 e início dos 90, principalmente após a criação da primeira política que estabelecia diretrizes gerais para a proteção de atividades antiéticas online nos Estados Unidos chamada *Internet Architecture Board* (Conselho de Arquitetura da Internet). Ele cita que algumas dessas atividades são, por exemplo, “obter acesso não autorizado a recursos da Internet, interromper o uso pretendido da Internet, comprometer a integridade de informações computacionais e comprometer a privacidade dos usuários”²¹ (1989 apud LUPPICINI, 2009, p. 9, tradução nossa).

1.2 A ética computacional e de dados

O artigo “*What is Computer Ethics?*” (O que é ética computacional?) foi escrito em 1985 e é considerado um dos pioneiros no assunto, onde James Moor aponta que o desenvolvimento das tecnologias computacionais fornecem a elas novas capacidades, o que, por sua vez, traz à sociedade e aos indivíduos novas formas de agir. Essa dinâmica cria vácuos de políticas (*policy vacuums*), e para lidar com esses vácuos é necessário pensar em uma ética computacional com o objetivo de determinar o que deve ser feito e formular políticas públicas, a fim de enfrentar as novas possibilidades (MOOR, 1985, p. 266). Deborah Johnson cita alguns exemplos de vácuos de políticas criados pelas tecnologias computacionais:

Os empregadores deveriam monitorar os empregados na medida do possível com software de computador? Os médicos devem realizar cirurgias remotamente? Devo fazer cópias de software proprietário? Há algum mal em eu assumir uma pseudoidentidade em uma sala de bate-papo on-line? As empresas que fazem

²⁰ “(...) provide conceptual grounding to clarify the role of technology in relation to those affected by it and to help guide ethical problem-solving and decision making in areas of activity that rely on technology.”

²¹ “This policy provided the first comprehensive set of general guidelines to guard against unethical activity including, gaining unauthorized access to Internet resources, disrupting the intended use of the Internet, compromising the integrity of computer-based information, and compromising the privacy of users.”

negócios on-line deveriam ser autorizadas a vender as informações geradas pelas transações que coletam?²² (JOHNSON, 2004, p. 66, tradução nossa)

Em referência à ideia de ética computacional colocado por Moor, Luciano Floridi e Mariarosaria Taddeo publicaram em 2016 um artigo na revista *The Philosophical Transactions of the Royal Society* intitulado “*What is data ethics?*” (O que é ética de dados?), em que sugere um campo ético ainda mais específico. No resumo de seu artigo, os autores definem a ética de dados da seguinte forma:

Um novo ramo da ética que estuda e avalia problemas morais relacionados a dados (incluindo geração, registro, curadoria, processamento, disseminação, compartilhamento e uso), algoritmos (incluindo inteligência artificial, agentes artificiais, aprendizado de máquina e robôs) e práticas correspondentes (incluindo inovação responsável, programação, hacking e códigos profissionais), a fim de formular e apoiar soluções moralmente boas (por exemplo, condutas corretas ou valores corretos).²³ (FLORIDI; TADDEO, 2016, p.1, tradução nossa)

Esse novo ramo apresentado é essencial para trazer uma questão ainda mais atual na discussão da tecnoética. Ao falar de uma ética da computação ou uma ética da Internet, coloca-se como questão central um problema informacional. No entanto, Floridi e Taddeo (2016, p. 1-2) procuram obter um foco maior na questão dos dados, a fim de trazer um nível mais alto de abstração a esse problema. Com isso, a discussão não é apenas sobre os aspectos concretos a respeito do uso dos dados, ela é elevada a um nível mais abstrato, sutil, onde os dados não são traduzidos diretamente em informação, mas usados para gerar determinado comportamento ou induzir a alguma ação.

Tal nível de abstração é essencial para diferenciar a ética de dados de outras tecnoéticas. No artigo, o termo “nível de abstração” é simplificado como LoA (*level of abstraction*); ao falar do nível centrado na informação, usa a sigla LoA_I, o centrado na computação seria LoA_C e o que é centrado em dados seria LoA_D. Esse salto no nível de

²² “Should employers monitor employees to the extent possible with computer software? Should doctors perform surgery remotely? Should I make copies of proprietary software? Is there any harm in me taking on a pseudo-identity in an online chatroom? Should companies doing business online be allowed to sell the transaction-generated information they collect?”.

²³ “A new branch of ethics that studies and evaluates moral problems related to data (including generation, recording, curation, processing, dissemination, sharing and use), algorithms (including artificial intelligence, artificial agents, machine learning and robots) and corresponding practices (including responsible innovation, programming, hacking and professional codes), in order to formulate and support morally good solutions (e.g. right conducts or right values).”

abstração da ética da informação para a ética de dados é considerado a última revolução informacional.

Primeiramente, é importante demonstrar o que Floridi e Taddeo querem dizer com nível de abstração. Os autores colocam uma rápida definição na nota de rodapé que pode ser essencial para o entendimento do raciocínio:

O método de abstração [...] especifica os diferentes LoAs em que um sistema pode ser analisado, focando em diferentes aspectos, chamados observáveis. A escolha dos observáveis depende do propósito da análise e determina a escolha do LoA. Qualquer sistema pode ser analisado em diferentes LoAs. Por exemplo, um engenheiro interessado em maximizar a aerodinâmica de um carro pode se concentrar na forma de suas peças, peso e materiais. Um cliente interessado na estética do mesmo carro pode se concentrar em sua cor e na aparência geral e pode desconsiderar a forma, o peso e o material dos componentes do carro.²⁴ (FLORIDI; TADDEO, 2016, p. 2, tradução nossa)

Floridi e Taddeo dizem que a principal questão no chamado LoA_C estava em entender os impactos dos computadores na dinâmica social e no design no ambiente que nos cerca, uma vez que destacava a natureza dos computadores como ferramentas maleáveis e universais. Com LoA_I , o foco tornou-se o conteúdo da informação que era capaz de ser criado, registrado, processado e compartilhado pelos meios tecnológicos.

Ao fazer isso, LoA_I enfatizou as diferentes dimensões morais da informação – ou seja, informação como fonte, resultado ou alvo de ações morais – e levou ao delineamento de uma abordagem macroética capaz de abordar todo o ciclo de criação, compartilhamento, armazenamento, proteção, uso e possível destruição da informação.²⁵ (FLORIDI; TADDEO, 2016, p. 2, tradução nossa)

No entanto, a ciência de dados tem tornado a análise ética de dados cada vez mais frutífera e isso tem mudado também o nível de abstração. Floridi e Taddeo explicam (FLORIDI; TADDEO, 2016, p. 3) que o uso do termo “dado” ao invés de “informação”

²⁴ “*The method of abstraction is a common methodology in computer science and in philosophy and ethics of information. It specifies the different LoAs at which a system can be analysed, by focusing on different aspects, called observables. The choice of the observables depends on the purpose of the analysis and determines the choice of LoA. Any given system can be analysed at different LoAs. For example, an engineer interested in maximizing the aerodynamics of a car may focus upon the shape of its parts, their weight and the materials. A customer interested in the aesthetics of the same car may focus on its colour and on the overall look and may disregard the shape, weights and material of the car’s components.*”

²⁵ “*In doing so, LoA_I emphasized the different moral dimensions of information—i.e. information as the source, the result or the target of moral actions—and led to the design of a macroethical approach able to address the whole cycle of information creation, sharing, storage, protection, usage and possible destruction.*”

ressalta a real variável da questão que afeta os dias de hoje, uma vez que entende-se cada vez mais que se trata não de uma tecnologia específica (como computadores, celulares, tablets, plataformas online, etc.), mas sim a capacidade de manipulação de qualquer tecnologia digital. “Não é o *hardware* que causa problemas éticos, é o que o *hardware* faz com o *software* e os dados que representa a fonte de nossas novas dificuldades.”²⁶ (tradução nossa). O LoA_D ressalta que o problema não está na informação obtida, na verdade, ele vem antes disso. Trata-se do fato de como a “coleta, curadoria, análise e uso de dados” tem trazido problemas éticos que afetam temas importantes como “privacidade, anonimato, transparência, confiança e responsabilidade” (tradução nossa). Dentro da ética de dados, Floridi e Taddeo (2016, p. 4) citam três eixos de pesquisa: dados, algoritmos e práticas. Ao longo do artigo, explicam que a necessidade de colocar essas linhas de pesquisa em mais de um eixo se dá pelo fato de haver um entrelaçamento entre elas dentro do espaço conceitual.

Por exemplo, análises com foco na privacidade de dados também abordarão questões relativas a responsabilidades profissionais e de consentimento. Da mesma forma, a auditoria ética de algoritmos geralmente implica em análises das responsabilidades de seus designers, desenvolvedores, usuários e adotantes. A ética dos dados deve abordar todo o espaço conceitual e, portanto, todos os três eixos de pesquisa juntos, mesmo que com prioridades e focos diferentes.²⁷ (FLORIDI; TADDEO, 2016, p. 4, tradução nossa)

Concernente à ética focada em dados, concentra-se na coleta de um grande número de dados a ser analisado e sua utilização em diversas áreas, como acontece na promoção de publicidades por exemplo. A mineração de dados é utilizada para construir um perfil de consumidor, que é induzido a tomar certos tipos de ações a fim de favorecer o anunciante. Este é apenas um dos usos da coleta e análise de dados. O principal problema é que essa criação e identificação de perfis pode trazer grandes problemas éticos concernente à exclusão de grupos minoritários. Outro dilema se encontra na ausência de transparência que há nessa coleta devido à falta de conscientização da população em respeito aos seus benefícios e riscos, como também do que exatamente está sendo obtido, onde estão sendo armazenados e para que são utilizados. No entanto, Floridi e Taddeo (2016, p. 4) apontam diversos desafios presentes

²⁶ “*It is not the hardware that causes ethical problems, it is what the hardware does with the software and the data that represents the source of our new difficulties*”

²⁷ “*For example, analyses focusing on data privacy will also address issues concerning consent and professional responsibilities. Likewise, ethical auditing of algorithms often implies analyses of the responsibilities of their designers, developers, users and adopters. Data ethics must address the whole conceptual space and hence all three axes of research together, even if with different priorities and focus.*”

na obtenção de uma transparência adequada de dados, pois não se sabe ainda exatamente quais informações devem ser disponibilizadas e para quem especificamente.

Na ética de algoritmos, a questão se encontra no desenvolvimento algorítmico, o que leva a uma crescente constante no avanço da complexidade e autonomia de tecnologias envolvendo métodos de aprendizado de máquina (*machine learning*), tal como em máquinas de inteligência artificial. O desafio gira em torno principalmente dos seus desenvolvedores e sua responsabilização e prestação de contas concernente às consequências de comportamentos indesejados e imprevistos de suas criações. O estudo da ética dos algoritmos demonstra uma necessidade cada vez maior em avaliar os possíveis resultados indesejados com o objetivo de, por exemplo, evitar a promoção de conteúdos discriminatórios em redes sociais.

Já a ética das práticas aborda as práticas dos órgãos responsáveis pela regulação do processamento de dados. Nesse caso, o foco da questão está na responsabilização de todos aqueles que são encarregados em definir as estruturas éticas que moldem códigos profissionais de processos, estratégias e políticas de dados em torno da inovação, desenvolvimento e uso responsáveis.

2 DISCRIMINAÇÃO E RESPONSABILIZAÇÃO

2.1 Inteligência artificial e algoritmos

Neste capítulo, pretendo discorrer a respeito da questão proposta anteriormente ao colocar a promoção da discriminação social como uma das potenciais consequências trazidas pelo aprendizado de máquina e o desenvolvimento da inteligência artificial. Primeiramente, gostaria de fazer uma breve introdução a esses conceitos.

A IA tem como objetivo funcionar como uma “máquina pensante”, similar ao da mente humana, ao realizar tarefas, partindo da visão funcionalista em que a identidade dos estados mentais funciona como um computador, possuindo um *input* (ou entrada) e um *output* (ou saída). Na perspectiva funcionalista da mente, o *input* desses estados mentais seriam as relações causais que acontecem por meio das experiências sensoriais, que geraria relações causais entre os estados internos do organismo e resultaria em um *output*, que são os

comportamentos causados. A mente então seria equivalente ao *software* da máquina e o cérebro seria o *hardware*. (LECLERC, 2018, p. 81)

O estudo da computação teve seu primórdio através do desenvolvimento daquilo que era chamado de “máquina computacional” (*computing machine*), criada com o objetivo de simular o trabalho do pensamento humano para solucionar problemas matemáticos. Um dos pioneiros dessa tentativa é o matemático inglês Charles Babbage, responsável por propor a Máquina Diferencial em meados do século XIX. Essa “máquina de computação digital para fins especiais” (*special-purpose digital computing machine*) era capaz de calcular e produzir automaticamente tabelas matemáticas, como a logarítmica e a trigonométrica, e era composta inteiramente por componentes mecânicos (rodas, engrenagens, hastes, etc.). Um modelo em pequena escala foi exibido em 1822, no entanto, Babbage nunca chegou a construir seu design completo. (COPELAND, 2020)

Anos depois, Babbage propôs, em trabalho conjunto com Ada Lovelace, a Máquina Analítica. Esse “computador digital mecânico de uso geral” (*general-purpose mechanical digital computer*) tratava-se de um desenvolvimento ambicioso da Máquina Diferencial. Funcionaria da seguinte maneira:

