INSTITUTO BRASILIENSE DE DIREITO PUBLICO – IDP
INSTITUTO DE ENSINO SUPERIOR – ICEV
MESTRADO EM DIREITO CONSTITUCIONAL

Nazareno César Moreira Reis

DIREITO À PROTEÇÃO DE DADOS E DECISÕES AUTOMATIZADAS: os direitos do titular à luz da LGPD.

Teresina
2021

Nazareno César Moreira Reis

DIREITO À PROTEÇÃO DE DADOS E DECISÕES AUTOMATIZADAS: os direitos do titular à luz da LGPD.

Dissertação apresentada ao Programa de Pós-Graduação em Direito Constitucional do IDP/iCEV, como requisito para obtenção do título de Mestre em Direito.

Orientadora: Profª. Drª. Laura Schertel Mendes

Teresina
2021

Nazareno César Moreira Reis

DIREITO À PROTEÇÃO DE DADOS E DECISÕES AUTOMATIZADAS: os direitos do titular à luz da LGPD.

Dissertação apresentada ao Programa de Pós-Graduação em Direito Constitucional do IDP/iCEV, como requisito para obtenção do título de Mestre em Direito.

Aprovado em:____/_____/_____

BANCA EXAMINADORA

_____

Profª. Drª. Laura Schertel Mendes- IDP (Orientadora)

_____

Prof. Dr. Danilo Doneda - IDP (Avaliador)

_____

Prof. Dr.Gabriel Rocha Furtado- iCEV/UFPI (Avaliador)

À minha esposa, Cláudia, e aos meus filhos, Mário e Rafael.

AGRADECIMENTOS

A minha orientadora, Professora Laura Schertel Mendes, pelas seguras orientações. Ao meu amigo Bruno Oliveira, pelas discussões fecundas. À Ana Luísa Melo, pelo inestimável auxílio. A todos que, de alguma forma, contribuíram para este trabalho.

RESUMO

A presente dissertação tratou sobre a temática da relação entre o direito e as decisões automatizadas produzidas por máquinas eletrônicas à luz do marco legal sobre o tema no Brasil: A Lei Geral de Proteção de Dados. A pesquisa buscou investigar que direitos e garantias individuais têm as pessoas contra quem são tomadas decisões automatizadas com base no tratamento de seus dados pessoais. Para desenvolver a pesquisa, além da análise da Lei Geral de Proteção de Dados, buscou-se investigar também a proteção de dados como um direito fundamental de base constitucional. A metodologia usada foi a pesquisa bibliográfica de doutrinas jurídicas e a pesquisa documental de legislação e jurisprudência. Diante da complexidade e interdisciplinaridade do tema com outras áreas, mostra-se necessária a análise de textos fora do Direito, tais como de Ciência da Computação e de Filosofia da Tecnologia, em determinados momentos da pesquisa. Para concluir o trabalho, enumerou-se e definiu-se quais são, à luz da LGPD, os direitos do titular em relação a quem foi ou será tomada uma decisão automatizada.

Palavras-chave: proteção de dados, dados pessoais, decisões automatizadas, inteligência atificial, Lei Geral de Proteção de Dados.

ABSTRACT

The present dissertation discussed the theme of the relationship between law and automated decisions produced by electronic machines based on the legal framework on the subject in Brazil: The General Data Protection Law. The research investigated what individual rights and guarantees have the people against whom automated decisions are made based on the treatment of their personal data. In order to carry out the research, in addition to the analysis of the General Data Protection Law, we also sought to investigate data protection as a fundamental constitutionally based right. The methodology used was the bibliographic research of legal doctrines and the documentary research of legislation and jurisprudence. In view of the complexity and interdisciplinarity of the theme with other areas, it is necessary to analyze texts outside the law, such as Computer Science and Philosophy of Technology, at certain times of the research. To conclude the work, it was enumerated and defined what, based on the LGPD, the rights of the holder are in relation to whom an automated decision was or will be taken.

Keywords: data protection, personal data, automated decisions, Artificial Intelligence, General Data Protection Law.

LISTA DE ABREVIATURAS E SIGLAS

AGI – sigla em inglês para Inteligência Artificial Geral

ANN - Artificial Neural Networks

ANPD - Agência Nacional de Proteção de Dados

AI – sigla em inglês para Inteligência Artificial

CC – Código Civil

CDC – Código de Defesa do Consumidor

CEO – sigla em inglês para Diretor Executivo

CF – Constituição Federal de 1988

COMPAS - Correctional Offender Management Profiling for Alternative Sanctions

CR – Constituição da República de 1891

DL - Deep Learning

EUA – Estados Unidos da América

FISA - Foreign Intelligence Surveillance Act

GDPR - General Data Protection Regulation

LCP – Lei de Cadastro Positivo

LGPD – Lei Geral de Proteção de Dados

MCI - Marco Civil da Internet

ML – Machine Learning

OCDE – Organização para a Cooperação e Desenvolvimento Econômico

PbD - Privacy by design

STF – Supremo Tribunal Federal

STJ – Superior Tribunal de Justiça

UE – União Europeia

VSD - Value Sensitive Design

XAI - sigla em inglês para Inteligência Artificial Explicável

SUMÁRIO

INTRODUÇÃO **10**
1 TECNOLOGIAS DA COMUNICAÇÃO E ADAPTAÇÕES JURÍDICAS **15**
1.1 Tecnologia e transformações sociais **15**
1.2 Oralidade, escrita e informática: da memória ao hipertexto **16**
1.2.1 O jornalismo: um fruto imprevisto da prensa móvel *19*
1.2.2 Eletricidade e telecomunicações: tecnologia em rede e internacionalização *20*
1.2.3 Telefone, rádio e televisão *22*
1.3 Eletricidade e lógica: computadores e internet **25**
1.4 A indústria dos dados **26**
1.4.1 Dados digitais e seu tratamento *29*
1.4.2 Inteligência Artificial. *31*
1.4.2.1 Aprendizado de Máquina (Machine Learning – ML) 32
1.4.2.2 Aprendizado Supervisionado 34
1.5 O ecossistema digital **36**
2 DO DIREITO À PROTEÇÃO DE DADOS: EVOLUÇÃO E CARACTERÍSTICAS **38**
2.1 A evolução do direito à privacidade **38**
2.2 O direito à proteção de dados pessoais **45**
2.2.1 A proteção de dados no ecossistema da internet *49*
2.2.1.1 O dilema do indivíduo nas redes 49
2.2.1.2 A dimensão política da proteção de dados pessoais 52
2.3 Natureza complexa do direito à proteção de dados: direito da personalidade, direito fundamental e direito básico do consumidor **59**
2.3.1 Reconhecimento do direito fundamental à proteção de dados pelo STF *61*
2.3.2 Proteção de dados e direitos do consumidor *63*
2.4 A configuração técnica da internet e o direito à proteção de dados **68**
2.5 Internet de Quinta Geração (5G), Economia da Abundância, Internet das Coisas e Privacidade **70**
2.6 O consentimento do titular dos dados **72**
2.6.1 Problemas cognitivos *74*
2.6.2 Problemas estruturais *78*
3 DECISÕES AUTOMATIZADAS: DEFINIÇÃO, BENEFÍCIOS E RISCOS **84**
3.1 Conceito **84**
3.1.1 Uso de dados pessoais *87*

3.1.2 Tratamento automatizado *89*

4.1.2.1 Tratamentos automatizados excluídos do alcance da LGPD (tratamentos domésticos, jornalísticos, artísticos e acadêmicos) 92

4.1.2.2 Tratamento automatizado regulado subsidiariamente pela LGPD (segurança pública, defesa nacional, segurança do Estado e atividades de investigação e repressão de infrações penais) 92

3.1.3 Ameaça ou lesão a interesse juridicamente tutelado *95*

3.1.4 Definição *96*

3.2 Perfilização 97

3.3 Benefícios 101

3.3.1 Ciclo Virtuoso da Inteligência Artificial 103

3.4 Riscos 110

4 DECISÕES AUTOMATIZADAS: OS DIREITOS DO TITULAR 115

4.1 Introdução 115

4.1.1 Direitos dos titulares na LGPD 116

4.1.2 Outros direitos 117

4.2 Direitos Relativos à Formação da Decisão Automatizada 119

4.2.1 Direito ao design adequado do sistema (art. 49), inclusive com preferência para padrões técnicos que possibilitem ao próprio titular a fiscalização (art. 51) 119

4.2.2 Direito à confirmação da existência do tratamento (LGPD, art. 18,I c/c art. 19) 123

4.2.3 Direito de opor-se ao tratamento (art. 18, VIII e §2º, LGPD) 124

4.2.4 Direito de acesso aos dados pessoais e aos compartilhamentos (art. 6º, IV e VI c/c art. 9º e art. 18, II e VII, LGPD) 126

4.2.5 Direito de correção (art. 18, III, LGPD) 129

4.2.6 Direito à anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou tratados em desconformidade (art. 18, IV, LGPD) 131

4.2.7 Direito de consentir ou não no tratamento (arts. 7°, 11 e 14, LGPD) e Direito de revogar o consentimento (art. 8º, §5º, art. 15, III c/c art. 18, XI, LGPD) 133

4.3 Direitos Relativos aos Resultados da Decisão Automatizada 137

4.3.1 Direito à explicação 137

4.3.2 Direito de Revisão das Decisões Automatizadas (art. 20, LGPD e art. 5º, VI da LCP) 140

4.4 Direitos Instrumentais 142

4.4.1 Direito de petição (CF, art. 5º, XXXIV, "a"; LGPD, art. 18,§1° c/c art. 55-J, V) 142

4.4.2 Direito ao devido processo legal (CF, art. 5º, LIV; LGPD, art. 4º, §1º) 143

CONCLUSÃO 145

REFERÊNCIAS 147

INTRODUÇÃO

O tema desta pesquisa diz respeito à articulação entre o direito e as decisões automatizadas produzidas por máquinas eletrônicas. Havendo já um marco legal no país sobre o assunto (Lei 13.709, de 14 de agosto de 2018, conhecida como Lei Geral de Proteção de Dados - LGPD), a investigação volta-se para tentar esclarecer que direitos são pertinentes em casos de decisões automatizadas.

A relevância do tema é grande. O modo de vida atual torna-se cada vez mais dependente das tecnologias da comunicação, em especial dos dispositivos computacionais (fixos ou móveis) e da internet. O trabalho, as compras, o lazer e, enfim, as relações humanas em geral passam por um processo de incorporação à atmosfera digital.

A representação de crescente coleção de fatos da vida em termos de códigos e objetos digitais, por seu turno, tem reduzido a distância entre o plano *off-line* o plano *on-line*, gerando influências recíprocas entre eles e abalando modelos mentais outrora consolidados, em especial aqueles ligados à intimidade e à privacidade.

Como todas as operações com códigos digitais são potencialmente gravadas e tendem a convergir para a internet, no ambiente das redes não há nada informal, nada local e nada completamente oculto à esfera pública. Todas as categorias jurídicas que se utilizam, portanto, das ideias de informalidade, territorialidade ou de sigilo precisam ser reconsideradas à luz dessa nova realidade material e social.[1]

Mesmo as mais prosaicas ações humanas — como aquelas praticadas no recinto do lar, por exemplo —, assumem peso muito maior quando são executadas em meio digital. E isso decorre da circunstância de que qualquer manifestação no mundo virtual transforma-se em documento, que tende a inserir-se na rede mundial de computadores (a internet), com as inúmeras consequências que disso podem resultar.

A massa sempre crescente de dados, produzidos simultaneamente em vários contextos da vida coletiva, desde o âmbito doméstico até à política e à cultura, passando pela agricultura, indústria, comércio, ensino, pesquisa, etc., não fica sob o controle absoluto de nenhum indivíduo ou governo. Esses dados aglutinam-se em grupos geralmente volumosos (os *big datas*), fundem-se, refundem-se, apartam-se e se propagam inexoravelmente nos meios digitais; e, mesmo quando protegidos por mecanismos de criptografia que imitam no mundo *on-line* os

---

[1] LÉVY, Pierre. **As tecnologias da inteligência**: o futuro do pensamento na era da informática. São Paulo: Editora 34, 1993. Tradução de Carlos Irineu da Costa, p. 115-134.

muros, as paredes e os cofres do mundo *off-line*, apresentam suscetibilidades próprias de sua conformação, que precisam ser consideradas pelo direito.[2]

Embora os dados sejam gravados, no modo digital, sob a mesma lógica e segundo um padrão físico homogêneo (como um sinal eletromagnético), os fatos aos quais eles se referem concedem-lhes pesos jurídicos muito variados. Enquanto o meio físico os iguala, os valores humanos subjacentes os hierarquizam, donde surge uma tensão entre a técnica e a política, que acaba se expressando em termos jurídicos.

Assim é que, quando os dados estão ligados a uma pessoa natural identificada ou identificável, isto é, quando dizem respeito a fatos ou atos da vida de um ser humano, indicando aspectos específicos dos seus comportamentos, dos seus gostos, das suas preferências, eles são chamados de "dados pessoais" (LGPD, art. 5º, I), e têm uma proteção legal especial. Se, ademais, forem considerados especificamente os dados da pessoa natural que estão ligados à sua origem racial ou étnica, à sua convicção religiosa, à sua opinião política, à sua filiação a sindicato ou a organização de caráter religioso, filosófico ou político, bem como os referentes à sua saúde ou à sua vida sexual, dados genéticos ou biométricos — então se fala em "dados sensíveis" (LGPD, art. 5º, II), cuja proteção legal é ainda mais forte.

Esses dados (os pessoais e, mais ainda, os sensíveis) apresentam valor maior para o direito porque são expressões da personalidade humana, estando por isso no centro do ordenamento jurídico (CF, art. 1º, III). O esquema doutrinário tradicional, que explica a relação do homem com as coisas por meio dos direitos reais, em especial o direito de propriedade, não atende às necessidades ligadas à proteção dos dados pessoais. É que, por meio desses rastros digitais, com o devido "tratamento", pode-se reconstituir fatos, atos e até pensamentos relacionados a alguém, acossando o ser humano na intimidade de sua vida intelectual, afetiva, moral, política, econômica e social. Os dados pessoais não são, portanto, em relação à pessoa a quem se referem, coisas sobre as quais ela exerce algum direito real, mas sim emanações da sua personalidade, daí porque se fala de um "direito à proteção de dados", e não de um direito de propriedade sobre dados[3].

---

[2] HILDEBRANDT, Mireille. Privacy as Protection of the Incomputable Self: from agnostic to agonistic machine learning. **Theoretical Inquiries In Law**, Tel Aviv, v. 20, n. 1, p. 83-121, jan. 2019. Disponível em: https://www7.tau.ac.il/ojs/index.php/til/article/view/1622/1723. Acesso em: 17 jul. 2020; CRAWFORD, Kate; SCHULTZ, Jason. Big Data and Due Process: toward a framework to redress predictive privacy harms. **Boston College Law Review**, Boston, v. 55, n. 1, p. 93-128, 29 jan. 2014. Disponível em: https://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4. Acesso em: 17 jul. 2020.

[3] MENDES, Laura Schertel. **Privacidade, proteção de dados e defesa do consumidor**: linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014. p.120-124.

Com efeito, métodos sofisticados de tratamento de dados, chamados genericamente de Inteligência Artificial, permitem a recopilação de dados dispersos para reconstituir ações humanas e para analisar, prever ou mesmo induzir comportamentos futuros. Enfim, permitem formar uma imagem completa do indivíduo, a partir dos vestígios digitais dos seus movimentos nas redes, prognosticando as suas ações, características, interesses e até pensamentos.

Os usos desse poder novo e espantoso têm sido muito diversificados. As máquinas têm sido usadas para avaliar a capacidade de pagamento de pessoas que pedem empréstimos[4], para estimar preços de mercadorias conforme o consumidor que as queira comprar[5], para prever locais que devem receber maior atenção das rondas policiais[6], para dirigir carros autônomos[7], para fazer diagnósticos de doenças[8], em reconhecimento facial ou detecção de objetos por imagens para diversos fins, em *sites* de busca, veículos autônomos, classificação de crédito, publicidade comercial, seleção de pessoal para vagas de emprego, avaliação de produtos, etc.[9]

Na base de toda essa revolução está a Inteligência Artificial e os múltiplos usos que ela é capaz de fazer do grande volume de dados disponíveis na internet ou fora dela, sobretudo por meio do chamado Aprendizado de Máquina (*Machine Learning*). O conjunto dos efeitos sociais desses usos ainda é um território desconhecido, pleno de possibilidades, de esperanças e de muitos receios também.