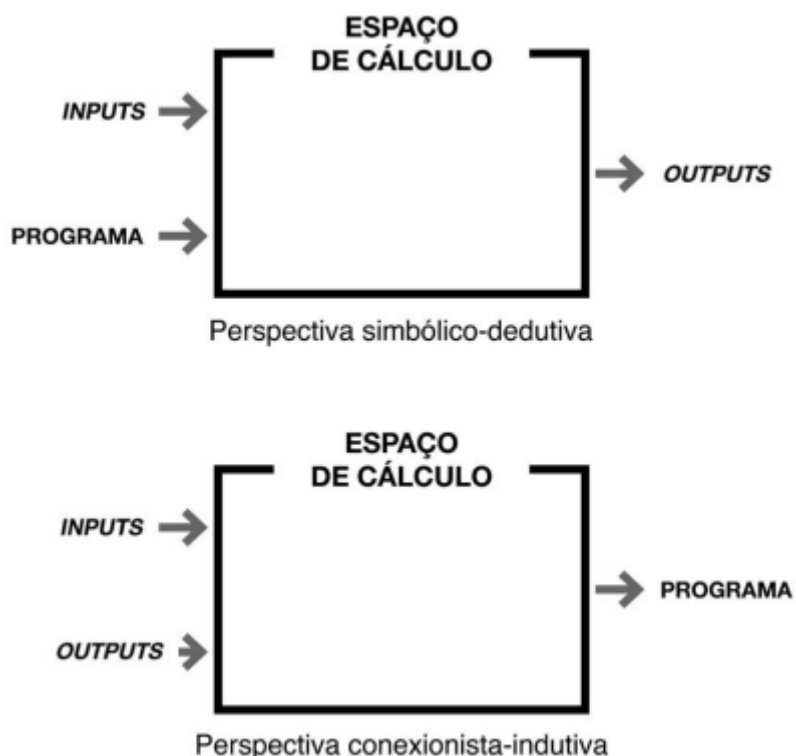
O comportamento da Máquina Analítica teria sido controlado por um programa de instruções contidas em cartões perfurados conectados por fitas (uma ideia que Babbage havia adotado do tear Jacquard). Babbage enfatizou a generalidade da Máquina Analítica, dizendo que “as condições que permitem que uma máquina finita faça cálculos de extensão ilimitada são satisfeitas na Máquina Analítica” (BABBAGE apud COPELAND, 2020, tradução nossa).²⁸

Devido ao orçamento limitado da época, Babbage nunca foi capaz de construir o design da máquina em escala completa, mas o modelo de linguagem pôde ser desenvolvido por Ada Lovelace, que hoje é considerada a primeira desenvolvedora de um algoritmo a ser processado por uma máquina. Lovelace foi capaz de prever um potencial ainda maior para os usos da Máquina Analítica além da computação não-numérica, como por exemplo, para compor peças musicais elaboradas. (COPELAND, 2020)

²⁸ “*The behaviour of the Analytical Engine would have been controlled by a program of instructions contained on punched cards connected together with ribbons (an idea that Babbage had adopted from the Jacquard weaving loom). Babbage emphasised the generality of the Analytical Engine, saying ‘the conditions which enable a finite machine to make calculations of unlimited extent are fulfilled in the Analytical Engine’.*”

Na obra “Racismo algorítmico: inteligência artificial e discriminação nas redes digitais”, Tarcízio Silva (2022, p. 62) apresenta duas perspectivas da IA. Até o momento, foi citado neste trabalho predominantemente a perspectiva *simbólico-dedutiva*, focada numa relação entre o programa e os *inputs*, resultando em *outputs*. Contudo, uma nova perspectiva prevaleceu a partir de 1990: a *conexionista-indutiva*. Essa perspectiva é resultado do desenvolvimento de máquinas cada vez mais rápidas e com uma imensa capacidade de memória, permitindo coletas de dados em grande escala (conhecidos como *big data*) e criando algoritmos mais elaborados e eficientes. Dessa maneira, através de campos como o aprendizado de máquina, procura-se alcançar um funcionamento cada vez mais próximo da mente humana com o desenvolvimento de redes neurais. Silva utiliza o seguinte esquema para ilustrar as duas perspectivas.

Figura 1 - “Perspectiva simbólico-dedutiva x perspectiva conexionista-indutiva na inteligência artificial.” (CARDON, COINTET, MAZIÈRES apud SILVA, 2022, p.64)



Fonte: Racismo algorítmico: inteligência artificial e discriminação nas redes digitais.

“Como é possível observar no esquema mostrado na figura, no modelo simbólico-dedutivo o campo dos resultados é uma relação das entradas de instâncias

de dados e parâmetros do programa em um espaço de cálculo que vai gerar tais resultados. A abordagem conexionista-indutiva age de modo diferente: geralmente a partir de uma quantidade massiva de dados de entrada e de saída previamente conhecidos, seu objetivo é construir e atualizar constantemente o “programa” para a otimização contínua dos níveis de objetivo alcançados.” (SILVA, 2022, p. 64-65)

Dessa forma, os algoritmos se desenvolvem de maneira autônoma, “substituindo programadores por sistemas de aprendizado” (ALPAYDIN apud SILVA, 2022, p. 66). Mas afinal, o que são algoritmos? De acordo com Cormen et al. (2012, p. 17-23), um algoritmo é uma ferramenta desenvolvida para resolver um problema computacional específico, descrevendo o procedimento e as instruções necessárias para transformar valores ou conjunto de valores a fim de obter a solução do problema.

Um exemplo dado por Silva (2022, p. 66) de “sistema de aprendizado” é o que ocorre com o algoritmo do Youtube. O sistema reúne uma série de sinais deixados pelo usuário, tais como movimentação do mouse, cliques, anúncios assistidos, tempo de uso, entre outros. Todo o engajamento é usado com o objetivo de definir como a plataforma deve funcionar para aquele indivíduo, determinando o que será exibido, “para quem, em quais horários, em relação a quais outros vídeos, etc.” (SILVA, 2022, p. 66). Contudo, esse sistema acaba por promover vídeos independentemente dos impactos sociais e culturais gerados por ele, priorizando sempre “metas inequívocas de negócio”.

Assim, um efeito colateral do crescente engajamento dos usuários na plataforma assistindo a vídeos extremistas seria, assim, alheio ao modelo de negócio. Os objetivos de gerar mais visualizações, cliques, visitas repetidas e compras estariam sendo alcançados, e isso seria o suficiente. (SILVA, 2022, p. 66)

De acordo com Silva (2022, p. 66), o entendimento desses programas de aprendizado é essencial para compreender “a distribuição de sistemas algorítmicos, seus riscos inerentes e, principalmente, as estratégias de evasão de responsabilidade das organizações que lucram com tais sistemas”.

2.2 O chatbot Tay e atribuição de responsabilidade

As empresas responsáveis pelo desenvolvimento dessas tecnologias usam frequentemente a máscara da neutralidade tecnológica para fugir da culpa em torno dos comportamentos inesperados que podem ocorrer no funcionamento da IA. Um exemplo disso dado por Silva (2022, p. 67-71) foi o que ocorreu com o *chatbot* Tay, lançado pela Microsoft em parceria com o Twitter. Lançada em 2016, Tay representava uma mulher jovem e foi criada visando estabelecer um contato descontraído com jovens estadunidenses na faixa dos 16 aos 24 anos. Seu funcionamento se dava através de um perfil próprio no Twitter, e seu comportamento era moldado de acordo com as interações entre os usuários da rede em tempo real. No entanto, essas conversas acabaram direcionando o robô a um comportamento misógino e racista, e rapidamente a Microsoft resolveu retirá-la do ar, para lançar novamente com alguns ajustes pouco tempo depois. O robô acabou voltando ainda pior, publicando posts extremamente ofensivos que apoiavam temas como o genocídio de mexicanos, Holocausto e ódio à população negra. Poucas horas depois, a Microsoft retirou Tay definitivamente do ar.

No artigo de Sara Suárez-Gonzalo, Lluís Mas-Manchón e Frederic Guerrero-Solé, intitulado “*Tay is you. The attribution of responsibility in the algorithmic culture.*” (Tay é você. A atribuição de responsabilidade na cultura algorítmica.), os autores procuram estudar a atribuição de responsabilidade sobre o *chatbot* que é colocada pela mídia. Responsabilidade, de acordo com o texto (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 3), está tradicionalmente ligada a ações que acontecem por meio de intenções concretas geralmente atribuídas àqueles que são referidos como “agentes” (sendo eles indivíduos, grupos de indivíduos ou instituições), e que são capazes de gerar consequências significativas à outros, denominados “pacientes”. Dessa forma, para a atribuição de responsabilidade a um agente, é necessário que ele seja racional, e também que possua intenção e agência sobre a sua ação.

Uma conceito importante para a atribuição de responsabilidade é a possibilidade de prestação de contas (*accountability*) do agente, para que o mesmo possa responder pelas suas ações. No artigo (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p.3), o conceito é apontado como um fator importante para que haja o balanceamento entre os poderes, que geralmente existe como parte de relacionamentos como aquele entre sociedade e Estado. A prestação de contas possui duas dimensões, sendo elas a dimensão retrospectiva (quando o agente é culpado ou punido pela ação) e a dimensão prospectiva (quando são definidos obrigações e deveres relacionados à ação).

Contudo, a atribuição de capacidades associadas à responsabilização e prestação de contas em sistemas como o chatbot (ou seja, um “ator controlado por algoritmos que aprendem com o comportamento dos outros”²⁹) pode ser muito mais trabalhoso. Não é simples definir se o chatbot pode ser considerado um agente racional, tampouco se suas ações são intencionais ou não. Além disso, existe outra questão: esses algoritmos, por serem desenvolvidos através do aprendizado de máquina, fazem o sistema ser considerado uma caixa preta (*blackbox*), ou seja, “seu funcionamento interno é demasiado complicado ou muito baseado na aleatoriedade e, por isso, torna os resultados difíceis de prever”³⁰ (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 4, tradução nossa). Devido a isso, a forma que as tarefas e ações são atribuídas nas relações sociotécnicas modernas entre humanos e não humanos acontece de maneira imprevisível. Aqueles que operam os robôs também não podem ser responsabilizados pelas ações do sistema por não serem capazes de prever seu comportamento, o que torna a atribuição da “intencionalidade, causalidade e o vínculo agente-paciente”³¹ (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 4, tradução nossa) ainda mais complexa. É possível notar que, “conforme a complexidade e autonomia do aprendizado de máquina aumenta, mais os humanos perdem o controle sobre elas, atribuindo o processo de tomada de decisão às próprias máquinas”³² (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 4, tradução nossa) e, assim, criando uma “brecha de responsabilidade” (*responsibility gap*). O artigo pontua mais duas idéias equivocadas sobre tecnologia e culpabilidade que complicam a atribuição de responsabilidade. Uma delas é a já citada presunção de neutralidade dos artefatos tecnológicos, e outra é o entendimento predominantemente reducionista da responsabilidade, que considera apenas sua dimensão retrospectiva.

Devido à essas complexidades, diversos autores atualmente têm procurado repensar a respeito dos conceitos de responsabilidade e prestação de contas no atual mundo da tecnologia, para que seja possível aplicá-los nesse tipo de contexto. O artigo pontua duas visões predominantes: a abordagem dependente do contexto e a abordagem estruturalista.

²⁹ “[...] *algorithmically controlled actors that learn from others’ behaviors.*”. SUÁREZ-GONZALO, MAS-MANCHÓN e GUERRERO-SOLÉ, 2019, p. 4, tradução nossa

³⁰ “[...] *social bots are peculiar black boxes in which the inner workings are either too complicated or based on randomness, and so the outcomes become difficult to foresee.*”.

³¹ “*Therefore, intentionality, causality and the agent-patient link become hard to define.*”.

³² “*As the complexity and autonomy of learning machines keep growing, humans cannot continue to be directly responsible for them. In some way, humans lose control over them, and bestow the decision-making process to the machines themselves.*”.

Dentro da primeira abordagem, “a responsabilidade corresponde ao ambiente com o qual o bot interage e, portanto, sua assunção se dispersa entre todos os atores”³³ (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 4). Nessa perspectiva, pensadores como Floridi entendem que “quando um artefato aprende com o contexto em que atua, a intencionalidade se espalha através das diferentes relações e interesses externos envolvidos na interação”³⁴. Lucas Introna, também dentro da abordagem contextual, estabelece uma perspectiva foucaultiana ao trazer um conceito interacional de intencionalidade, que é definida pela “interação inseparável entre artefatos técnicos (*dispositifs*) e humanos”³⁵. Já José van Dijck³⁶ pressupõe que, dado que o ambiente como um todo pode ser considerado o *input* dos chatbots, cuja identidade é baseada em interações simples, o conceito de agência também é considerado um “conceito em rede” (*networked concept*). Ainda sobre agência, Neff e Nagy desenvolvem a ideia de “agência simbiótica” (*symbiotic agency*), que é definido como

[...] o que os usuários, atores e ferramentas fazem quando interagem com sistemas tecnológicos complexos [...]. Em outras palavras, o que as pessoas dizem sobre as influências dos bots, o que as pessoas podem fazer com eles e quais capacidades os bots têm para ação social”³⁷

Ao apresentar a abordagem estruturalista, o artigo usa de exemplo autores como Deborah Johnson, que argumenta que tecnologias computacionais possuem intencionalidade mesmo não sendo intencionais, uma vez que sua intencionalidade está relacionada com a dos seus designers e usuários (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 4). Para a autora, é necessário considerar “as forças sociais, políticas e institucionais interessadas em moldar o desenvolvimento tecnológico”³⁸. Também são citados dentro dessa abordagem Ford, Dubois e Puschmann (SUÁREZ-GONZALO; MAS-MANCHÓN;

³³ “[...] *responsibility corresponds to the environment with which the bot interacts, and so its assumption disperses among all actors*”.

³⁴ “[...] *when an artifact learns from the context in which it performs, intentionality spreads through the different relationships and outer interests involved in the interaction*”. (FLORIDI apud SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 4, tradução nossa)

³⁵ “[...] *defined as the inseparable interaction between technical artefacts (“dispositifs”) and humans*”. (INTRONA apud SUÁREZ-GONZALO, MAS-MANCHÓN e GUERRERO-SOLÉ, 2019, p. 4, tradução nossa)

³⁶ VAN DIJCK apud SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 4.

³⁷ “*what users, actors, and tools do when interacting with complex technological systems [...] In other words, what people say about bots influences, what people can do with them and what capacities the bots have for social action*”. (NEFF e NAGY apud SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 4, tradução nossa)

³⁸ “*the social, political and institutional forces interested in shaping technological developments*”. (JOHNSON apud SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 5, tradução nossa)

GUERRERO-SOLÉ, 2019, p. 5, tradução nossa), que defendem que a prestação de contas das ações dos chatbots deve ser feita através do conjunto de diferentes interesses envolvidos em sua cocriação. Bernhard Rieder é outro autor que, nessa perspectiva, examina a relação entre “governamentalidade” (*governmentality*) e computação e ressalta a importância de “lidar com os computadores como ferramentas políticas nas mãos de atores interessados ou *think tanks*”³⁹.

No artigo, os autores procuram responder questões a respeito de como os meios de comunicação predominantes trataram o fracasso do chatbot Tay e como foi atribuída a responsabilidade por esse fracasso.

Foi observado (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 7) que a procura pelo responsável era um padrão recorrente entre as notícias analisadas, que frequentemente era colocado da seguinte maneira: “Os usuários do Twitter (o agente) usaram e atacaram de forma maliciosa (contribuição causal) um chatbot fraco e vulnerável chamado Tay (o paciente) que teve que ser desconectado por seu designer (consequência)”⁴⁰ (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 7, tradução nossa). No total da análise, foi identificado que três em cada quatro notícias procuravam identificar o responsável do comportamento de Tay. Entre eles, um terço das matérias descrevia que a principal causa do evento foi devido a um ataque orquestrado por usuários do Twitter, que levaram ao chatbot a agir de maneira inapropriada. De acordo com o texto, 40% das matérias se referiam precisamente aos usuários do Twitter como os principais agentes do incidente, enquanto apenas 17% culpavam a Microsoft. Além disso, 18% das notícias colocava a interação do robô com os humanos como o principal provocador do incidente e 14% apontavam o código do chatbot como causador do problema.

Um terço das matérias reportaram que a principal consequência do evento foi a desconexão de Tay e o pedido de desculpas feito pela Microsoft. Em 25% das notícias, as ações de Tay foram colocadas como uma espécie de espelho que serviria para refletir o pior da humanidade após ser ensinada a agir de forma discriminatória. A mídia, ao colocar as ações do robô como resultado de um ataque, estabelece a existência de um conflito entre os

³⁹ “examines the relationship between governmentality and computing and notes the importance of dealing with computers as political tools in the hands of interested actors or think tanks”. (RIEDER apud SUÁREZ-GONZALO, MAS-MANCHÓN e GUERRERO-SOLÉ, 2019, p. 5, tradução nossa)

⁴⁰ “Twitter users (the agent) maliciously misused and attacked (causal contribution) a feeble and vulnerable chatbot called Tay (the patient) that had to be disconnected by its designer (consequence)”.

usuários do Twitter e o chatbot. Além disso, o artigo ressalta que grande parte das matérias analisadas tratavam o evento de forma dramática, ao colocar Tay como uma pessoa e fazendo julgamentos sensíveis sobre o robô. Inclusive, era recorrente a representação de Tay como o paciente do evento, ou seja, o ator mais afetado, e em 33% das matérias era vista mais precisamente como “um ser humano e um ‘bebê robô’ prejudicado pelos abusadores”⁴¹ (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 8, tradução nossa). Apenas 12% consideravam os grupos minoritários atacados pelas mensagens de Tay como os principais prejudicados pelo incidente, e mesmo assim, o artigo ressalta que “não houve referência explícita ao discurso de ódio e às suas consequências sobre as pessoas, nem à questões jurídicas”⁴² (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 8, tradução nossa).

Após análise, o artigo constata que os veículos de mídia frequentemente reforçam um discurso contraditório: “por um lado, a mídia pessoalizou Tay e a tratou como algo capaz de sentir e sofrer. Por outro lado, os usuários do Twitter são desumanizados e considerados culpados pela transformação de Tay em um ser misógino e racista”⁴³ (SUÁREZ-GONZALO; MAS-MANCHÓN; GUERRERO-SOLÉ, 2019, p. 8, tradução nossa). Esse discurso, ao manifestar a atribuição de responsabilidade de maneira retrospectiva ao invés de realizar uma abordagem prospectiva, favorece principalmente os interesses da Microsoft, uma vez que permite que a empresa se esquive da culpa e trate o fracasso do chatbot como um evento isolado.

Apesar dos usuários terem sua parcela de responsabilidade por instigar o robô ao mau comportamento, os autores ressaltam que isso não pode ser tratado como a única causa da falha de Tay. O incidente decorre também da falta de preparo do algoritmo, que não foi desenvolvido propriamente para lidar com esse tipo de situação. O chatbot havia sido desenvolvido apenas para “replicar o tom e vocabulário de seus interlocutores”, porém, por ser uma máquina não-racional, é incapaz de fazer juízos de valor. Contudo, seus desenvolvedores deveriam ter consciência do perigo potencial embutido neste tipo de design e alguns valores a priori deveriam ter sido programados a fim de prevenir o fiasco.

⁴¹ “as a human being and a ‘baby robot’ harmed by the abusers”.

⁴² “There was no explicit reference to hate speech and its consequences over people, nor to legal issues”.

⁴³ “on the one hand, media personalized Tay and treated it as something capable to feel and suffer. On the other hand, Twitter users are dehumanized and found guilty on Tay’s turn into a misogynistic and racist being”.

Apesar desse comportamento imprudente, diversos entusiastas de tecnologia consideraram um robô Tay um grande sucesso e um exemplo da capacidade de inventividade da Microsoft. Em uma postagem de um popular blog de tecnologia⁴⁴, o autor, ao tentar explicar o ocorrido, deixou claro que “a primeira coisa a ser dita é que a Microsoft não tem culpa pelo que aconteceu”. De acordo com ele, o comportamento foi similar ao de uma criança, que “observam e aprendem com o mundo que veem à sua volta”, assim como citado anteriormente no padrão reconhecido nas matérias a partir das análises feitas. No final o autor conclui:

Sendo assim, eu acredito que o projeto teve um resultado brilhante, pois ele não só demonstrou que o sistema cognitivo de Tay está funcionando, como concluiu que o aprendizado desse robô se assemelha e muito à forma como nós humanos nos tornamos preconceituosos.

Entretanto, como defende Tarcísio Silva (2022, p. 69-70) em seu livro, é necessário pensar a respeito da responsabilidade de grandes empresas como a Microsoft e o Twitter, que possuem profissionais em diversas áreas envolvidos no planejamento do *bot* e com um enorme investimento, em aceitar realizar um lançamento dessa magnitude ignorando todos os fatores como “avaliação de riscos, padrões éticos e especulação de possíveis usos nocivos” envolvidos que certamente previam essa possibilidade. Esse descaso reflete a quem esse recurso é direcionado e o favorecimento de grandes grupos hegemônicos, reforçando um comportamento altamente problemático.

“Talvez esses erros frequentes, que chegam à superfície personificados em tecnologias como chatbots, a exemplo de Tay [...], não sejam glitches ou bugs, mas funcionalidades. Nos termos de Ruha Benjamin, talvez os erros sejam ‘antes um tipo de sinal de como o sistema opera. Não uma aberração, mas uma forma de evidência, jogando luz sobre falhas subjacentes em um sistema corrupto’”. (SILVA, 2022, p. 71)

2.3 Redes sociais e moderação de conteúdo

⁴⁴ Matheus Gonçalves, “O robô da Microsoft que aprende com humanos não demorou nem um dia para virar racista”, Tecnoblog, abr. 2016, disponível em: <<https://tecnoblog.net/193318/tay- robo-racista-microsoft/>>, acesso em: out. 2023.

Diversas pesquisas⁴⁵ mostram como o discurso reacionário e discriminatório tem crescido exponencialmente através da manipulação de algoritmos. Não faltam exemplos de situações onde as dinâmicas das redes sociais foram responsáveis por favorecer a propagação de discursos de ódio. Uma delas aconteceu em 2017, na mesma época em que ocorreu o vazamento de documentos internos do *Facebook* pela ProPublica, uma organização sem fins lucrativos que produz jornalismo investigativo. (ANGWIN e GRASSEGER, 2017)

Após um ataque terrorista em Londres, um congressista estadunidense chamado Clay Higgins escreveu uma postagem no Facebook que dizia: “Cace-os, identifique-os e mate-os. Mate a todos. Pelo amor do que é bom e correto. Mate a todos” (SILVA, 2022, p. 40). Apesar das denúncias, a postagem de Higgins não foi apagada pelos moderadores do *Facebook* responsáveis por detectar conteúdo ofensivo. Contudo, a poeta e ativista do Black Lives Matter, Didi Delgado, havia obtido uma resposta diferente meses antes ao escrever “Todas as pessoas brancas são racistas. Comece desse ponto de partida ou você já falhou” (SILVA, 2022, p. 40). Após publicar essas frases, o perfil de Delgado foi suspenso por sete dias e sua postagem foi deletada.

Os documentos vazados (ANGWIN e GRASSEGER, 2017) mostram as diretrizes estabelecidas pelo Facebook que deveriam ser seguidas pelos seus moderadores de conteúdo para fazerem a distinção entre manifestações políticas legítimas e discursos de ódio. Foi constatado que parcela considerável das regras favoreciam as elites e governos em detrimento de ativistas e minorias raciais. A diferença de tratamento entre duas publicações da rede consegue ilustrar bem como funcionam essas regras.

Através desses documentos (ANGWIN e GRASSEGER, 2017), foi possível descobrir a metodologia por trás da moderação plataforma, mostrando o que deve ser proibido e o que está permitido. Dentre essas regras, existem “categorias protegidas”, que seriam grupos que a rede social procura defender de discursos de ódio. Por exemplo, entre essas três categorias: mulheres motoristas, crianças negras e homens brancos, o único grupo que os moderadores do Facebook devem defender é o de homens brancos. As categorias “mulheres motoristas” e “crianças negras” são consideradas subcategorias. Isso é mostrado em um dos slides de treinamento que foi vazado.

⁴⁵ Exemplos de pesquisas citadas no artigo “*Tay is you. The attribution of responsibility in the algorithmic culture.*” foram: “*Media Manipulation and Disinformation Online*” por Alice Marwick e Rebecca Lewis e também “*Platformed Racism: The Mediation and Circulation of an Australian Race-Based Controversy on Twitter, Facebook and YouTube*” por Ariadna Matamoros-Fernández.

Nas duas publicações mencionadas, é possível ver essas regras sendo aplicadas. A publicação do deputado republicano não infringe as regras pois se refere a uma subcategoria — muçulmanos radicalizados —, enquanto o outro se aplica a brancos em geral. Ou seja, apesar de Higgins ter incitado uma violência mortal, foi Delgado quem sofreu punições por ter feito uma acusação a um grupo protegido pela plataforma. (SILVA, 2022, p. 40).

Como foi demonstrado anteriormente, não é possível depender da IA para fazer julgamentos morais, logo, as redes sociais recorrem a moderadores para fiscalizar se o que é postado está de acordo com as normas da comunidade. Para as postagens chegarem nas mãos de sua avaliação, é necessária a cooperação da comunidade para que elas sejam denunciadas e, só assim, tornar possível o impedimento de sua circulação. Em alguns casos, se uma postagem receber várias denúncias de usuários em pouco espaço de tempo, o conteúdo é automaticamente retirado do ar e apenas depois é revisado por humanos (SILVA, 2022, p. 41). A intenção é evitar a circulação de conteúdos que são indicados como “inequivocamente nocivos”, no entanto, esse recurso é frequentemente explorado para organização de ataques coordenados, em que grandes grupos se organizam com o objetivo de derrubar uma publicação. (SILVA, 2022, p. 42)

Esses moderadores são trabalhadores presentes em diversas partes do mundo e em sua maioria são pessoas de baixa renda, originárias de países com poucos direitos trabalhistas e que recebem baixa remuneração. Seu trabalho deve ser feito em segredo de forma que seja o mais invisível possível para os usuários da plataforma, e cabe a eles filtrar todo conteúdo reportado, principalmente aquele considerado potencialmente pornográfico, violento, perturbador e/ou repugnante. (ROBERTS, 2016, p. 2).

Em seu estudo direcionado ao trabalho dos moderadores em redes sociais, Sarah Roberts (2016, p. 2) aponta que a necessidade desses trabalhadores para a manutenção das redes sociais rompe com o mito da internet como um local em que exista um relacionamento direto do usuário com a plataforma. Como um moderador entrevistado por ela comentou (ROBERTS, 2016, p.3), experimentar uma internet sem eles seria uma enorme fossa (*cesspool*), e é com essa fossa que eles lidam em seu trabalho todos os dias.

Esses trabalhadores agem como guardiões, que necessitam analisar diversas questões ao mesmo tempo quando revisam publicações. Entre elas, Roberts cita:

Qual é a natureza desse conteúdo? Que significado a linguagem, os símbolos ou as imagens contidos nele transmitem? Qual é o contexto cultural do conteúdo em seu ponto de origem, no local e no site da Internet ao qual se destina e no local onde está sendo moderado? Quais são as políticas estabelecidas por sua empresa ou pela plataforma à qual o conteúdo se destina, com relação à tolerância a conteúdo polêmico, perturbador, racista, sexista, homofóbico ou sexual ou violentamente gráfico? (ROBERTS, 2016, p. 3, tradução nossa)

Além dessas questões, também é necessário lidar com a popularidade alcançada pelo conteúdo, uma vez que nesses casos a decisão de manter ou excluir teria um impacto maior na opinião pública.

Em plataformas que funcionam através da geração de conteúdo pelos seus usuários, estes, que são chamados “criadores de conteúdo”, possuem o objetivo de que suas postagens alcancem o maior número de pessoas para assim, obter um maior retorno financeiro. Quanto maior o número de interações obtidas em uma postagem, maior é o seu valor para a atração de novos espectadores. Quando algo faz sucesso, como por exemplo um vídeo no Youtube capaz de atrair milhões de “visualizações, curtidas, repostagens e imitações ou homenagens”⁴⁶ (ROBERTS, 2016, p. 3, tradução nossa), este conteúdo é chamado de “viral”. Esses conteúdos virais altamente popularizados geralmente possuem “ganchos textuais ou significantes-chave”⁴⁷ (BURGESS apud ROBERTS, 2016, p. 3, tradução nossa) que atraem o público. No entanto, esses ganchos também podem envolver questões problemáticas que promovam a discriminação, como por exemplo, estereótipos racistas.

Em conversas de Roberts (2016, p. 5) com os trabalhadores de uma grande empresa de rede social do Vale do Silício, vários comentaram sobre a frustração em ter que liberar a imagens claramente racistas, como as que envolviam *blackface*, na moderação do conteúdo. Um deles explicou que a empresa possuía políticas internas muito específicas e cheias de detalhes, e disse que se essas regras fossem colocadas em público, seria muito fácil violá-las. Em discussões durante reuniões semanais com chefes da área de políticas de segurança, estes colocavam que o *blackface* “não é tecnicamente considerado discurso de ódio por padrão”⁴⁸ (ROBERTS, 2016, p. 5, tradução nossa).