O indiscutível é que essas novas tecnologias não podem ser ignoradas pelas ciências sociais. Para se ter uma ideia da estimativa de impacto da Inteligência Artificial entre os especialistas da área, Sundar Pichai, o atual CEO do Google, recentemente afirmou que ela terá mais consequências sobre a vida em sociedade do que teve o fogo ou a eletricidade[10].

---

[4] LEE, Tian-Shyug; CHEN, I-Fei. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. **Expert Systems with Applications**, [*s. l.*], v. 28, n. 4, p. 743-752, mai. 2005.

[5] HANNÁK, Anikó et al. Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr. **Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing**, New York, p. 1914–1933, fev. 2017.

[6] PERRY, Walter L. *et al*. **Predictive Policing**: The Role of Crime Forecasting in Law Enforcement Operations. [*S. l.*]: RAND Corporation, 2013. E-book.

[7] MASSACHUSETTS INSTITUTE OF TECHNOLOGY (ed.). **MIT Technology Review**. Self-driving cars. Topics. Disponível em: https://www.technologyreview.com/topic/smart-cities/self-driving-cars/. Acesso em: 02 jun. 2020.

[8] MIT TECHNOLOGY REVIEW INSIGHTS. How AI is humanizing health care: Artificial intelligence is helping health-care professionals do their jobs better, giving them the tools to build a smarter, more efficient ecosystem. *In*: MASSACHUSETTS INSTITUTE OF TECHNOLOGY (ed.). **MIT Technology Review**. [*S. l.*], 22 jan. 2020. Disponível em: https://www.technologyreview.com/2020/01/22/276128/how-ai-is-humanizing-health-care/. Acesso em: 2 jun. 2020.

[9] RAHWAN, -Iyad et al. Machine behaviour. **Nature**, [s. l.], v. 568, p. 477–486, 24 abr. 2019. Disponível em: https://www.nature.com/articles/s41586-019-1138-y. Acesso em: 2 jun. 2020.

[10] THOMSON, Amy; BODONI, Stephanie. Google CEO Thinks AI Will Be More Profound Change Than Fire. *In*: **Bloomberg**. [*S. l.*], 22 jan. 2020. Disponível em: https://www.bloomberg.com/news/articles/2020-01-22/google-ceo-thinks-ai-is-more-profound-than-fire. Acesso em: 2 jun. 2020.

Em semelhante contexto, o direito precisa se ocupar da disciplina dessas máquinas cognoscentes, visto que elas estão gerando fatos com importantes consequências jurídicas na vida das pessoas. Levanta-se, entre outras, uma questão que promete ocupar crescente atenção dos juristas em toda parte, mas nesta pesquisa com enfoque no Brasil: que direitos e garantias individuais têm as pessoas contra quem são tomadas decisões automatizadas com base no tratamento de seus dados pessoais?[11]

Esse é o problema de pesquisa que será desenvolvido, restringindo-se a investigação ao caso brasileiro, à luz da legislação interna, especialmente da Lei Geral de Proteção de Dados – LGPD (Lei 13.709, de 14 de agosto de 2018, com as alterações promovidas pela Lei n. 13.853, de 8 de julho de 2019) e da Constituição Federal de 1988.

O enfoque se dará nos direitos previstos na LGPD, notadamente no art. 18, mas sempre à luz da ideia mais ampla de proteção de dados como um direito fundamental, decorrente do direito à privacidade, bem como da própria dignidade humana.

O objetivo geral da pesquisa, desse modo, é avaliar o que são as decisões automatizadas, quais os riscos e benefícios que elas trazem, bem como quais as soluções jurídicas oferecidas pela LGPD para acomodar essa importante inovação tecnológica dentro do sistema jurídico. Admite-se como referência que existe um *direito fundamental à proteção dos dados pessoais*, como emanação do direito à privacidade, conforme tem apontado a mais moderna doutrina nacional[12], e que a efetivação desse direito fundamental depende não apenas de procedimentos e arranjos institucionais específicos, senão também de direitos instrumentais adequados.

As decisões automatizadas têm potencial para gerar riscos significativos ao *direito à proteção de dados pessoais*, em várias das suas dimensões. Tanto interesses individuais, como coletivos e difusos podem ser atingidos pelo uso algoritmos enviesados ou por coleta inidônea de dados.

As respostas jurídicas para situações do tipo dependem, antes de tudo, da compreensão do próprio fenômeno de concepção de decisões por máquinas. Por isso, será necessário para a pesquisa atingir seu objetivo geral, antes alcançar certos objetivos específicos, quais sejam: proceder a um escorço histórico da evolução das tecnologias da informação e a sua relação com transformações jurídicas; analisar o direito à proteção de dados, escrutinando os seus elementos

---

[11] CITRON, Danielle Keats. Technological Due Process. **Washington University Law Review**, Washington, D.c., v. 85, n. 6, p. 1249-1313, ago. 2008. Disponível em: https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview. Acesso em: 17 jul. 2020.

[12] MENDES, Laura Schertel. **Privacidade, proteção de dados e defesa do consumidor**: linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014.

principais; verificar o que são e como se formam as decisões automatizadas; pesquisar, no texto da LGPD, quais são os direitos do titular de dados que são especificamente aplicáveis em casos de decisões automatizadas.

A metodologia consiste basicamente em pesquisa bibliográfica de doutrinas jurídicas e a pesquisa documental de legislação e jurisprudência. Como o tema apresenta aspecto interdisciplinar, mostra-se necessária eventualmente também a análise de textos ligados a outras áreas, tais como de Ciência da Computação e de Filosofia da Tecnologia.

A partir de uma reflexão inicial sobre a influência das tecnologias da comunicação sobre o modo de vida das sociedades, a pesquisa busca desvendar o que se compreende por Algoritmo, Inteligência Artificial e Aprendizado de Máquina, de modo a apurar os fundamentos dos mecanismos subjacentes à tomada de decisões por máquinas, detendo-se sobre questões ligadas à transparência e aos possíveis vieses desses modelos de manipulação de dados.

A seguir, a pesquisa direciona-se para a avaliação do chamado "direito fundamental à proteção de dados" buscando extrair o seu fundamento e extensão. Considera-se a hipótese de que este direito tem raiz no direito à intimidade, previsto na Constituição Federal, e que estão compreendidos no seu conteúdo os "princípios" a que alude o art. 6º da LGPD, bem como certos direitos do titular.

Após, a pesquisa volta-se para a anatomia da decisão automatizada, intentando analisar os elementos fundamentais que a compõem, de modo a verificar como o iter decisório automático pode violar direitos individuais e quais são os mecanismos jurídicos de proteção do indivíduo nesse contexto.

Em conclusão, a pesquisa responde o problema de pesquisa, que consiste justamente em enumerar e definir quais são, à luz da LGPD, os direitos do titular em relação a quem foi ou será tomada uma decisão automatizada.

1 TECNOLOGIAS DA COMUNICAÇÃO E ADAPTAÇÕES JURÍDICAS

1.1 Tecnologia e transformações sociais

Está bem estabelecida relação entre a ocorrência de mudanças tecnológicas e a verificação de alterações jurídicas correspondentes, notadamente no conceito de privacidade ao longo da história[13]. Em particular, as mudanças ligadas à produção, armazenamento e fluxo das informações são manifestamente relevantes para o direito.

O direito nasce dos fatos históricos, não do acaso, nem porventura de uma sabedoria superior[14] que o crie *ex nihilo*. Contudo, o direito tem uma relação peculiar com os fatos sociais: por um lado é um reflexo da cultura e das formas de organização econômica e política no tempo e no espaço; por outro, é um projeto de atuação sobre essa mesma realidade, visando a dirigi-la.

Disso decorre que, quando sucedem mudanças rápidas e profundas numa dada sociedade — sejam elas políticas, econômicas, culturais ou tecnológicas —, isso naturalmente leva à obsolescência dos modelos jurídicos mais diretamente afetados por essas mudanças, donde emerge a necessidade de novas soluções jurídicas para assimilar as demandas e expectativas conaturais a essas transformações, potencializando-as ou refreando-as.