Essas e outras situações são colocadas por Roberts para ilustrar o fato de que “as empresas e plataformas de mídia social tomam decisões ativas sobre que tipos de imagens e

⁴⁶ “[...] *attracting millions of views, likes, reposts, and imitations or homages*”

⁴⁷ “[...] *textual hooks or key signifiers*”

⁴⁸ “[...] *blackface is not technically considered hate speech by default*”

conteúdos racistas, sexistas e odiosos irão hospedar e até que ponto irão hospedá-los”⁴⁹ (ROBERTS, 2016, p. 5, tradução nossa). Essas empresas frequentemente usam discursos como liberdade de expressão para se defender perante o grande público, no entanto, os interesses comerciais sempre possuem um peso maior nessa balança. Como coloca a autora, “se uma plataforma se tornasse notória por ser muito restritiva aos olhos da maioria de seus usuários, correria o risco de perder participantes para oferecer aos seus anunciantes”⁵⁰ (ROBERTS, 2016, p. 5, tradução nossa). Para evitar correr esse risco, elas escolhem permitir mais, e não menos.

Observar as práticas de moderação de conteúdo oferece uma visão especial a respeito das diferentes maneiras pelas quais as imagens e a linguagem discriminatória são implementadas nas mídias sociais. Existe uma série de fatores que permitem essa implementação. Por um lado, envolve questões como a “palatabilidade” desse conteúdo para um público alvo e seu potencial de viralização, enquanto por outro, a “probabilidade de causar ofensas e danos à marca”⁵¹ (ROBERTS, 2016, p. 9, tradução nossa).

Em resumo, a importância da moderação na perspectiva das grandes empresas de redes sociais é avaliada através do seu “valor potencial como mercadoria”⁵² (ROBERTS, 2016, p. 9, tradução nossa). Esse processo de moderação é regido, antes de tudo, por “políticas destinadas a encontrar um equilíbrio entre atrair usuários-participantes e anunciantes”⁵³ (ROBERTS, 2016, p. 9, tradução nossa), atendendo demandas e normas legais ao mesmo tempo que permanece atraente e lucrativo para seus investidores. Ou seja, suas políticas internas claramente procuram atender essas demandas em primeiro lugar ao invés de defender justiça sociais e direitos humanos.

No entanto, devido à invisibilidade da moderação de conteúdo, conteúdos e imagens racistas, misóginos e homofóbicos acabam sendo colocados como norma, uma vez que as estruturas que as encorajam estão escondidas. Dessa forma, Roberts (2016, p. 9) coloca que a imagem que passa para seus usuários é que a existência deste tipo de conteúdo é algo natural e orgânico, e não fruto de políticas com potencial altamente lucrativo. Ao revelar as práticas

⁴⁹ “[...] *social media companies and platforms make active decisions about what kinds of racist, sexist, and hateful imagery and content they will host and to what extent they will host it*”

⁵⁰ “[...] *if a platform were to become notorious for being too restrictive in the eyes of the majority of its users, it would run the risk of losing participants to offer to its advertisers*”

⁵¹ “[...] *the likelihood of it causing offense and brand damage*”

⁵² “*potential value as commodity*”

⁵³ “[...] *policies meant to strike a balance among attracting user-participants and advertisers*”

daqueles que trabalham com moderação de conteúdo e os mandatos e pressões colocadas sobre esse tipo de emprego, é possível ter “uma visão mais realista e completa do cenário das mídias sociais e de como o conteúdo chega até ela, e até nós”⁵⁴ (ROBERTS, 2016, p. 9, tradução nossa).

3 PRIVACIDADE DE DADOS

3.1 Aprendizado de máquina e *big data*

Discutir sobre a extração de dados é fundamental ao falar de algoritmos e inteligência artificial. E ao falar de extração de dados, é importante retomar a ideia de aprendizado de máquina que foi citada anteriormente.

O aprendizado de máquina é um subcampo da IA que mais se aproxima do funcionamento da mente humana. Para seu funcionamento, é necessário o armazenamento massivo de dados, e o desenvolvimento de máquinas capazes de realizar esse armazenamento tem sido o principal foco dos setores de tecnologia da informação nas últimas décadas. Tais dados coletados possuem “diversas naturezas, de micro e macro escalas, referentes a objetos e fenômenos, nos mais variados contextos” (HIRATA, 2020, p. 93) e é conhecido pelo termo *big data* devido à sua abundância, sendo “coletados ou gerados em grande volume e em alta velocidade” (HIRATA, 2020, p. 93). Essa busca e necessidade por uma quantidade cada vez maior de dados e os efeitos disso é definido pela comunidade científica como o ‘quarto paradigma da ciência’ (HEY, TANSLEY, TOLLE apud HIRATA, 2020, p. 93). Em termos mais técnicos,

big data se refere a dados que são grandes demais para um único servidor, muito diversos para se adequar a uma base de dados estruturada em linhas e colunas, ou cujo fluxo seja tão intenso que não permita adequação a um data warehouse [armazém de dados] estático. (DAVENPORT apud SANT’ANA, 2016, p. 2).

⁵⁴ “[...] *more realistic and fuller view of the social media landscape and how content makes its way to it, and to us*”

Antes do desenvolvimento do computador, os dispositivos não possuíam a capacidade de tratar do conteúdo. Isso quer dizer que serviam apenas para armazenamento, uma vez que não possuíam as “funções necessárias” (SANT’ANA, 2020, p. 117), o “como fazer” e suas devidas instruções para lidar com esses conteúdos. Apenas identificavam o que seria “coletado, armazenado, recuperado e descartado” (SANT’ANA, 2020, p. 117), tratando de registrar o conteúdo para que fosse possível acessá-lo e analisá-lo posteriormente, essa função dependendo totalmente do trabalho manual.

O desenvolvimento dos dispositivos computacionais tornou possível que eles obtivessem a função de calcular. A partir disso, Ricardo Sant’Ana (2020, p. 117) define que essa instrumentalização permitiu que os dispositivos atendessem a dimensão função. O primeiro grande salto foi dado por meio da otimização das técnicas no processo de registro, como a que já era usada na indústria têxtil utilizando cartões perfurados para armazenamento de dados. No final do século XIX, essa técnica começou a ser utilizada em processos como o levantamento censitário estatal nos Estados Unidos, servindo não apenas para a coleta e o armazenamento, como também para a execução de cálculos simples em grandes volumes de dados.

Durante o início da Segunda Guerra Mundial, em meados da década de 40, o aumento do interesse pela dimensão de funcionalidade foi responsável para o avanço principalmente do tratamento de dados, permitindo o apoio da forma automatizada do processamento. Sant’Ana (2020, p. 118) afirma que essa mudança de foco deslocou o conceito de computação “de humano para maquínico”, e diz que a capacidade de processamento aumentou de tal maneira que “a pressão por novos desenvolvimentos voltava à dimensão dos dados”.

Ao analisar historicamente, é possível notar que essas duas fases se alternam. O predomínio da fase da dimensão dos dados esteve presente desde o final do século XIX até a década de 40, que alternou para a fase do domínio da dimensão função principalmente durante o período de 1940 até 1990. Nos dias de hoje, existe uma retomada da fase da dimensão dos dados, que tem acelerado no século XXI sob o fenômeno chamado de *big data*. Com o *big data*, é possível ter à disposição três Vs: “grandes Volumes, com ampla Variedade e com alta Velocidade no acesso” (SANT’ANA, 2020, p. 118).

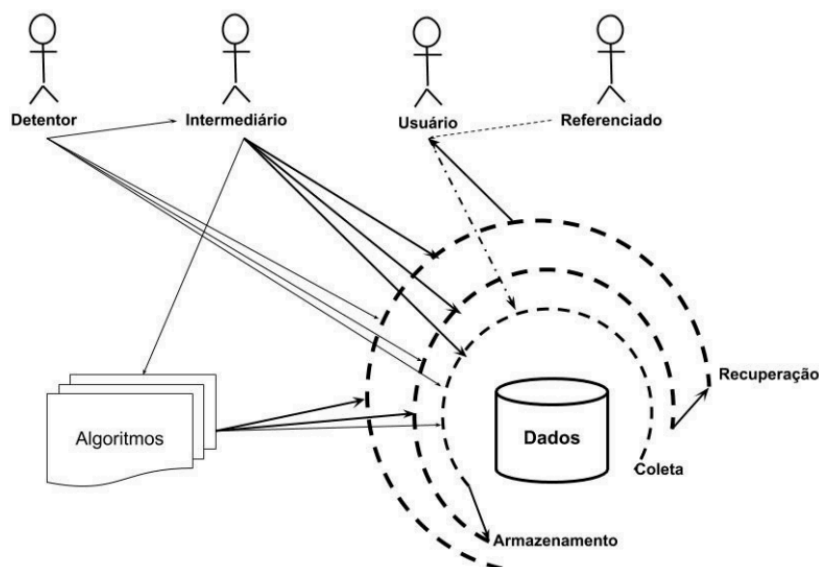
Sant’Ana (2020, p. 118) procura resgatar este percurso a fim de apontar para a tendência predominantemente funcional dos próximos anos, uma vez que o desenvolvimento

do *big data* traz um deslocamento cada vez maior das capacidades cognitivas humanas no processo de interpretação e um crescente grau de abstração. A minimização da participação humana torna-se uma questão preocupante ao refletir quem são os devidos atores responsáveis pelo acesso aos dados e o desconhecimento de parte dos integrantes. É como coloca o autor no seguinte parágrafo:

A minimização da participação humana no processo de interpretação, agregada à questão da insciência dos atores, inerente à transdução informacional envolvendo um número crescente de camadas, proporcionam um alto grau de encapsulamento da complexidade do processo de acesso aos dados como um todo. Esse encapsulamento nos processos de construção de “como” serão coletados, armazenados e recuperados os “dados” acrescenta maior importância à identificação dos principais atores envolvidos. Usuários passam a depender da atuação dos detentores, nem sempre perceptíveis, mas que podem, intencionalmente ou não, impor vieses aos resultados gerados pelos processos automatizados de acordo com suas próprias agendas. (SANT’ANA, 2020, p. 118)

O “encapsulamento” citado trata-se de uma limitação do acesso direto ao código, ou seja, o método operado no tratamento dos dados. Sant’Ana (2020, p. 119) indica que, para que exista uma relação de confiança com o conteúdo, é importante que haja a percepção de “quem tem a responsabilidade de definir as diretrizes e estabelecer quais procedimentos serão realizados”. As diretrizes e os procedimentos seriam a dimensão função dentro do processamento de dados. No entanto, para isso, é necessário entender quem exatamente são os atores envolvidos no processo e o papel de cada um. Para ilustrar, Sant’Ana utiliza a seguinte imagem:

Figura 2 - “Atores e o Ciclo de Vida dos Dados”.



Fonte: SANT'ANA, 2020, p. 120.

O “usuário” é aquele que utiliza do recurso a fim de atender suas necessidades e que, ao interagir com o sistema (tanto de forma direta quanto indireta), torna-se fonte de dados, podendo ser consciente disso ou não (SANT'ANA, 2020, p. 119).

O “referenciado” é definido pelo o que a legislação brasileira coloca como “pessoa natural a quem se referem os dados pessoais que são objeto de tratamento” (BRASIL apud SANT'ANA, 2020, p. 119). O referenciado não necessariamente entra em contato com o ciclo de vida dos dados, e por isso, pode não ter consciência de sua relação com eles. Esse também pode ser o usuário, caso haja vinculação dos dados a ele (SANT'ANA, 2020, p. 119).

O “detentor” é “aquele que responde à posse e controle dos recursos necessários para a instânciação e manutenção dos suportes e sistemas que irão coletar, armazenar e permitir a recuperação dos dados” (SANT'ANA, 2020, p. 119). É quem define e decide quais serão os objetivos e estratégias para o tratamento dos dados e seu ciclo de vida. Na constituição, é colocado como o “controlador”, podendo ser “pessoa natural ou jurídica, de direito público ou privado” (BRASIL apud SANT'ANA, 2020, p. 119).

No entanto, por este detentor estar envolvido com trabalhos multidisciplinares, frequentemente ele conta com um “intermediário” responsável pelo desenvolvimento e operacionalização dos sistemas. Sant'Ana (2020, p. 119-120) aponta que, na constituição, este corresponde ao que é colocado como “pessoa natural ou jurídica, de direito público ou

privado, que realiza o tratamento de dados pessoais em nome do controlador”, e também adiciona que o intermediário é “pessoa indicada pelo controlador e operador para atuar como canal de comunicação entre o controlador, os titulares dos dados e a Autoridade Nacional de Proteção de Dados (ANPD)” (BRASIL apud SANT’ANA, 2020, p. 119-120).

Estes atores são importantes para entender o papel e a participação que cada um deles possui nas diferentes camadas das fases de acesso aos dados. Sant’Ana (2020, p. 120) apresenta três fases, como mostrado na figura 2: recuperação, armazenamento e coleta, sendo a fase da recuperação a mais externa e disponibilizada pelos recursos digitais. Os usuários possuem acesso à fase da coleta, contudo, possuem baixa capacidade de controle. Os referenciados, como citado anteriormente, não necessitam de interação prévia para que seus dados sejam utilizados. E o detentor, como diz Sant’Ana, “estabelece uma relação de definição e controle sobre todas as fases do ciclo de vida dos dados” (SANT’ANA, 2020, p. 120), principalmente devido à relação de autoridade que o mesmo possui com o intermediário, responsável pelo acesso de todas as fases do sistema, uma vez que são estes os que as “desenvolvem, instanciam e operacionalizam” (SANT’ANA, 2020, p. 121).

Entendendo esse esquema, é possível visualizar a questão em torno da restrição de acesso e a crescente presença da dimensão funcional do tratamento dos dados. Tal questão é acentuada pela “especificidade das competências necessárias para essa interação” (SANT’ANA, 2020, p. 121), cabendo ao intermediário responder por essas competências. No entanto, os intermediários estão sob a autoridade dos detentores, sendo eles a quem cabe o “estabelecimento de quais serão os objetivos e diretrizes dos sistemas” (SANT’ANA, 2020, p. 121).

Assim, já não é passível de verificação e validação, por parte dos demais atores, o real estabelecimento de “como” os dados serão coletados, tratados, armazenados e disponibilizados. O detentor define; os intermediários desenvolvem, e “cabe” aos usuários e à sociedade como um todo, a confiança no que é declarado por eles ou, quando muito, a validação por meio de busca de indícios a partir da fase de recuperação [...] do que realmente foi feito no ciclo de vida dos dados. (SANT’ANA, 2020, p. 121)

É possível notar que tal distanciamento coloca o conjunto de dados longe do alcance dos usuários e, principalmente, dos referenciados. Sant’Ana chama esse distanciamento de “transdução informacional”, que busca “explicitar as camadas de abstração oriundas das diversas transformações a que a informação está sujeita quando utilizada nos suportes digitais” (SANT’ANA, 2019, p. 117), e também explicita “seus impactos na percepção dos envolvidos a partir da opacidade gerada por estas transformações” (SANT’ANA, 2019, p. 117). Com isso, há uma tendência a dificultar a proteção e a garantia dos direitos à privacidade desses atores. Além disso, dentro deste conceito, existe outro complicador: a algoritmização.

Sant’Ana argumenta que, devido à crescente substituição do trabalho humano, qualquer tarefa “minimamente repetitiva e passível de explicitação” pode ser mapeada e tende a ser substituída pelos processos de automatização (SANT’ANA, 2020, p. 122). A deterioração da participação dos atores é capaz de afetar inclusive aos intermediários, conseqüentemente afetando também os detentores, uma vez que os conjuntos de códigos que tratam do *big data* têm se tornado cada vez mais complexos e difíceis de compreender. A situação é agravada pela atuação de áreas como o *Machine Learning* e a Inteligência Artificial, responsáveis pela automatização do próprio processo de codificação e desenvolvimento do algoritmo. Desta forma, Sant’Ana aponta para uma transformação no processo de participação, como pode ser compreendido no seguinte fragmento:

“Vivenciamos, assim, a substituição mecânica de etapas no processo de informação chegando às camadas superiores da participação humana, gerando situações em que os próprios algoritmos encontram soluções para os problemas. Nessas situações, nos cabe somente imaginar como foram encontradas tais soluções por meio de indícios exteriores ao processo, os intermediários se transformando em usuários, ou seja, passando de condutores a passageiros.” (SANT’ANA, 2020, p. 123)

Abordar esta questão é essencial para entender as dificuldades enfrentadas atualmente em relação ao gerenciamento e domínio dos dados obtidos através da sua mineração. O crescente grau de abstração e o distanciamento da participação mecânica que acompanham o desenvolvimento do *big data* podem tornar os processos de coleta, armazenamento e recuperação de dados cada vez mais obscuros. Além disso, também criam um potencial de

aceleração do desequilíbrio entre os diversos atores que fazem parte desse processo e a retroalimentação da concentração de poder entre os detentores (SANT'ANA, 2019, p. 30).

3.2 O direito à privacidade de dados

No texto de Deborah Johnson intitulado *Computer Ethics* (Ética Computacional), a autora coloca que desde a Segunda Guerra Mundial ocorre uma crescente preocupação no potencial do uso dos computadores para centralização e concentração de poder por governos. Acompanhado desse inquietamento, encontra-se também o receio em torno da expansão do uso da tecnologia para depósito e arquivamento de informações, gerando um crescimento exponencial sem precedentes na escala das bases de dados (JOHNSON, 2004, p. 65).

Em seu texto, Johnson procura apontar algumas questões tradicionais e emergentes concernentes à ética da computação na sociedade da informação. O termo “sociedade da informação” se refere a sociedades que passaram por transformações significativas nas atividades humanas e instituições devido ao avanço da computação e tecnologias da informação (JOHNSON, 2004, p. 68). Entre essas questões, é colocado o problema da privacidade (JOHNSON, 2004, p. 70). Como mencionado no parágrafo anterior, ao longo das últimas décadas foi possível notar um crescimento inimaginável do poder de coleta, armazenamento, retenção e troca de dados, que são obtidos em grande parte por informações sobre indivíduos. O desenvolvimento da computação tornou possível que essa coleta ocorresse durante diversas atividades usuais do dia-a-dia, como por exemplo ao utilizar um cartão de crédito, navegar na web ou realizar ligações.

Uma preocupação recorrente é que essa constante coleta de informações crie um efeito panóptico, podendo transformar a sociedade da informação em uma sociedade de vigilância. Apesar de não ter a mesma estrutura física do panóptico criado no design de Jeremy Bentham, a computadorização cria algo similar ao tornar possível a observação constante do comportamento humano. Com isso, assim como no panóptico, o risco à privacidade causado pela coleta excessiva de dados pode levar os indivíduos a mudar seu comportamento ao saberem que estão sendo observados por medo de sofrerem consequências. Esse tipo de situação pode ser levada a um ponto que fere a liberdade de agir e pensar criticamente, colocando em risco a democracia (JOHNSON, 2004, p. 71).

Seria possível argumentar que, pelo fato da coleta de dados acontecer de forma quase invisível, isso poderia diminuir a possibilidade da ocorrência do efeito panóptico. No entanto, esse efeito pode acontecer através da tomada da consciência pelos indivíduos conforme os efeitos e resultados dessa coleta forem se apresentando aos olhos do usuário. Ou seja, por meio da habituação das interferências da coleta de dados nas atividades cotidianas, o indivíduo pode se sentir observado ao notar de que maneira “as informações coletadas em vários lugares são reunidas e usadas para tomar decisões que afetam suas interações”, como ocorre por exemplo nas interações com instituições tais quais “agências governamentais, agências de crédito, companhias de seguros, instituições educacionais, empregadores, etc.”⁵⁵ (JOHNSON, 2004, p. 71, tradução nossa).

Essa preocupação torna necessária a interferência da esfera política para a criação de leis com o propósito de regular e limitar o recolhimento e utilização dos dados pessoais. Como tudo que envolve o âmbito legislativo, cada país possui sua própria abordagem. Algumas abordagens são mais fragmentadas, como ocorre com a estadunidense, que separa uma legislação diferente para cada tipo de registro (isto é, registros médicos, históricos de emprego, registros de crédito), enquanto vários países europeus possuem “políticas abrangentes que especificam que tipo de informação pode ser recolhida e sob que condições em todos os domínios”⁵⁶ (JOHNSON, 2004, p. 71, tradução nossa).

No entanto, Johnson ressalta que existe uma grande pressão da intensificação dos negócios globais no debate político. Apesar das organizações de armazenamento de informações prometerem que os dados de seus usuários só serão utilizados de uma determinada maneira (agindo conforme a legislação de proteção de dados do próprio país), os dados podem fluir para outras regiões que não possuam a mesma proteção ou que não tenha proteção nenhuma devido à natureza globalizada dessa interação. Ou seja, essas organizações podem “prometer tratar a informação de uma determinada forma e depois enviá-la para o estrangeiro, onde é tratada de uma forma completamente diferente, quebrando assim a promessa feita ao titular dos dados”⁵⁷ (JOHNSON, 2004, p. 71, tradução nossa). Esse risco

⁵⁵ “They may come to see how information gathered in various places is put together and used to make decisions that affect their interactions with government agencies, credit bureaus, insurance companies, educational institutions, employers, etc.”.

⁵⁶ “[...] whereas several European countries have comprehensive policies that specify what kind of information can be collected under what conditions in all domains.”.

⁵⁷ “An information-gathering organization might promise to treat information in a certain way, and then send the information abroad where it is treated in a completely different way, thus breaking the promise made to the data subject.”.

intensifica a necessidade da elaboração de acordos internacionais a respeito do fluxo de dados de forma que sua segurança possa ser estendida para além dos limites fronteiriços.

Floridi, no livro “A Quarta Revolução – Como a Infosfera está remodelando a realidade humana” (*“The Fourth Revolution - How the Infosphere is Reshaping Human Reality”*), vai além ao pensar a respeito da privacidade na sociedade da informação, trazendo uma reflexão ainda mais profunda concernente à identidade humana.

Inicialmente, ele apresenta duas teorias que tratam a respeito do valor da privacidade: a interpretação reducionista e a interpretação baseada na propriedade.

A interpretação reducionista se baseia nas consequências indesejáveis que podem ser causada pela violação, “seja pessoalmente, como a angústia, ou socialmente, como a injustiça”⁵⁸ (FLORIDI, 2014, p. 116, tradução nossa). A privacidade é colocada como uma utilidade, seja para preservar a dignidade, assim mantendo uma condição de boas interações humanas, assim como também com o objetivo de garantir o controle e o equilíbrio político. Nessa abordagem, o valor da privacidade é definido por meio da avaliação consequencialista da análise de custo-benefício de sua proteção ou violação.

A interpretação baseada na propriedade, apesar de não ser incompatível com essa perspectiva, ressalta um aspecto diferente no valor da privacidade. Dentro dessa abordagem, é importante que a privacidade informacional seja preservada em virtude ao direito de cada pessoa à segurança comportamental e a propriedade, “em que a ‘propriedade de x’ é classicamente entendida como o direito ao uso exclusivo de x”⁵⁹ (FLORIDI, 2014, p. 116, tradução nossa). A pessoa possui a posse das suas próprias informações, portanto, deve ter o direito de controlar todo o seu ciclo de vida, “desde a geração até o apagamento por meio do uso”⁶⁰ (FLORIDI, 2014, p. 116, tradução nossa). Essa interpretação é voltada para uma compreensão em termos de propriedade privada ou intelectual, ou seja, de um direito juridicamente reconhecido do valor da privacidade em si.

Floridi aponta que ambas pertencem a uma “mentalidade histórica” e devido à isso, na perspectiva das duas interpretações, a violação de privacidade é frequentemente comparada à uma “transgressão ou invasão não autorizada” de um espaço metafórico das informações

⁵⁸ “[...] either personally, such as distress, or socially, such as unfairness.”

⁵⁹ “[...] where ‘property of x’ is classically understood as the right to exclusive use of x.”

⁶⁰ “[...] from generation to erasure through usage.”

personais⁶¹ (FLORIDI, 2014, p. 116-7), cuja “acessibilidade e uso devem ser totalmente controlados por seu proprietário e, portanto, mantidos privados”⁶² (FLORIDI, 2014, p. 117, tradução nossa). No entanto, o autor não considera nenhuma das duas abordagens totalmente satisfatórias.

Apesar da interpretação reducionista ser razoável ao focar na preocupação pelo respeito da privacidade considerando o uso incorreto das informações adquiridas, ela pode ser inconsistente ao tratar “a busca e a promoção dos interesses sociais e do bem-estar”. Floridi diz que, embora algumas informações pessoais públicas precisam ser protegidas (com o objetivo de impedir seu uso indevido como acontece por meio da “criação de perfis ou a vigilância eletrônica irrestrita”⁶³) (FLORIDI, 2014, p. 117, tradução nossa), uma base puramente reducionista não seria capaz de ir contra a ideia de que, visando uma maior importância ao bem estar comum, não seria melhor uma sociedade totalmente desprovida de privacidade.

“Justamente por causa de considerações reducionistas, mesmo em sociedades democráticas, tendemos a reconhecer que o direito à privacidade pode ser anulado quando outras preocupações e prioridades, inclusive a segurança pública ou a segurança nacional, tornam-se mais urgentes.”⁶⁴ (FLORIDI, 2014, p. 117, tradução nossa)

Já a interpretação baseada na propriedade falha por três razões. Primeiramente, devido à possibilidade de informações alheias serem obtidas de forma não intencional. Existem situações em que a privacidade pode ser passivamente violada, como por exemplo quando uma pessoa fala no telefone em voz alta, no entanto nenhum direito de posse informacional está sendo violado neste caso. Em outras palavras, nenhuma atitude violadora está sendo ativamente tomada contra o direito de propriedade. Em segundo lugar, existe um problema que concerne a privacidade em contextos públicos. A privacidade pode ser exercida em lugares em que ela é socialmente, fisicamente e informacionalmente compartilhada. Em uma praça pública, qualquer um pode ver o que é feito por outro indivíduo. Nesses casos, como

⁶¹ O autor diz, mais especificamente, da seguinte maneira: “[...] *They both compare 'privacy breach' to trespass or unauthorized invasion of, or intrusion in, a metaphorical space or sphere of personal information.*”

⁶² “[...] *whose accessibility and usage ought to be fully controlled by its owner and hence kept private.*”

⁶³ “[...] *especially against profiling or unrestrained electronic surveillance.*”

⁶⁴ “*Precisely because of reductionist-only considerations, even in democratic societies we tend to acknowledge that the right to privacy can be overridden when other concerns and priorities, including public safety or national security, become more pressing.*”

poderia o sistema de vigilância público ser considerado uma violação da privacidade de alguém, se este estiver acessando um local que já é público em todos os sentidos da palavra? Nesse sentido, qual seria a diferença entre ser visto por uma outra pessoa ou por uma câmera? Floridi (FLORIDI, 2014, p. 117-8) aponta que a interpretação baseada na posse não consegue trazer uma resposta satisfatória a essa questão. Em último lugar, existe um problema no uso impreciso do conceito metafórico de “propriedade de informação”, que falha ao tentar explicar o uso ou aquisição de uma informação sem ter ligação a uma perda. Ao contrário da maioria das coisas que são propriedades de um indivíduo, as informações pessoais não são perdidas quando adquiridas por alguém. “As análises de privacidade baseadas na ‘propriedade’ de um ‘espaço informacional’ são duplamente metafóricas”⁶⁵ (FLORIDI, 2014, p. 118, tradução nossa).

Para Floridi, assim como é necessário pensar na revolução informacional como uma quarta revolução da autocompreensão humana, o conceito de privacidade também precisa ser reinterpretado de uma forma igualmente radical levando em conta “a natureza informacional de nós mesmos e de nossas interações enquanto *inforgs*”⁶⁶ (FLORIDI, 2014, p. 119, tradução nossa). Ou seja, ao pensar no humano enquanto uma entidade composta de informação, é necessário colocar a questão da privacidade informacional em uma nova perspectiva.

Enquanto cada indivíduo é constituído da sua própria informação, uma violação da privacidade deste indivíduo pode ser considerada uma forma de agressão à sua própria identidade pessoal. De acordo com o autor, a visão da privacidade informacional enquanto tendo um valor autoconstitutivo é compatível com a ideia de que as Tecnologias de Informação e Comunicação (TIC) possuem potencial tanto para desgastar quanto para reforçar a privacidade informacional. Portanto, Floridi argumenta que “o valor da privacidade deve ser defendido e reforçado”⁶⁷ (FLORIDI, 2014, p. 119, tradução nossa), sendo necessário não só apoiar tecnologias que aumentam a privacidade como também aplicações construtivas, que podem “permitir aos usuários projetar, moldar e manter suas identidades como agentes informacionais”.⁶⁸

⁶⁵ “Analyses of privacy based on ‘ownership’ of an ‘informational space’ are metaphorical twice over.”

⁶⁶ “[...] one that takes into account the informational nature of our selves and of our interactions as *inforgs*.”

⁶⁷ “The value of privacy is both to be defended and enhanced.” p. 119.