Foi assim, por exemplo, que uma revolução social e política, como foi a Revolução Francesa, no século XVIII, ao assolar o Antigo Regime, inspirou a criação do Direito Constitucional[15] e do Direito Administrativo[16], para submeter o poder político ao Direito — afinal o *leitmotiv* dessa Revolução. Foi desse modo também que a Revolução Industrial, um conjunto de mudanças mais da técnica de produção econômica, no século XIX, ao criar condições de trabalho degradantes, induziu a criação dos chamados "direitos sociais", para rearticular a relação da capital e do trabalho. Foi por causa da intensa degradação da Natureza, ligada a tantos acontecimentos novos do século XX (aumento da população, uso massivo de combustíveis fósseis e de substâncias tóxicas, surgimento de resíduos nucleares, etc), que surgiu o Direito Ambiental, especialmente depois da II Guerra[17].

---

[13] Cf.: DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos de formação da lei geral de proteção de dados. 2. ed. São Paulo: Thomson-reuters Brasil, 2019, p. 35-70.
[14] SAVIGNY, F. Carl Von. **Sistema del Derecho Romano Actua**. 2. ed. Madrid: Centro Editorial de Góngora, 2004. Tradução de: Espanhola de Jacinto Mesía e Manuel Poley. p. 69
[15] ACCIOLI, Wilson. **Instituições de Direito Constitucional.** 3ªed. Rio de Janeiro: Forense, 1984, p. 1-4; MIRANDA, Jorge. **Manual de Direito Constitucional**. 4. ed. Coimbra: Coimbra Editora Ltda, 1990. p. 14-15.
[16] TÁCIO, Caio. **Temas de Direito Público**: estudos e pareceres. Rio de Janeiro: Renovar, 1997, p. 1.
[17] NASH, Roderick Frazier. **The Rights of Nature**: a history of environmental ethics. Madison: The University Of Wisconsin Press, 1989, p. 10.

Os exemplos poderiam ser multiplicados. Interessa-nos neste trabalho, porém, concentrar a atenção sobre transformações sociais provocadas diretamente por progressos tecnológicos — notadamente na área da comunicação — e algumas das respostas oferecidas pelo direito para acomodá-las.

O tema é relevante, dado o surgimento recente de tecnologias ditas disruptivas e inevitáveis[18], que estão transformando rapidamente o modo como as pessoas e organizações se inter-relacionam em todos os aspectos da vida, notadamente nas funções de governança e controle.

Para tanto, convém fazer uma breve digressão histórica. A avaliação de experiências anteriores, ou seja, de situações nas quais a criação de tecnologias reordenou a política, a cultura e o direito parece ser o caminho mais seguro para a especulação a respeito das possíveis alternativas que o direito poderá e deverá criar para oferecer respostas satisfatórias aos desafios apresentados pelas nova tecnologias da informação.

1.2 Oralidade, escrita e informática: da memória ao hipertexto

Pierre Lèvy divide a História em três tempos, que ele chama de "três tempos do espírito": a oralidade primária, a escrita e a informática[19].

A oralidade primária refere-se ao período em que a palavra não tinha registro algum — diferentemente da oralidade secundária, que é aquela que subsiste paralelamente ao texto escrito, sendo por ele influenciada.

Numa sociedade de oralidade primária, a palavra tem a função de fixar o sentido e ao mesmo tempo o registro das ideias. A produção do espaço-tempo numa sociedade estruturada sobre a oralidade primária é essencialmente dependente da memória humana e do que nela pode ser gravado[20].

As estratégias mnemônicas, num contexto assim, mostram-se como de fundamental importância, e as mais exitosas tendem a ocupar um lugar de destaque no meio social. As narrativas dramáticas e mitológicas, por serem úteis para criar memória de longo prazo, dada a carga emocional que envolvem, assumem o centro da vida cultural e política[21].

---

[18] KELLY, Kevin. **Inevitável**: as 12 forças tecnológicas que mudarão nosso mundo. Rio de Janeiro: Alta Books, 2018. Tradução de Cristina Yamagami..

[19] LÉVY, Pierre. **As tecnologias da inteligência**: o futuro do pensamento na era da informática. São Paulo: Editora 34, 1993. Tradução de Carlos Irineu da Costa, p. 75.

[20] Ibidem, p. 78.

[21] Ibidem, p. 83.

O tempo da oralidade tende a ser circular, e não retilíneo, pois a falta de registro externo das memórias impede que as gerações seguintes assumam a sua tarefa como uma continuidade ininterrupta de uma vivência anterior. É o tempo da pré-história.

A invenção da escrita funda a História. E funda também a própria ideia de governo estatal, como explica Pierre Lèvy[22]:

> Através da escrita, o poder estatal comanda tanto os signos como os homens, fixando-os em uma função, designando-os para um território, ordenando-os sobre uma superfície unificada. Através dos anais, arquivos administrativos, leis, regulamentos e contas, o Estado tenta de todas as maneiras congelar, programar, represar, ou estocar seu futuro e seu passado. E é perseguindo o mesmo objetivo que manda construir monumentos, depósitos e muralhas nas cidades, e que mantém, a um alto custo, os silos, os canais de irrigação e as estradas.

A escrita cria uma nova relação com a mensagem. Separada do contexto de sua emissão original e, ao mesmo tempo, plasmada num suporte físico exterior à memória humana, a mensagem agora exige uma atualização permanente por um intermediário. Surge aqui um nicho de poder, que vai ser ocupado em grande parte pelos sacerdotes e pelos juristas, e uma fonte de novas vivências culturais: a tradição hermenêutica, que vai da religião ao direito, passando pelo comércio, a ciência e a guerra.

A invenção da prensa de tipos móveis, no século XVI, deu um novo impulso à escrita, ao conferir escala à sua produção, tornando acessíveis os textos a um número crescente de pessoas. Embora não se possa estabelecer uma relação de causa e efeito entre a prensa de tipos móveis e muitas transformações políticas e sociais que a sucederam imediatamente, é certo que ela ao menos moldou a forma como essas transformações ocorreram.

De fato, a conformação do Estado Moderno, por exemplo, deve-se, em grande parte, à prensa de tipos móveis. Embora a escrita em alfabeto fonético e o papel fossem já bastante antigos, foi só com a prensa de Gutenberg (século XVI) que se pôde imprimir e reproduzir com fidelidade e rapidez cópias de grande quantidade de livros e documentos, tais como leis, tábuas de arrecadação de impostas, anotações de pagamentos, nomeações, títulos, censos demográficos, etc.

O monopólio da força e da tributação, assegurado por uma burocracia especializada e diferenciada, não seria possível sem a prensa. Antigas organizações políticas — como o Império Romano, por exemplo — também chegaram a deter forte controle sobre áreas territoriais definidas e montaram incipientes corpos de funcionários, entretanto sem jamais terem alcançado o nível complexo de institucionalização burocrática que os estados modernos, com

[22] LÈVY, op. cit., p. 88.

seus aparelhos numerosos, ordenados, especializados, hierarquizados, impessoais e contínuos, conseguiriam — graças, em grande parte, à prensa de tipos móveis e aos amplos arquivos que elas geraram.

Como explica Norbert Elias[23], desde quando a burocracia estabeleceu-se de modo claro e incontestável, açambarcando praticamente todas as funções de governança, os conflitos sociais deixaram de versar sobre a eliminação do governo e passaram a concentrar-se exclusivamente sobre como e quem deveria ocupar o aparato do governo, ou seja, daquilo que se convencionou chamar de Estado.

Tal evolução seria impraticável sem o uso massivo de textos, amplamente acessíveis, com designação de órgãos, cargos, funções, competências e de procedimentos.

E aqui se pode licitamente extrair uma conclusão importante: as tecnologias, especialmente aquelas ligadas à produção e difusão da informação, mesmo quando não são causas diretas e imediatas de mudanças políticas, culturais e econômicas, acabam por formatar o modo como essas mudanças ocorrem, estabelecendo com elas naturalmente uma relação simbiótica.

Em paralelo à ação sobre a organização do Estado, a prensa de tipos móveis agiu de maneira transformadora também sobre a sociedade civil. Em razão da prensa e da tipografia, livros tornaram-se mercadoria acessível a número cada vez maior de pessoas. Uma explosão de descobertas científicas ocorreu quase simultaneamente em muitas áreas do conhecimento, graças ao acesso à informação exata. Pierre Lévy lembra, por exemplo, que grande parte das descobertas astronômicas da Renascença foram possíveis sem o uso de telescópio porque, "graças à impressão, Kepler e Tycho Brahe puderam servir-se de compêndios de observações antigas ou modernas que eram exatos e estavam disponíveis, assim como tabelas numéricas precisas."[24]

Mas não apenas a ciência foi transformada pelo acesso ampliado aos livros. A vida privada e familiar assumiu novos contornos. Roger Chartier[25] observa que, nas partes da Europa onde o acesso aos livros tornou-se significativamente maior [por causa da impressão em massa], o nível de alfabetização — e, por consequência da leitura — da população cresceu muito. A leitura silenciosa nos escritórios domésticos, proporcionada pelo livro impresso, cada vez mais

---

[23] ELIAS, Norbert. **O processo civilizador**: formação do estado e civilização. Rio de Janeiro: Jorge Zahar Editor, 1993. 2 v. Tradução de Ruy Jungmann.

[24] LÈVY, op. cit, p. 99.

[25] CHARTIER, Roger. As práticas da escrita. In: CHARTIER, Roger; ARIÈS, Philippe (org.). Histórias da Vida Privada: da renascença ao século das luzes. Da Renascença ao Século das Luzes. São Paulo: Companhia das Letras, 1991. 3 v. Tradução de Hidelgard Feist.p. 121.

barato e acessível, modificou as relações do homem com a divindade e do homem com o estado, dando nova dimensão à ideia de privacidade individual.

Diz Chartier[26]:

> Da maior ou menor familiaridade com a escrita depende, pois, uma maior ou menor emancipação, com relação a formas tradicionais de existência que ligam estreitamente o indivíduo a sua comunidade, que o imergem num coletivo próximo, que o tornam dependente de mediadores obrigatórios, intérpretes e leitores da Palavra divina ou das determinações do soberano.

A difusão de textos escritos em grande escala gerou sinergia entre estado e sociedade, reconduzindo as comunidades políticas para modelos cada vez mais burocráticos e técnicos — por assim dizer, artificiais —, em contraste com as antigas formas mais ligadas à natureza, ao hábito e à tradição. As Constituições foram o coroamento desse processo de reordenação do poder, mercê, entre outras coisas, da nova dimensão que a prensa trouxe para as mensagens escritas.

1.2.1 O jornalismo: um fruto imprevisto da prensa móvel

Outro padrão interessante que se constata no surgimento de novas tecnologias de comunicação é que elas geram microculturas que, por sua vez, produzem novas demandas e novos direitos.

O caso do jornalismo é um exemplo bem representativo. Subproduto cultural relevante da prensa móvel, com amplos reflexos políticos, culturais e econômicos, o jornalismo, embora pareça hoje algo absolutamente conatural à vida política de uma nação civilizada, é relativamente recente (séculos XVII em diante)[27].

A razão disso é prosaica: antes das novas técnicas de impressão em papel, por meio de tipos móveis, simplesmente não havia meios práticos para levar a informação, de modo preciso, rápido e regular até as pessoas. Embora houvesse veículos de notícia ao menos desde a *Acta Diurna*, de Júlio César, apenas com a prensa foi possível dar escala a essa espécie de produto cultural e, portanto, criar as condições objetivas para que houvesse demanda constante por informações.

O jornal foi um passo à frente do livro. Escrevendo em 1859, um então muito jovem literato brasileiro, que se tornaria depois a nossa maior expressão na Literatura — Machado de Assis — já percebia que o jornal, fruto mais amadurecido da prensa, modificaria ainda mais

---

[26] Ibidem.

[27] SOUSA, Jorge Pedro. **Uma história breve do jornalismo no Ocidente**. Porto: Edições Universidade Fernando Pessoa, 2008, p. 75.

que os livros não apenas a cultura, senão também outras relações sociais, que já vinham de várias revoluções mundo afora:

> O jornal é a verdadeira forma da república do pensamento. É a locomotiva intelectual em viagem para mundos desconhecidos, é a literatura comum, universal, altamente democrática, reproduzida todos os dias, levando em si a frescura das ideias e o fogo das convicções.(...) O jornal apareceu, trazendo em si o gérmen de uma revolução. Essa revolução não é só literária, é também social, é econômica, porque é um movimento da humanidade abalando todas as suas eminências, a reação do espírito humano sobre as fórmulas existentes do mundo literário, do mundo econômico e do mundo social.[28]

A prática social do jornalismo, ao incorporar-se ao cotidiano da vida, criou direitos inéditos, tais como a liberdade de imprensa, o direito à informação, a imunidade tributária dos insumos para a publicação de jornais, etc. Essa parece ser também uma tendência recorrente: as novas tecnologias de comunicação acabam por espraiar-se por diferentes áreas não antevistas, abrindo novos e inesperados campos que reclamam a atuação do direito.

1.2.2 Eletricidade e telecomunicações: tecnologia em rede e internacionalização

Se a escrita, o livro e a impressão tinham atendido à aspiração humana de superar a fugacidade da palavra oral, projetando o pensamento para muito além do seu tempo, outra ambição permanecia insaciada em pleno século XIX: o desejo de comunicar-se à distância em tempo real.

A primeira tentativa de criar um mecanismo artificial de comunicação à distância foi o telégrafo de Chappe[29], ainda no século XVIII. Era algo ainda bastante rudimentar, não distando muito de comunicações por fumaça ou tambores, já praticadas desde tempos imemoriais, exceto pelo fato de que era semafórico, isto é, usava símbolos para representar as mensagens.

Foi somente depois do domínio da eletricidade que se tornou possível a criação de artefatos capazes de mandar mensagens à distância com grande velocidade e precisão.

O primeiro deles foi o telégrafo elétrico. Em meados do século XIX, a produção e a transmissão da eletricidade já eram dominadas, e a ideia de que isso poderia ser usado para enviar e receber mensagens também já havia surgido. O problema maior estava em codificar (na origem) e decodificar (no destino) a mensagem. A eletricidade aparentemente não apresentava suscetível a modulação que permitisse expressar diferentes símbolos de comunicação (como o alfabeto, por exemplo). Até que Samuel Morse percebeu que, mesmo

---

[28] ASSIS, J.M. Machado de. **O jornal e o livro**. São Paulo: Companhia das Letras, 2011.

[29] GLEICK, James. **A informação**: uma história, uma enxurrada. São Paulo: Companhia das Letras, 2013. Tradução de Augusto Kalil, p. 138.

uma modulação mínima, tal como ligado/desligado, poderia ser suficiente para comunicar uma mensagem, desde que se transpusesse o alfabeto para um código com menor número de símbolos. E foi assim que a primeira fusão de eletricidade com palavras ocorreu, permitindo a transmissão de mensagens na velocidade da luz. Em maio de 1844, Morse transmitiu aquela que seria a primeira mensagem à distância com o uso da eletricidade, entre Baltimore e Washington[30].

O desenvolvimento da comunicação por telégrafo pressupunha uma rede de estações transmissoras e receptoras, ligadas por fios elétricos. As fronteiras dos estados, que delimitavam as soberanias políticas, não poderiam ser obstáculos para o avanço do telégrafo, inclusive porque era do interesse dos próprios estados e do comércio em geral que essas comunicações se amplificassem.

Assim, o telégrafo foi a tecnologia pioneira no processo de internacionalização das comunicações, juntamente com as estradas de ferro. Para os estados, essa internacionalização conduziu a um processo de mitigação crescente da ideia de soberania, especialmente no campo da técnica da telegrafia, pois seria impossível para qualquer país desenvolver um sistema idiossincrático de comunicação. Toda comunicação pressupõe enviar e receber mensagens, ou seja, uma rede. E o isolamento não é uma alternativa viável, pois seu custo é altíssimo, podendo levar mesmo à autodestruição da soberania[31]. Internamente, porém, os estados buscaram monopolizar a infraestrutura de comunicação telegráfica, reconhecendo desde logo o poder que ela projetava (ver, por exemplo, arts. 9º, §4º e art. 34, item 15 da Constituição brasileira de 1891).