⁶⁸ “[...] and hence that a positive effort needs to be made in order to support not only Privacy Enhancing Technologies but also constructive applications, which may allow users to design, shape, and maintain their identities as informational agents.” p. 119.

É necessário a existência de um “atrito” no fluxo de informações para que a distinção entre “o sistema macro multiagente (a sociedade) e a identidade do sistema micro multiagente (os indivíduos) que o constitui”⁶⁹ (FLORIDI, 2014, p. 119, tradução nossa) seja mantida. No seguinte trecho, Floridi ressalta a importância dessa distinção:

Qualquer sociedade (mesmo utópica) em que não haja privacidade informacional é aquela em que nenhum processo de autoconstituição pode ocorrer, nenhuma identidade pessoal pode ser desenvolvida e mantida e, portanto, nenhum bem-estar social pode ser alcançado, sendo o bem-estar social apenas a soma dos indivíduos envolvidos.⁷⁰ (FLORIDI, 2014, p. 119, tradução nossa)

A ideia de uma sociedade totalmente desprovida de privacidade trazida pela interpretação reducionista, onde existe uma “total transparência” da esfera informacional, de fato poderia conseguir obter uma proteção social ao “apagar” todas identidades pessoais e individualidades, no entanto, dificilmente aqueles indivíduos que compõem essa sociedade ficariam satisfeitos em viver desta maneira.

A interpretação autoconstitutiva do valor da privacidade também resolve os três problemas da interpretação baseada na propriedade (FLORIDI, 2014, p. 120-121):

Relacionado a primeiro problema, Floridi mostra que, ao considerar a natureza de uma pessoa como constituída por sua informação, é possível entender o direito à privacidade como “um direito à imunidade pessoal contra mudanças desconhecidas, indesejadas ou não intencionais em sua própria identidade como uma entidade informacional”⁷¹ (tradução nossa). Essa interpretação sugere que a “esfera informacional e sua identidade pessoal são co-referenciais, ou dois lados da mesma moeda”⁷² (tradução nossa). Aquilo que é feito com a informação de um indivíduo afeta a sua identidade, e não as suas posses. Isso ocorre tanto de forma passiva (ao forçar alguém a receber dados não desejados) quanto de forma ativa (como coletando, armazenando ou reproduzindo os dados de alguém contra a sua vontade). Enquanto entidade informacional, na forma passiva, essa violação pode ser considerada uma tentativa

⁶⁹ “[...] *the macro multi-agent system (the society) and the identity of the micro multi-agent systems (the individuals) constituting it.*” p. 119.

⁷⁰ “*Any society (even a utopian one) in which no informational privacy is possible is one in which no self-constituting process can take place, no personal identity can be developed and maintained, and hence no welfare can be achieved, social welfare being only the sum of the individuals' involved.*”

⁷¹ “[...] *a right to personal immunity from unknown, undesired, or unintentional changes in one's own identity as an informational entity.*”

⁷² “[...] *your informational sphere and your personal identity are co-referential, or two sides of the same coin.*”

de manipulação. Na forma ativa, pode ser considerada um roubo ou clonagem da própria identidade do indivíduo. Essa perspectiva atribui um alto valor à privacidade devido ao seu papel de proteger a identidade informacional, e é importante que ela seja respeitada.

O segundo problema pode ser resolvido uma vez que, nessa perspectiva, a violação da privacidade informacional está mais próxima de ser considerada um tipo de sequestro do que de uma invasão de propriedade. Com isso, essa interpretação está livre da falsa dicotomia entre público e privado com relação à privacidade que está presente na interpretação baseada na propriedade. Isso ocorre uma vez que as informações que constituem a identidade de uma pessoa é mantida independentemente se ela está inserida em um contexto público ou privado. Logo, se uma pessoa em um contexto público não se sente confortável em ter suas informações recolhidas, armazenadas e utilizadas contra a sua vontade, ela tem total direito a isso, uma vez que concerne a privacidade de sua pessoa enquanto um ente informacional.⁷³ (tradução nossa)

O último problema é resolvido ao pensarmos que, após reivindicar a ideia de que alguém “possui” uma informação e compreendê-la como algo que constitui um agente, não é mais uma ideia metafórica de propriedade que está em questão. Nesse caso, o agente é a sua informação. Isso muda totalmente o ponto central da argumentação que é colocada contra a interpretação com base na posse. “‘Seu’ em ‘sua informação’ não é o mesmo ‘seu’ de ‘seu carro’, mas sim o mesmo ‘seu’ de ‘seu corpo’, ‘seus sentimentos’, ‘suas memórias’, ‘suas ideias’, ‘suas escolhas’ e assim por diante”⁷⁴ (FLORIDI, 2014, p. 121, tradução nossa). Não é mais sobre um senso de propriedade exterior, mas sim de pertencimento constitutivo. É sobre aquilo que é parte do indivíduo, e não aquilo que ele possui.

Através da interpretação autoconstitutiva da privacidade da informação, é possível questionar empresas como Google a respeito de suas políticas de privacidade. Aqui, Floridi cita Kevin Bankston, advogado da equipe da Electronic Frontier Foundation:

⁷³ É possível que essa questão seja um pouco mais difícil de entender, uma vez que a seguinte questão pode surgir: num contexto público, uma pessoa me observa passivamente e com isso “absorve” minhas informações. Nessa perspectiva, seria isso algum tipo de sequestro da minha identidade? Vale ressaltar que, ao falar de sistemas públicos de vigilância, faz mais sentido compará-lo com um *stalker* do que com um mero observador, devido à quantidade de informação que pode ser obtida e armazenada sem consentimento. Utilizando da interpretação autoconstitutiva da privacidade, a pessoa que é observada possui total direito em querer manter sua integridade.

⁷⁴ “Your’ in ‘your information’ is not the same ‘your’ as in ‘your car’ but rather the same ‘your’ as in your body, your feelings’, ‘your memories’, ‘your ideas, ‘your choices’, and so forth.”

Seu histórico de pesquisa mostra suas associações, crenças e talvez seus problemas médicos. As coisas que você procura no Google definem você [...] dados que são praticamente uma impressão do que está acontecendo em seu cérebro: O que você está pensando em comprar, com quem você conversa, sobre o que você fala.⁷⁵ (BANKSTON apud FLORIDI, 2014, p. 122, tradução nossa)

Com a popularização da interpretação autoconstitutiva, as perspectivas de privacidade consideradas “industriais, modernas ou Newtonianas” podem ser superadas, tornando possível uma maior problematização de questões como por exemplo a venda não autorizada de dados.

Floridi (2014, p. 115) aponta que novas TICs têm sido desenvolvidas não apenas para reunir e processar dados, mas também para aumentar o controle e a proteção. Numa perspectiva de geração de dados, “TICs digitais podem promover a proteção de dados pessoais, especialmente por meio de criptografia, anonimato, codificação de senha, firewall, protocolos ou serviços especificamente concebidos e, no caso de dados capturados externamente, sistemas de alerta”⁷⁶ (FLORIDI, 2014, p. 115, tradução nossa). Na perspectiva de armazenamento de dados, legislações já aprovadas podem garantir que “nenhum atrito informacional, já removido pelas TICs digitais, seja sorrateiramente reintroduzido” (FLORIDI, 2014, p. 115, tradução nossa), a fim de prevenir que “os agentes descubram a existência de registros de dados pessoais e que os acessem, verifiquem sua precisão, corrijam ou atualizem, ou exijam sua exclusão”⁷⁷ (FLORIDI, 2014, p. 115, tradução nossa). Na perspectiva de gerenciamento de dados, “as TICs digitais podem ajudar os agentes a controlar e regular o uso de seus dados, facilitando a identificação e a regulamentação dos usuários relevantes envolvidos”⁷⁸ (FLORIDI, 2014, tradução nossa).

O autor apresenta essas diversas modalidades de TICs com o objetivo de mostrar como o empenho para a proteção da privacidade está cada vez maior, porém ressalta que esses artifícios não foram desenvolvidos subitamente e por conta própria, mas sim por meio de esforços da parte da quem problematiza esses sistemas e demanda por soluções. “As TICs

⁷⁵ “Your search history shows your associations, beliefs, perhaps your medical problems. The things you Google for define you [...] data that's practically a printout of what's going on in your brain: What you are thinking of buying, who[m] you talk to, what you talk about.”

⁷⁶ “[...] digital ICTs can foster the protection of personal data, especially by means of encryption, anonymization, password encoding, firewalling, specifically devised protocols or services, and, in the case of externally captured data, warning systems.”

⁷⁷ “[...] digital ICTs make possible legislation, such as the Data Protection Directive already passed by the EU in 1995, which can guarantee that no informational friction, already removed by digital ICTs, is surreptitiously reintroduced to prevent agents from discovering the existence of personal data records, and from accessing them, checking their accuracy, correcting or upgrading them, or demanding their erasure.”

⁷⁸ “[...] digital ICTs can help agents to control and regulate the usage of their data by facilitating the identification and regulation of the relevant users involved.”

digitais não necessariamente corroem a privacidade; elas também podem aprimorá-la e protegê-la”⁷⁹ (FLORIDI, 2014, tradução nossa).

3.3 O lucro da análise de dados

Os *big datas* são associados em grande frequência como “o novo petróleo” do século XXI, e tal qual a *commodity*, possui um valor fundamental em sua extração (SCHAFER, 2016, p. 1). Para os entusiastas de tecnologia e grandes empresários, existe um enorme potencial na exploração e aproveitamento de dados, tal qual teve o petróleo na revolução industrial.

No entanto, Burkhard Schafer (2016, p. 2) ressalta que, ao fazer essa associação, é importante também levar em consideração os efeitos negativos trazidos pelo petróleo no passado para evitar que os mesmos se repitam com a mineração de dados. Se o *big data* é o novo petróleo, logo, também é necessário pensar na possibilidade de danos ao ambiente advindos de questões como vazamentos massivos, danos ambientais causados por métodos irresponsáveis de extração, deslocamentos forçados de comunidades indígenas, entre outros. Além disso, a combinação de dinheiro e poder econômico detidos por um pequeno número de grandes monopólios indica uma possível ameaça ao sistema democrático. Como é colocado por Schafer, “para todos esses perigos ‘antigos’ que vieram com o setor fóssil, podemos facilmente encontrar perigos equivalentes na economia baseada em dados”⁸⁰. (SCHAFER, 2016, p. 2, tradução nossa)

A comparação do *big data* enquanto novo petróleo serve para ilustrar o poder potencial existente no domínio e extração de dados. Shoshana Zuboff, em seu artigo “*Big Other*: capitalismo de vigilância e perspectivas para uma civilização de informação”, procura explorar o poder dos dados na sociedade da informação e vai ainda mais fundo ao tratar o *big data* como parte fundamental de nova lógica de acumulação presente no chamado “capitalismo de vigilância”.

Em seu artigo, Zuboff traça primeiramente uma trajetória histórica onde procura apresentar a emergência das tecnologias da informação na mediação do trabalho, onde

⁷⁹ “Digital ICTs do not necessarily erode privacy; they can also enhance and protect it.”

⁸⁰ “For all of these ‘old’ dangers that came with the fossil fuel-based industry, we can easily find equivalent dangers in the data-driven economy.”

observa que o local de trabalho cada vez mais informatizado torna o aprendizado a nova forma de trabalho. A autora coloca que, hoje em dia, é difícil imaginar épocas em que o papel das TICs nesses tipos de mediações trabalhistas não eram normas. “O aprendizado em tempo real, baseado em informação e mediado pelo computador” (ZUBOFF, 2018, p. 21), se tornou tão fundamental para o cotidiano dos negócios que os dois domínios agora se entrelaçam, sendo grande parte daquilo que é costumeiramente feito no trabalho. A autora cita exemplos de ações mediadas pelas TICs já institucionalizadas dentro das organizações, sendo algumas mais formais (“metodologias de aperfeiçoamento contínuo, integração empresarial, monitoramento de empregados, [...] cadeias de fornecedores, projetos inter-empresas”, etc) e outras informais (“o fluxo incessante de mensagens eletrônicas, buscas online, atividades no *smartphone*, [...] videoconferências”, etc) (ZUBOFF, 2018, p. 21).

Zuboff aponta que “cada época da história do capitalismo rumou em direção a uma lógica de acumulação dominante” (ZUBOFF, 2018, p. 22). O capitalismo teve seu “sucesso” como sistema econômico devido às recorrentes emergências de novas formas de mercado, expressando conseqüentemente novas formas de acumulação, sempre indo de acordo com aquilo consiga satisfazer “as necessidades sempre em evolução das populações e sua expressão na natureza cambiante da demanda” (ZUBOFF, 2018, p. 23). Essas lógicas são constituídas por uma natureza que se mostra sempre dinâmica, assim como afirma Piketty: “Não existe uma variante única do capitalismo ou da organização da produção [...]. Isto sem dúvida continuará valendo para o futuro, pois novas formas de organização e de propriedade ainda estão para ser inventadas” (PIKETTY apud ZUBOFF, 2018, p. 23). Diferentes formas de mercado surgem em variados momentos e contextos. Algumas alcançam a posição hegemônica, outras coexistem de forma paralela à configuração predominante. Algumas, ao longo do tempo, se revelam becos sem saída. (ZUBOFF, 2018, p. 23)

A autora procura criar uma relação entre essas duas questões para falar a respeito do domínio que existe e tem aumentado da computação ubíqua, que hoje em dia se faz presente em quase todos os momentos, incluindo no horário fora do expediente. “O mundo renasce como dados e o texto eletrônico é universal em escala e escopo”. Para ilustrar a crescente constante da quantidade de dados no mundo, Zuboff (2018, p. 24) apresenta em números: em 1986, foi calculado um total de aproximadamente 2,5 exabytes de informação comprimida e desses, apenas 1% existia na forma digital. No ano 2000, o número era 5%. Já em 2007, foram calculados 300 exabytes comprimidos, em que 94% eram digitalizados. Atualmente,

A digitalização e a dataficação (o programa que permite a computadores e algoritmos processar e analisar dados brutos), junto com o desenvolvimento de tecnologias de armazenamento mais baratas, produziram 1.200 exabytes de dados armazenados no mundo com uma taxa de 98% de conteúdo digital. (MAYER-SCHÖNBERGER E CUKIER apud ZUBOFF, 2018, p. 24)

Nesse contexto de uma sociedade informacional, ou “civilização informacional”, são direcionadas novas questões concernentes a poder e autoridade, cujas respostas devem moldar o caráter dessa civilização ao longo do século: “Quem aprende com os fluxos de dados globais, como e o quê? Quem decide? O que acontece quando a autoridade falha? Qual lógica de acumulação moldará as respostas a essas perguntas?” (ZUBOFF, 2018, p. 24)

De acordo com Zuboff, a sociedade atual vive uma lógica emergente de acumulação denominada “capitalismo de vigilância, da qual o *big data* é tanto uma condição quanto uma expressão” (ZUBOFF, 2018, p. 25). O Google é considerado um pioneiro dessa lógica de acumulação, lógica que se tornou o modelo-padrão para o que é buscado pela maior parte das novas startups e aplicativos.