O telégrafo deu ensejo para a criação da primeira organização internacional com poder de ditar regras para os estados soberanos: a *International Telegraph Union*, criada em Paris no ano de 1865, e que foi sucedida pela *International Telecommunication Union*, hoje vinculada à Organização das Nações Unidas[32].

Assim como a prensa de tipos móveis fizera séculos antes, o telégrafo elétrico também lançou sua influência sobre diferentes práticas sociais, além do próprio governo. O telégrafo mudou radicalmente a linguagem do jornalismo, por exemplo. As notícias, que eram enviadas

---

[30] HISTORY.COM EDITORS. Morse Code & the Telegraph. *In*: A&E TELEVISION NETWORKS. **HISTORY**. [*S. l.*], 6 jun. 2019. Disponível em: https://www.history.com/topics/inventions/telegraph. Acesso em: 30 mai. 2019

[31] CREVELD, Martin Van. **Ascensão e declínio do Estado**. São Paulo: Martins Fontes, 2004. Tradução de Jussara Simões, p. 544.

[32] INTERNATIONAL TELECOMMUNICATION UNION. Overview of ITU's History. In: INTERNATIONAL TELECOMMUNICATION UNION (org.). **ITU**: Committed to connecting the world. Disponível em: https://www.itu.int/en/history/Pages/ITUsHistory-page-2.aspx. Acesso em: 30 maio 2019.

em tempo real pelas estações telegráficas, tinham de ser objetivas e despojadas de tudo que não fosse essencial, inclusive porque o envio das mensagens era cobrado por caractere. Isso fez nascer o estilo enxuto de escrita que ainda hoje vemos ser ensinado nas escolas de jornalismo[33]. E até na Literatura teria havido reflexos. O estilo de Ernst Hemingway teria sido forjado na linguagem telegráfica[34].

Outro importante aspecto trazido pelo telégrafo foi a convergência de diferentes relações sociais para uma mesma mídia. Em 1845, com apenas um ano de funcionamento da primeira estação de telegrafia, Alfred Vail fez um levantamento de tudo que tinha sido transmitido, e encontrou os mais diferentes tipos de mensagens: eram mensagens comerciais; trocas de informação entre funcionários do governo americano, bancos, corretores, oficiais da polícia; notícias, resultados de eleições, deliberações de casos julgados em tribunais, convites, recibos, consultas médicas, etc[35].

Essa tendência agora se repete na internet. Há forte propensão para fazer convergir diversificadas formas de comunicação para a mesma mídia, fazendo com que haja ressonâncias entre áreas outrora isoladas entre si.

A mais profunda e disruptiva transformação provocada pelo telégrafo, no entanto, foi a separação definitiva entre comunicação e transporte. Pela primeira vez na história da Humanidade, as mensagens, os símbolos, tinham uma rede de transmissão independente daquela por onde corriam os transportes das coisas. As mensagens ganhavam, assim, uma autonomia inimaginável.

1.2.3 Telefone, rádio e televisão

Como peças que vão sucessivamente se encaixando, novas tecnologias de comunicação instantânea à distância foram sendo descobertas e implementadas.

O telefone surgiu em meados da década de 1870, aproximadamente trinta anos depois da primeira mensagem de telégrafo ter sido enviada. Tratava-se, na visão da época, de um "telégrafo falante"[36]. Muitos não acreditavam que essa tecnologia pudesse ter futuro. O telefone era considerado pouco sério, por não deixar nenhum registro escrito da mensagem enviada.

---

[33] CAREY, James W.. Technology and Ideology: the case of the telegraph. **Prospects**, [s.l.], v. 8, p. 303-325, out. 1983. Cambridge University Press (CUP). http://dx.doi.org/10.1017/s0361233300003793. Disponível em: http://faculty.georgetown.edu/irvinem/theory/Carey-TechnologyandIdeology.pdf. Acesso em: 27 maio 2019.
[34] Ibidem.
[35] GLEICK, op. cit., p. 158.
[36] Ibidem, p. 196.

Essas análises eram meramente especulativas. No instante em que os telefones caíram nas mãos das pessoas comuns, logo se tornaram febre. Em 1890, estima-se que havia aproximadamente 500 mil usuários de telefone em todo o mundo; em 1914 esse número havia saltado para algo em torno de 10 milhões. Em 1907, quando fizeram um recenseamento nos Estados Unidos, para avaliar quais setores dependiam do telefone, verificaram que o aparelho já era usado por empresas ligadas à mineração, à agricultura, ao comércio, aos transportes, e que até mesmo sapateiros, limpadores e lavadeiras já usavam o telefone[37].

Inesperado subproduto do telefone foi o início da emancipação feminina no mercado de trabalho. As centrais telefônicas, criadas para conectar e desconectar as linhas, conforme as ligações telefônicas fossem ocorrendo, implicou a criação da figura da telefonista[38].

O telefone acelerou o processo de comunicação instantânea que havia sido iniciado pelo telégrafo, engendrando novas formas de comportamento e de negócio, que acabaram sendo percebidos e regulamentados pelo estado. É sintomático constatar que, para a instalação da primeira companhia telefônica no Brasil, em 1879, foi necessária a autorização do Imperador[39]. Isso para não falar nos pesados investimentos públicos e na própria criação de empresa pública nacional para gerir o serviço de telefonia, na segunda metade do século XX.

As primeiras transmissões de rádio ocorreram no começo do século XX; as de televisão apenas ocorreram no final dos anos 1930. Duas importantes novidades técnicas vieram com o rádio e a televisão: 1º) foram tecnologias para cujo nascimento combinaram-se progressos que ocorriam simultaneamente em diferentes áreas do conhecimento[40]; 2º) ademais, o rádio e televisão criaram uma nova forma de transmissão de mensagens, por difusão de ondas, de modo a estabelecer uma comunicação unilateral, entre uma estação difusora e incontáveis ouvintes receptores.

Não cabe neste trabalho a investigação minuciosa sobre as vastas e numerosas consequências que o advento do rádio e da televisão trouxeram para vida social, política,

---

[37] GLEICK, op. cit., p. 200.

[38] SCIENCE MUSEUM. Goodbye To The Hello Girls: Automating The Telephone Exchang. *In*: SCIENCE MUSEUM GROUP. **Science Museum**. [*S. l.*], 22 out. 2018. Disponível em: https://www.sciencemuseum.org.uk/objects-and-stories/goodbye-hello-girls-automating-telephone-exchange. Acesso em: 31 maio 2019.

[39] OI FUTURO. **Museu das telecomunicações**. Disponível em: http://museudastelecomunicacoes.org.br/historia-das-telecomunicacoes/. Acesso em: 31 maio 2019.

[40] BURROWS, Charles. The History of Radio Wave Propagation up to the End of World War I. **Proceedings Of The Ire**, [s.l.], v. 50, n. 5, p. 682-684, maio 1962. Institute of Electrical and Electronics Engineers (IEEE). http://dx.doi.org/10.1109/jrproc.1962.288097. Disponível em: https://ieeexplore.ieee.org/document/4066757. Acesso em: 31 maio 2019.

econômica e cultural dos países. Dois pontos, no entanto, merecem destaque, por se relacionarem com o objeto da pesquisa.

Primeiro, no campo do entretenimento e do jornalismo, o rádio e a televisão criaram um fluxo de comunicação contínuo que ocupou, dentro das residências, papel unificador da sociedade. No século XX, desenrolava-se intenso processo de urbanização, que fazia nascer grandes cidades e, assim, reduzia-se progressivamente o papel da vizinhança e da família no processo de formação do indivíduo e de sua intimidade[41]. O rádio e a televisão passaram a ocupar esse papel. Não é difícil, assim, avaliar o enorme poder que os órgãos da mídia concentraram em suas mãos.

Segundo, o rádio e a televisão mostraram-se apropriados para o discurso político sedutor e incansável, especialmente o populista. Há estudo que faz essa ligação do rádio com o populismo, aqui no Brasil e na Argentina[42], em meados do século XX.

O direito reagiu ao rádio e à televisão de forma muito semelhante. No exemplo brasileiro ficou estabelecido, antes de tudo, que a radiodifusão seria uma atividade pública, suscetível de concessão a particulares (Constituição brasileira de 1934, art. 5º, VIII; Constituição brasileira de 1946, art. 5º, XII; Constituição de 1988, art. 21, XII, "a").