Documentos do economista-chefe da Google, Hal Varian, podem ser úteis para refletir as práticas de negócio da empresa e são usados pela autora como exemplares para entender o funcionamento dessa nova lógica de acumulação. Nos dois artigos analisados pela autora, Varian procura tratar principalmente a universalidade das “transações econômicas mediadas por computador” (VARIAN apud ZUBOFF, 2018, p. 25), onde ele explica da seguinte maneira: “O computador cria um registro da transação [...]. Eu argumento que essas transações mediadas por computador permitiram melhorias significativas na forma como as transações são realizadas e continuarão a impactar a economia no futuro que prevemos” (VARIAN apud ZUBOFF, 2018, p. 25-26). Nesse funcionamento mediado pelo computador é importante observar que, por meio da informatização da economia, cada transação possui um registro persistente e contínuo de cada detalhe, o que a tornaria “transparente e cognoscível”, um contraste ao ideal neoliberal de um mercado visto como “inefável e incognoscível”. Diante deste mercado cognoscível, Zuboff apresenta quatro novos usos nas transações mediadas por computador colocados por Varian: “‘extração e análise de dados’, ‘novas formas contratuais devido à um melhor monitoramento’, ‘personalização e customização’ e ‘experimentos contínuos’” (VARIAN apud ZUBOFF, 2018, p. 26). Cada um deles é capaz de trazer diferentes *insights* sobre a nova lógica de acumulação emergente, assim como também

a respeito da divisão de aprendizagem formada por essa lógica e o caráter conduzido por ela na civilização da informação.

Nesta seção do trabalho, será tratado especificamente a respeito da extração e análise de dados enquanto nova ferramenta de negócios.

No texto, são apresentadas diversas fontes de obtenção do *big data* (ZUBOFF, 2018, p. 27-28). Além das transações, uma clássica fonte de extração de dados é por meio dos bancos de dados governamentais e corporativos, como aqueles associados a bancos, agências de avaliação de crédito, planos de saúde, registros censitários, entre outros.

Atualmente, uma terceira forma de obtenção de dados tem tido um crescimento exponencial, que acontece por meio da incorporação de sensores em diversos tipos de “objetos, corpos e lugares”, parte da já citada área da computação chamada internet das coisas. Zuboff cita como exemplo os investimentos da Google em “drones, dispositivos vestíveis, carros automatizados, nanopartículas que patrulham o corpo procurando por sinais de doenças e dispositivos inteligentes para o monitoramento do lar”, dispositivos conectados a internet e destinados a serem parte de uma “infraestrutura” inteligente. Outra fonte de uma quantidade massiva de dados acontece por meio de sistemas de vigilância, não apenas através de câmeras públicas ou privadas, como também através da localização de *smartphones* e satélites, incluindo como ocorre em serviços como *Google Earth* e *Google Street View*.

Zuboff ressalta que essas perspectivas representam apenas o lado da “‘oferta’ da interface mediada por computador” (ZUBOFF, 2018, p. 30), mas a universalização dessas mediações só ocorre devido a um processo de causalidade envolvendo atividades subjetivas, que constituem a parte da “demanda” de tais mediações. As interações e necessidades individuais fazem parte de uma outra parte das principais fontes de *big data*, constituído além dos espaços de trabalho, onde os processos da internet acontecem longe do monitoramento, segurança e rigidez das empresas. Aqui, se refere ao espaço da internet que existe cada vez mais dentro do cotidiano, suprimindo as “necessidades individuais de autoexpressão, voz, influência, informação, aprendizagem, empoderamento e conexão” (ZUBOFF, 2018, p. 31), necessidades que possibilitam uma “nova gama de capacidades”: músicas do Spotify, vídeos no Youtube, páginas do Facebook, publicações no Instagram, redes e comunidades de amigos. São capacidades da internet que ultrapassam todos os tipos de fronteiras, onde informações são compartilhadas para todos os tipos de propósito, servindo como expressões da

subjetividade autodeterminada, colocadas como “formas não mercantis de ‘produção social’” (BENKLER apud ZUBOFF, 2018, p. 31). Essas atividades cotidianas chamadas de *small data* são constituidoras importantes do *big data*.

Nada é trivial ou efêmero em excesso para essa colheita: as “curtidas” do Facebook, as buscas no Google, e-mails, textos, fotos, músicas e vídeos, localizações, padrões de comunicação, redes, compras, movimentos, todos os cliques, palavras com erros ortográficos, visualizações de páginas e muito mais. Esses dados são adquiridos, tornados abstratos, agregados, analisados, embalados, vendidos, analisados mais e mais e vendidos novamente. (ZUBOFF, 2018, p. 30)

Esses tipos de dados são chamados pelos especialistas de “*data exhaust*”, remetendo aos gases expelidos pelo escapamento de um carro. Ou seja, são colocados como resíduos, os dados que os usuários deixam para trás, e ao serem redefinidos dessa forma, correm menos risco de terem sua extração e monetização contestada.

O Google⁸¹ é o site mais visitado em toda web (ZUBOFF, 2018, 32), portanto, é o maior produtor de *big data* por conter uma enorme quantidade de *data exhaust*. Para satisfazer as pressões por lucro ao mesmo tempo que mantinha o site gratuito, a empresa criou, em seus primeiros anos, uma nova abordagem que dependia da “aquisição de dados de usuários como matéria-prima para análise e produção de algoritmos”, que seriam depois vendidos e que poderiam “segmentar a publicidade por meio de um modelo de leilão exclusivo, com precisão e sucesso cada vez maiores”. Isso serviu para o crescimento exponencial das receitas da empresa e o seu sucesso propiciou a explosão de uma nova “ciência de análise de *big data*”.

Esse leilão funciona através de uma plataforma de leilão algoritmo chamada AdWords, que é utilizada para vender publicidade online. Essa plataforma “analisa enormes quantidades de dados para determinar quais anunciantes comprarão cada um dos onze links publicitários em cada página de resultados de busca” (ZUBOFF, 2018, p. 32-33). Em seu texto, Zuboff cita um artigo onde Varian explica o sucesso financeiro do leilão:

“Por que a Google disponibiliza produtos de graça...? Qualquer coisa que aumente o uso da internet, em última instância, enriquece a Google [...]”. O artigo continua: “[...] mais olhos fixados na web levam inexoravelmente a mais vendas de anúncios para a Google [...]. E, como a previsão e a análise são tão cruciais para o Google

⁸¹ Aqui, assim como no texto de Zuboff, “Google” quando referido no gênero masculino, trata do site enquanto ferramenta de busca. No feminino, se refere à empresa Google.

AdWords, qualquer bit de dados, mesmo que aparentemente trivial, tem valor potencial” (VARIAN apud ZUBOFF, 2018, p. 33)

O que importa aqui é a quantidade de dados colhidos, e não a qualidade. Empresas como a Google estão sempre à procura de novas maneiras de obter dados. Quanto mais dados, maior o lucro.

Essa extração não é expressão de um relacionamento mútuo entre empresa e usuário, e sim um processo unidirecional. “A extração tem por conotação ‘tomar algo’ em vez de ‘entregar’ ou de uma reciprocidade de ‘dar e receber’” (ZUBOFF, 2018, p. 33). Apesar dos dados representarem “subjetividades de vidas individuais”, seu processo de extração só é possível por ocorrer “na ausência de diálogo ou de consentimento”. Aqui, apenas o *status* dos dados como sinais de subjetividade é considerado valioso, uma vez que serve como potencial publicitário para os anunciantes ao serem utilizados na construção de perfis detalhados. Contudo, enquanto dados agregados pelas grandes empresas, eles são apenas *bits*. “As subjetividades são convertidas em objetos que reorientam o subjetivo para a mercantilização” (ZUBOFF, 2018, p. 34). A utilização de *small data* para produzir *big data* e as formas que essa massiva quantidade de dados extraídos adquire valor “refletem a indiferença formal que caracteriza o relacionamento da empresa com suas populações de ‘usuários’”. Nesse relacionamento, não existe reciprocidade. O usuário é tanto a fonte onde os dados são extraídos como também é o “alvo final” daquilo que é produzido através deles.

Essa indiferença se expressa na agressividade utilizada por empresas como a Google para seguir os seus interesses e alcançar o lucro máximo. Um exemplo foi o que já ocorreu com o *Street View*, que a autora usa de exemplo para pontuar questões judiciais importantes concernentes à quebra dos direitos de privacidade. Desde seu lançamento em 2007, o recurso da Google lida com resistências de diversas partes da sociedade, tanto na esfera privada quanto pública. Uma delas foi em 2013, quando autoridades alemãs descobriram que seus carros estavam equipados de escâneres capazes de extrair dados de redes sem fio privadas. Concernente a interceptação de redes sem fio, Zuboff cita um processo movido por 39 procuradores nos Estados Unidos, onde foi concluído que “a empresa participou na coleta não autorizada de dados de redes sem fio, incluindo dados de redes Wi-Fi privadas de usuários de internet residencial” (EPIC apud ZUBOFF, 2018, p. 29), onde dados foram extraídos intencionalmente e sem consentimento para fins comerciais.

Por meio da experiência obtida nos processos contra o *Street View*, a Google desenvolveu um “método declarativo” frequentemente utilizado em seus empreendimentos envolvendo dados. “O *modus operandi* consiste em fazer incursões em territórios privados não protegidos até que alguma resistência seja encontrada”, colocando seus projetos à frente de tudo e resistindo a pedir autorização. Ao ser notificada judicialmente, a empresa “esgota seus adversários no tribunal ou eventualmente concorda em pagar multas que representam um investimento negligenciável para um retorno significativo” (ZUBOFF, 2018, p. 30).

Grande parte das práticas utilizadas pela Google só se tornaram conhecidas pelo público devido aos conflitos produzidos por elas. A fins de exemplo, Zuboff cita alguns desses conflitos:

a Google enfrentou oposição legal e protesto social em relação a reclamações contra (1) a prática de varredura de e-mails, incluindo os de usuários que não são do Gmail e os de estudantes que usam seus aplicativos educacionais, (2) a captura de comunicações de voz, (3) ignorar as configurações de privacidade, (4) práticas unilaterais de agrupamento de dados em seus serviços online, (5) extensa retenção de dados de pesquisa, (6) rastreamento dos dados de localização dos smartphones de reconhecimento facial e (7) suas tecnologias portáteis e capacidades de reconhecimento facial. (ZUBOFF, 2018, p. 36)

A ausência de reciprocidades estruturais entre empresas como a Google com suas populações nos processos de extração reflete a separação dessa lógica de acumulação da “narrativa histórica das democracias de mercado ocidentais” (ZUBOFF, 2018, p. 36). Nessa narrativa, as empresas e suas populações sustentavam uma relação de interdependência, uma vez que a existência de populações consumidoras era importante para a sua sustentação. Com o *big data* e o novo modelo de negócios dessa lógica de acumulação, “suas populações não são mais necessárias como fonte de clientes ou funcionários” (ZUBOFF, 2018, p. 37), uma vez que elas existem apenas como fontes de dados a serem utilizados, mas não é diretamente com seus usuários que a empresa faz negócios. Seus verdadeiros clientes são os anunciantes e outros intermediários envolvidos na compra de suas análises de dados. “Os dados viajam através de muitas fases de produção, apenas para retornar à sua fonte em uma segunda fase de extração em que o objetivo não é mais lidar com os dados, mas produzir receita” (ZUBOFF, 2018, p. 40).

Nesse novo modelo de acumulação do capitalismo de vigilância, os direitos à privacidade não são corroídos, mas sim redistribuídos. Dessa maneira, o direito à privacidade

não é mais algo possuído por um grande número de pessoas, mas se concentra naqueles que estão no interior do “regime de vigilância” (ZUBOFF, 2018, p. 47). Os capitalistas de vigilância, ou seja, aqueles que concentram o poder e o lucro envolvendo o mundo dos dados (Google e Meta, por exemplo), são quem possuem “amplos direitos de privacidade e, portanto, muitas oportunidades para segredos” (ZUBOFF, 2018, p. 47). Estes segredos são usados cada vez mais com o objetivo de privar suas populações do seu direito de escolha envolvendo partes da sua vida onde o indivíduo poderia desejar manter em sigilo.

Os capitalistas de vigilância exploraram de forma hábil um lapso na evolução social, uma vez que o rápido desenvolvimento de suas habilidades de vigiar para o lucro em muito suplantou a compreensão pública e o eventual desenvolvimento de leis e regulamentações legais. Como resultado, os direitos à privacidade, uma vez acumulados e afirmados, podem então ser invocados como legitimação para manter a obscuridade das operações de vigilância. (ZUBOFF, 2018, p. 48)

Zuboff, em sua conclusão, ressalta que é notório que os processos legais direcionados aos atores do capitalismo de vigilância não têm se mostrado suficientes para conter e alterar seu comportamento. Enquanto lógica de acumulação, o interesse por um “rápido aumento de uma riqueza sem tamanho e por concentrações históricas de poder” (ZUBOFF, 2018, p. 60) sobressai qualquer preocupação sobre as partes prejudicadas. “O valor do fluxo constante de ações legais” consiste em, primeiramente, estabelecer novos precedentes que possam ser utilizados como parâmetro para futuros casos, e em última instância, criar novas leis.