A percepção de que essas mídias tinham grande poder de conformação da sociedade certamente influenciou na concepção de normas que impediam — a ainda impedem — o acesso de estrangeiros ao controle de estações de rádio e televisão (Constituição de 1946, art. 160; Constituição de 1967, art. 166; Constituição de 1988, art. 222).

O estado brasileiro também arvorou-se no poder de censurar conteúdos: em 1937, por lei, a qualquer tempo, conforme o art. 122, item 15, "a" da Constituição então vigente; em 1946, quando decretado o estado de sítio, conforme art. 209, parágrafo único, I da Constituição brasileira de 1946; em 1967, por lei, a qualquer tempo, no interesse do regime democrático e do combate à subversão e à corrupção, conforme art. 166, §2º da Constituição de 1967; em 1988, em caso de estado de sítio, conforme art. 139, III da Constituição de 1988.

A reação do estado a essas tecnologias, portanto, foi de vigilância e controle. O estado tinha meios para isso e se dispôs a utilizá-los.

---

[41] SEVCENKO, Nicolau (org.). **História da vida privada no Brasil**. São Paulo: Companhia das Letras, 1998. 3 v, p. 585.

[42] HAUSSEN, Doria Fagundes. **Rádio e Política**: tempos de vargas e perón.. 1992. 324 f. Tese (Doutorado) - Curso de Doutorado em Ciências das Comunicações, Universidade de São Paulo, São Paulo, 1992.

1.3 Eletricidade e lógica: computadores e internet

Claude Shannon, no final dos anos 1940, disse que o problema fundamental da comunicação é reproduzir num determinado ponto, exata ou aproximadamente, uma mensagem selecionada num outro ponto, admitindo-se que as mensagens costumam ter um significado[43].

Isso foi escrito na sua tese de mestrado de 1948, que lançaria as bases para toda a revolução digital posterior. O problema da comunicação à distância, portanto, ainda ocupava as mentes mais férteis da ciência no pós-II Guerra. Mas havia outra questão que estava crescentemente sendo estudada: seria possível criar artificialmente um dispositivo capaz de imitar o pensamento humano?

Desde os anos 1930, os matemáticos já se ocupavam da questão da computação em nível abstrato. No começo dos anos 1900, o célebre matemático David Hilbert compilou vinte e três problemas não resolvidos, que serviriam como um programa de trabalho para a Matemática do século XX. Entre esses problemas, o décimo versava sobre saber se haveria como formular um procedimento formal e finito para demonstrar que certas equações, chamadas diofantinas, tinham raízes. A questão, embora um aparente arcano matemático, trazia à tona o problema da formalização do pensamento em uma linguagem extremamente abstrata.[44]

Alan Turing escreveu em 1936 um artigo[45] sobre o assunto, que na verdade não resolvia o problema proposto por Hilbert, mas inaugurava a ideia de "computablidade". Turing descreveu matematicamente como se poderia construir uma máquina capaz de computar. Essa máquina, que era então apenas uma ideia abstrata, pois não havia meios físicos de construí-la ainda, ficou conhecida como a Máquina de Turing, denominação essa que foi dada por Alonzo Church[46].

Estava dada, assim, a base teórica sobre a qual se poderia conceber um "computador". Faltavam, entretanto, meios físicos para concretizá-lo. É certo que já haviam sido construídos computadores mecânicos, mas eles não chegavam a operar com símbolos e lógica, senão com dispositivos materiais conectados entre si de forma apropriada.

---

[43] SHANNON, C. E.. A Mathematical Theory of Communication. **Mobile Computing And Communications Review**, S.l, v. 5, n. 1, p. 3-55, jan. 2001. Disponível em: https://culturemath.ens.fr/sites/default/files/p3-shannon.pdf. Acesso em: 21 maio 2019.

[44] ISAACSON, Walter. **Os inovadores**: uma biografia da revolução digital. São Paulo, Companhia das Letras, 2014. Tradução de Berilo Vargas, Luciano Vieira Machado e Pedro Maria Soares, p. 54-58.

[45] TURING, A. M.. On Computable Numbers, with an Application to the Entscheidungsproblem. **Proceedings Of The London Mathematical Society**, [s.l.], v. 2-42, n. 1, p. 230-265, jan. 1937. Disponível em: https://www.cs.virginia.edu/~robins/Turing_Paper_1936.pdf. Acesso em: 30 maio 2019.

[46] Ibidem, p. 58.

Em 1948 criou-se um dispositivo totalmente novo, que mudaria para sempre todas as tecnologias de produção, manipulação e transmissão da informação. Nos Laboratórios Bell, os cientistas conceberam algo que era semicondutor de energia, ou seja, algo que poderia modular a energia elétrica segundo um padrão semelhante aos estados ligado/desligado que permitira a comunicação pelo telégrafo. Não havia ainda um nome para a invenção e, depois de um concurso interno, escolheu-se o nome de Transistor[47] para designar aquele novo dispositivo

O semicondutor permitia a modulação da eletricidade, segundo um código binário previamente definido. Agora era possível, ao menos em tese, construir uma máquina, à base de eletricidade, que manipulasse informações e que efetuasse operações lógicas aproveitando-se de um fluxo de energia elétrica.

A evolução do computador logo se fez sentir e, dos anos 1950 em diante, máquinas com cada vez maior capacidade de computação foram sendo construídas à base de transistores interconectados: os *chips*.

Em paralelo, as redes de comunicação que agora se estabeleciam não interligavam mais apenas pessoas (como no telégrafo, no telefone, no rádio e na televisão), mas sim as próprias máquinas de computação entre si. Formava-se assim uma grande rede global de comunicação pessoa-máquina, pessoa-pessoa, máquina-máquina, que faria convergir para si uma massa gigantesca de informações digitalizadas: a internet.

A junção da internet com o computador fundiu todas as anteriores tecnologias da comunicação numa só rede. Então, todos avanços do telégrafo, do telefone, do rádio, da televisão e até do cinema e da fotografia, foram incorporados em um só *locus*, que de tão complexo que se tornou passou a ser designado como uma nova "realidade": o mundo virtual.

1.4 A indústria dos dados

No âmbito jurídico as palavras "dado" (ou "dados", no plural, como geralmente é usada) e "informação" são usadas com certa indiferença semântica, como nota Danilo Doneda[48].

A Lei Geral de Proteção de Dados contém vários usos indiscriminados dessas duas palavras. No art. 5º, I, por exemplo, ao definir o que é *dado pessoal*, a lei afirma que dado pessoal é a informação relacionada a pessoa natural identificada ou identificável.

---

[47] GLEICK, op. cit, p. 12.

[48] DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos de formação da lei geral de proteção de dados. 2. ed. São Paulo: Thomson-reuters Brasil, 2019, p. 136.

No entanto, a distinção exata entre esses dois vocábulos tem importância fundamental para que se construa adequada proteção jurídica aos direitos individuais e coletivos que podem ser atingidos pelo uso de máquinas inteligentes.

O dado é o material bruto, a partícula elementar que é utilizada no processo de comunicação. Assim, por exemplo, o conjunto das gravações eletromagnéticas de uma transação comercial realizada pela internet é constituído por dados. Ele tem potencial para gerar informação, a depender da capacidade de uma agente (humano ou eletrônico) para interpretá-los, segundo um código previamente definido. Mas, isoladamente, não é informação.

Quando os dados são lidos, mediante decodificação apropriada, eles então se transformam numa informação. No exemplo citado, alguma máquina poderia "ler" o conjunto de dados eletromagnéticos armazenado e inferir que houve certa transação comercial realizada, com todas as nuances da operação (sujeitos, objeto, data, forma de pagamento, etc.). A máquina poderia também expressar esse resultado numa linguagem imediatamente compreensível para seres humanos, manifestando a informação que extraiu dos dados.