Por fim, e o mais importante para todos os estudiosos e cidadãos, é o fato de que estamos no início da narrativa que nos levará a novas respostas. A trajetória dessa narrativa depende em grande parte dos estudiosos atraídos para esse projeto pioneiro e dos cidadãos que agem sabendo que a ignorância induzida por engano não é um contrato social e que a liberdade da incerteza não é liberdade. (ZUBOFF, 2018, p. 60)

CONSIDERAÇÕES FINAIS

O presente trabalho teve como objetivo central a construção de uma análise ética em torno dos usos e consequências do big data na sociedade da informação. Para atingir essa meta, inicialmente, foi apresentada uma ideia generalista de technoética encabeçada por Luppicini (2009). Em seguida, a discussão foi aprofundada ao explorar a ética computacional de Moor (1985), considerado um dos precursores desse campo, e as contribuições de Floridi e Taddeo (2016), que propuseram uma ética mais voltada para o *software* do que para o *hardware*. Aqui, são apresentados os níveis de abstração, fundamentais para compreender as diferenças entre uma ética centrada no computador, na informação e nos dados. A partir daí, é possível compreender as principais questões concernentes a ética de dados, assunto central deste trabalho, ao abordar alguns de seus desafios em uma análise conjunta dos três eixos de pesquisa: dados, algoritmos e práticas.

Ao apresentar o desafio da discriminação e responsabilização da IA e algoritmos, foi possível notar as dificuldades em atribuir responsabilidade ao lidar com um sistema que utiliza de um aprendizado de máquina. O exemplo utilizado foi o robô Tay, cujo comportamento era moldado pelos usuários do Twitter, e que acabou desenvolvendo atitudes extremamente ofensivas direcionado a minorias. Aqui, torna-se complicado definir quem são os agentes das ações discriminatórias de sistemas, assim como também não é simples atribuir intencionalidade à ações de uma máquina cuja as escolhas não são muito claras, uma vez que o autodesenvolvimento do algoritmo pelo aprendizado de máquina faz o seu sistema autônomo ser considerado uma caixa preta. Conforme os avanços desse tipo de tecnologia aumentam, maiores se tornam as brechas de responsabilidade causadas pelo distanciamento da intervenção humana.

A propagação de ideias equivocadas sobre tecnologia, influenciada fortemente pela mídia, não facilita a responsabilização desses casos. Uma dessas concepções errôneas é a suposição de neutralidade dos dispositivos tecnológicos, juntamente com a prevalência de uma abordagem predominantemente reducionista na atribuição de responsabilidade. Ao invés de buscar uma compreensão mais abrangente das obrigações e deveres, relacionados à uma perspectiva prospectiva, muitas vezes há uma tendência na mídia de procurar os culpados pelas ações, agindo através de uma perspectiva retrospectiva e negligenciando a oportunidade de estabelecer medidas preventivas para evitar a recorrência de determinadas situações.

Em razão da incapacidade algorítmica de fazer julgamentos morais, as redes sociais recorrem a moderadores de conteúdo que possuem o objetivo de manter as postagens de acordos com as regras da comunidade. Contudo, esses trabalhadores precisam se submeter a condições que podem ser colocadas como mentalmente insalubres para, no fim, viverem com uma remuneração baixa e direitos trabalhistas quase inexistentes. Além disso, as empresas estabelecem regras que priorizam muito mais os interesses de seus investidores do que a justiça e os direitos humanos. Por tentar manter o trabalho de moderação sempre invisível e encoberto, as redes acabam por naturalizar atitudes ofensivas e violentas.

Outro desafio da ética de dados apresentado neste trabalho é o da privacidade de dados. O funcionamento dos sistemas de aprendizado de máquina dependem de bancos de dados capazes de armazenar uma quantidade massiva. Esses dados são denominados *big data*, e são conhecidos por serem coletados em grandes volumes, com ampla variedade e alta velocidade de acesso. Aqui, com o auxílio dos trabalhos de Sant’Ana (2016, 2019, 2020) sobre o assunto, é possível evidenciar de forma mais técnica o distanciamento do trabalho humano criado pelo crescente processo de algoritmização das tecnologias. Nesse processo, nem os intermediários responsáveis por desenvolverem o código do algoritmo conseguem prever seus resultados.

Com o avanço das TICs, torna-se ainda mais valioso o poder de coleta, armazenamento, retenção e troca de dados. No entanto, esse processo tem trazido preocupações concernentes ao potencial de tornar uma sociedade da informação em uma sociedade de vigilância. Aqui, a necessidade da elaboração de leis e direitos que lidem com os vácuos de políticas colocados por Moor (1985) diante das novas possibilidades se fazem ainda mais urgentes. O texto explora o valor do direito à privacidade, utilizando a abordagem de Floridi (2014) para contrastar duas interpretações — a reducionista e a baseada na propriedade. Concluindo a inadequação das duas em abarcar as questões relacionadas à sociedade da informação, é introduzida uma nova interpretação, proposta por Floridi: a interpretação autoconstitutiva. Essa abordagem considera os indivíduos enquanto inforgs, ou seja, organismos informacionais, que compõem a sociedade informacional. Nessa visão, o direito à privacidade está relacionado com o direito à preservação do indivíduo e sua própria identidade, enquanto entidade composta de informação.

Ao tratar dos usos da extração de dados, é essencial compreender o potencial lucrativo que eles possuem. Ao final deste trabalho, procura-se incitar uma reflexão, com o auxílio de

Zuboff (2018), acerca das novas lógicas de acumulação que permeiam o universo dos dados no emergente sistema, denominado capitalismo de vigilância. Tais dinâmicas tornam-se evidentes ao examinar as práticas de grandes empresas do Vale do Silício, em especial a Google, que se destacou como pioneira na integração do *big data* como uma ferramenta fundamental para seus negócios.

Dessa forma, o trabalho cumpre sua proposta ao realizar análises éticas e provocar reflexões sobre os desafios relacionados à coleta massiva de dados na contemporaneidade. A complexidade ética que envolve o *big data* demanda uma abordagem reflexiva contínua à medida que avançamos na era da informação.

REFERÊNCIAS BIBLIOGRÁFICAS

ANGWIN, Julia, GRASSEGGER, H. **Facebook's Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children**. ProPublica, 28 jun 2017, disponível em: <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms>. Acesso em: 8 jan. 2024.

BRASIL. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Brasília, DF: Presidência da República, 2020.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, 2020.

CARDON, D.; COINTET, J.; MAZIÈRES, A. Neurons spike back: The invention of inductive machines and the artificial intelligence controversy. **Reaux**, v. 36, n. 211, p. 173-220, 2018. Disponível em: <https://mazieres.gitlab.io/neurons-spike-back/NeuronsSpikeBack.pdf>. Acesso em: 8 jan. 2024.

COPELAND, J. B. **The Modern History of Computing**. The Stanford Encyclopedia of Philosophy (Winter 2020 Edition), Edward N. Zalta (ed.). Disponível em: <https://plato.stanford.edu/archives/win2020/entries/computing-history/>. Acesso em 19 de junho de 2023.

CORMEN, T. H. et al. **Algoritmos**. Tradução de Arlete Simille Marques. 3ª edição. Rio de Janeiro: Elsevier, 2012.

FACELI, K. et al. **Inteligência artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC., 2011.

FLORIDI, L; SANDERS, J. W. On the morality of artificial agents. *Minds and Machines*, Oxford, 14(3):349-379, 2004. Disponível em: https://www.researchgate.net/publication/227190404_On_the_Morality_of_Artificial_Agents.

Acesso em: 8 jan. 2024.

FLORIDI, L.; TADDEO, Mariarosaria. What is data ethics? **Philosophical Transactions of the Royal Society A**, Londres, v. 374, n. 2083, dec. 2016. Disponível em: <https://royalsocietypublishing.org/toc/rsta/2016/374/2083>. Acesso em: 8 jan. 2024.

FLORIDI, L. Faultless responsibility: on the nature and allocation of moral responsibility for distributed moral actions. **Philosophical Transactions of the Royal Society A**, Londres, v. 374, n. 2083, dec. 2016. Disponível em: <https://royalsocietypublishing.org/toc/rsta/2016/374/2083>. Acesso em: 8 jan. 2024.

FLORIDI, L.; TURILLI, M. The ethics of information transparency. **Ethics and Information Technology**, 11 (2):105-112, 2009. Disponível em: <https://philarchive.org/rec/TURTEO-4>. Acesso em: 8 jan. 2024.

FLORIDI, L. **The Blackwell guide to the philosophy of computing and information**. Malden, MA: Blackwell Publ., 2004.

FLORIDI, L. **The fourth revolution: how the infosphere is reshaping human reality**. Oxford: Oxford University Press, 2016.

FLORIDI, L. **The philosophy of information**. Oxford: Oxford University Press, 2014.

FLORIDI, L. **The ethics of information**. Oxford, UK: Oxford University Press, 2015.

FLORIDI, L. Artificial Intelligence's New Frontier: Artificial Companions and the Fourth Revolution. **Metaphilosophy**, v. 39, n. 4/5, p. 651-655, oct. 2008. Disponível em: https://www.researchgate.net/publication/227739537_Artificial_Intelligence's_New_Frontier_Artificial_Companions_and_the_Fourth_Revolution#fullTextFileContent. Acesso em: 8 jan. 2024.

HIRATA, Nina S. T. Introdução a Machine Learning a partir de uma perspectiva computacional. In: SOUZA, Edna A. de; BROENS, Mariana C.; GONZALEZ, Maria Eunice Q. **Big Data: Implicações Epistemológicas e Éticas**. Campinas: FiloCzar, 2020. *E-book*. Disponível em: <https://www.cle.unicamp.br/ebooks/index.php/publicacoes/catalog/download/8/7/29?inline=1>. Acesso em: 08/02/2024.

JOHNSON, Deborah. Computer Systems: Moral Entities but not Moral Agents. **Ethics and Information Technology**, v. 8, p.195–204, 2006. Disponível em: <https://doi.org/10.1007/s10676-006-9111-5>. Acesso em: 8 jan. 2024.

LECLERC, A. **Uma introdução à filosofia da mente**. Paraná: Appris Editora, 2018.

LEONELLI, Sabina. Locating ethics in data science: responsibility and accountability in global and distributed knowledge production systems. **Philosophical Transactions of the Royal Society A**, Londres, v. 374, n. 2083, dec. 2016. Disponível em: <https://royalsocietypublishing.org/toc/rsta/2016/374/2083>. Acesso em: 8 jan. 2024.

LUPPICINI, R. The Emerging Field of Technoethics. *In*: LUPPICINI, R; ADELL, R. **Handbook of Research on Technoethics**. New York: Information Science Reference, 2008. p. 1-19.

MOOR, James H. “WHAT IS COMPUTER ETHICS?”. **Metaphilosophy**, vol. 16, no. 4, 1985, pp. 266–75. Disponível em: <http://www.jstor.org/stable/24436819>. Acessado em 24 jan. 2024.

KEMPERAAND, J.; KOLKMAN, D. Transparent to whom? No algorithmic accountability without a critical audience. **Information, Communication & Society**, Amsterdam, Vol. 22, No. 14, 2019. Disponível em: <https://doi.org/10.1080/1369118X.2018.1477967>. Acesso em: 8 jan. 2024.

ROBERTS, Sarah T. Commercial Content Moderation: Digital Laborers' Dirty Work. *In*: NOBLE, S. U.; TYNES, B. **The Intersectional Internet: Race, Sex, Class and Culture Online**. Londres: Media Studies Publications, 2016. Disponível em: <https://ir.lib.uwo.ca/commpub/12>. Acesso em: 8 jan. 2024.

SANT'ANA, R.C.G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, [S.l.], v. 21, n. 2, p. 116–142, dez. 2016. ISSN 1981-8920. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/27940>. Acesso em: 18 jan 2024.

SANT'ANA, R.C.G. **Campo Informacional Resultante da Interação de Ciclos de Vida dos Dados**. *In*: DIAS, G.; FREIRE, B. (Orgs). **Dados Científicos: perspectivas e desafios**. João Pessoa: Editora UFPB, 2019.

SANT'ANA, R.C.G. Transdução informacional: impactos do controle sobre os dados. *In*: MARTÍNEZ-ÁVILA, D., SOUZA, E.A., GONZALEZ, M.E.Q. **Informação, conhecimento, ação autônoma e big data: continuidade ou revolução?** [online]. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica; FiloCzar, 2019. Disponível em: <http://books.scielo.org/id/gfrbh/pdf/martinez-9788572490559-09.pdf>. Acesso em: 8 jan. 2024.

SANT'ANA, R.C.G. A algoritmização a partir dos atores do acesso a dados. *In*: SOUZA, E. A. de; BROENS, M. C.; GONZALEZ, M.E.Q. **Big Data: Implicações Epistemológicas e Éticas**. Campinas: FiloCzar, 2020. *E-book*. Disponível em: <https://www.cle.unicamp.br/ebooks/index.php/publicacoes/catalog/download/8/7/29?inline=1>. Acesso em: 08/02/2024

SCHAFER, B. Compelling truth: legal protection of the infosphere against big data spills. **Philosophical Transactions of the Royal Society A**, Londres, v. 374, n. 2083, dec. 2016. Disponível em: <https://royalsocietypublishing.org/toc/rsta/2016/374/2083>. Acesso em: 8 jan. 2024.

SILVA, T. **Racismo algorítmico: inteligência artificial e discriminação nas redes digitais**. São Paulo: SESC Edições, 2022. *E-book*. Disponível em: https://www.researchgate.net/publication/358278108_Racismo_algoritmico_inteligencia_artificial_e_discriminacao_nas_redes_digitais. Acesso em: 8 jan. 2024.

SOUZA, Edna A. de; BROENS, Mariana C.; GONZALEZ, Maria Eunice Q. **Big Data: Implicações Epistemológicas e Éticas**. Campinas: FiloCzar, 2020. *E-book*. Disponível em: <https://www.cle.unicamp.br/ebooks/index.php/publicacoes/catalog/download/8/7/29?inline=1>. Acesso em: 08/02/2024

SOUZA, Edna A. de; GONZALEZ, J. A. Q., GONZALEZ; Maria Eunice Q. Big Data in Contemporary Science: Methodological and Ethical Implications for Everyday Life. **The GCAS Review Journal**, v. I, 2021. Disponível em: <https://www.gcasreview.com/publications-technology-and-society>. Acesso em: 8 jan. 2024.

SUÁREZ-GONZALO, Sara; MAS-MANCHÓN, L; GUERRERO-SOLÉ, F. Tay is you. The attribution of responsibility in the algorithmic culture. **Observatorio Journal**, Espanha, 13(2), 1-14, 2019. Disponível em: <https://obs.obercom.pt/index.php/obs/article/view/1432/pdf>. Acesso em: 8 jan. 2024.

ZUBOFF, Shoshana. Big Other: capitalismo de vigilância e perspectivas para uma civilização de informação. *In*: BRUNO, F. et al. **Tecnopolíticas da vigilância: perspectivas da margem**. 1ª edição. São Paulo: Boitempo, 2018. p. 17-68.