Vê-se, portanto, que os dados são *commodities*, cujo valor cresce à medida que deles são extraídas informações por um agente com capacidade para decodificá-los. Há certa semelhança com o processo industrial de extração de riqueza de um minério

Como unidades elementares, os dados não apresentam grande valor se permanecerem isolados. Porém, quando eles são colocados em contato uns com os outros, se entre eles há um nó (por exemplo, todos podem se referir à mesma pessoa natural), e se há um processador (humano ou eletrônico) com potencial para correlacioná-los adequadamente, então esses dados podem gerar muitas informações, e estas, por sua vez, correlacionadas entre si, geram *conhecimentos*, que geram riqueza para quem os possui.

Credita-se a Clivy Humby, um matemático inglês, a melhor tradução do significado dos dados para a economia atual (chamada de 4.0). Segundo a frase atribuída a ele, "Os dados são o novo petróleo. É valioso, mas se não refinado, não pode realmente ser usado. Ele precisa ser transformado em gás, plástico, produtos químicos, etc. para criar uma entidade valiosa que impulsiona atividades lucrativas; assim, os dados devem ser discriminados, analisados para que tenham valor."[49]

É, portanto, na transformação dos dados em conhecimentos que reside o processo de geração de riqueza típico da economia 4.0. Nesse processo, sobretudo quando os dados

[49] GIACAGLIA, Giuliano. Data is the New Oil. In: HACKER NOON. **Hacker Noon**. [S. l.], 9 fev. 2019. Disponível em: https://hackernoon.com/data-is-the-new-oil-1227197762b2. Acesso em: 17 jul. 2020.

utilizados são pessoais — ou seja, se referem à vida, à liberdade ou ao patrimônio de uma pessoa natural —, podem ocorrer violações ao direito fundamental à proteção de dados pessoais, sobre os quais falarei mais adiante.

Fala-se hoje da Quarta Revolução Industrial. Klaus Schwab resume em três pontos as características dessa Revolução[50]:

a) Velocidade — a evolução das tecnologias agora se dá em razão exponencial, e não mais linear, em razão das numerosas interconexões geradas pela internet;
b) Amplitude e profundidade — além de as mudanças que ela produz serem amplas, atingindo praticamente todas as áreas do relacionamento humano, não se trata mais de apenas mudanças sobre o que fazer e como fazer, mas também sobre quem somos;
c) Impacto sistêmico — as mudanças afetam sistemas inteiros, dentro dos países e entre eles, na indústria, na economia e em toda a sociedade.

A Humanidade conhece no século XXI avanços semelhantes e disruptivos em três grandes temas: digital, biotecnologia e empregos. Os líderes em tecnologia estão avançando até aqui sem maiores preocupações com aspectos éticos. Mas, a médio e longo prazo, isso deve mudar.

As técnicas computacionais atualmente disponíveis criaram, de fato, aptidões até há pouco inimagináveis para gerar conhecimento a partir de dados. Como explica Laura Schertel Mendes[51], "o valor das informações obtidas não reside apenas na capacidade de armazenamento de grande volume de dados, mas, principalmente, na possibilidade de obtenção de novos elementos informativos a respeito dos cidadãos a partir do tratamento de dados."

Daí a importância crucial, para a pesquisa, de investigar minimamente os mecanismos dessas técnicas de produção de conhecimento a partir de dados, de modo a poder verificar quando elas operam fora do campo da legalidade.

Como o tema é novo, é preciso não apenas recensear os aspectos históricos que aproximam essas técnicas de experiências anteriores, mas se faz necessário também desenvolver a imaginação jurídica para que se compreenda toda a complexidade dos novos tempos e se possa formular soluções de maneira tal a fazer com que as tecnologias da informação sejam sensíveis a valores (*value sensitive design*)[52].

---

[50] SCHWAB, Klaus. **A quarta revolução industrial**. São Paulo: Edipro, 2016, Tradução de Daniel Moreira Miranda, p. 13.
[51] Apud MENDES, op.cit.
[52] FRIEDMAN, Batya; HENDRY, David G. **Value Sensitive Design**: shaping technology with moral imagination. Cambridge (ma): The Mit Press, 2019, p. 173.

Já existem alguns standards sendo apresentados por organismos internacionais. A Organização para a Cooperação e Desenvolvimento Econômico – OCDE lançou os seus princípios para o uso da Inteligência Artificial - IA. O texto aponta grandes linhas a serem seguidas, sem maiores detalhamentos técnicos, os quais sabidamente estariam sujeitos a obsolescências prematuras.

Seguem os princípios da OCDE para a Inteligência Artificial[53]- IA:

a) A IA deve beneficiar as pessoas e o planeta, impulsionando o crescimento inclusivo, o desenvolvimento sustentável e o bem-estar.
b) Os sistemas de IA devem ser concebidos de forma a respeitar o estado de direito, os direitos humanos, os valores e a diversidade democráticos e devem incluir salvaguardas adequadas - por exemplo, possibilitando a intervenção humana quando necessário - para assegurar uma sociedade honesta e justa.
c) Deve haver transparência e divulgação responsável em torno dos sistemas de IA para garantir que as pessoas entendam os resultados baseados em IA e possam desafiá-los.
d) Os sistemas de IA devem funcionar de maneira robusta, segura e segura ao longo de seus ciclos de vida, e os riscos em potencial devem ser continuamente avaliados e gerenciados.
e) Organizações e indivíduos que desenvolvem, implantam ou operam sistemas de IA devem ser responsabilizados por seu funcionamento adequado, de acordo com os princípios acima.

1.4.1 Dados digitais e seu tratamento

Num sentido amplo, os dados referem-se à circunstância de que alguns fatos da realidade podem ser representados ou codificados para posterior uso, leitura, processamento ou entendimento. Essa representação ou codificação se dá por meio da produção de dados.

No âmbito da computação eletrônica, os dados analógicos são aqueles cuja voltagem pode variar num intervalo, assumindo valores contínuos, isto é, sem "pular" de um número a outro. Já os dados digitais (binários) apenas podem assumir dois valores: 0 ou 1[54].

---

[53] OECD.ORG. **OECD**: better policies for better lives. Better Policies for Better Lives. Disponível em: https://www.oecd.org/going-digital/ai/principles/. Acesso em: 30 jun. 2020.

[54] SARPESHKAR, Rahul. Analog Versus Digital: extrapolating from electronics to neurobiology. **Neural Computation**, [s.l.], v. 10, n. 7, p. 1601-1638, out. 1998. MIT Press - Journals. Disponível em: https://ieeexplore.ieee.org/abstract/document/6790538. Acesso em: 10 jun. 2020.

Embora se possa fazer computadores analógicos, e muitos até digam que esse será o futuro da computação, atualmente os computadores são quase que exclusivamente digitais, por trabalharem com *microchips*, que são estruturados para esse tipo de dado.

A Lei Geral de Proteção de Dados – LGPD, em seu art. 1º, assume essa presunção técnica e determina que a sua disciplina é aplicável à proteção de dados digitais, querendo se referir evidentemente a dados que são manipulados em computadores eletrônicos. Como as decisões automatizadas são tomadas por computadores digitais, resulta que a LGPD incide sobre essas decisões, quando elas colidem contra a proteção de dados pessoais.

O uso da técnica de digitalização faz com que todos os movimentos em dispositivos eletrônicos (sons, imagens, textos, compras, navegação na internet, etc.) possam ser de algum modo tratadas uniformemente pelo computador como parcelas de uma operação aritmética, ainda que bastante complexa.

Os códigos binários que estão associados a movimentos de pessoas naturais nos computadores são os chamados dados pessoais. É a partir deles que, com as técnicas computacionais apropriadas, pode-se extrair informação e construir conhecimento. O fato de a digitalização reduzir informações completamente diferentes entre si a um formato padrão faz com que a manipulação desses dados torne-se muito mais praticável, e, ao mesmo tempo, oferece riscos às proteções que normalmente são colocadas no mundo físico aos dados que se consideram mais merecedores de proteção, em particular dados pessoais.

Assim, antes dos computadores e da coleta massiva de dados em formato digital, o indivíduo podia, por exemplo, levar a sua vida de tal maneira que o banco onde estava a sua conta-corrente não sabia onde ele fazia compras, para onde ele viajava ou com quem se comunicava. Além disso, ele podia também criar anteparos físicos para impedir o acesso seus dados: edifícios, caixas com fechaduras, cadeados, etc. Em suma, o indivídio gerenciava a maior parte dos seus próprios dados, e mesmo quando não gerenciava, os seus dados estavam fisicamente dispersos de tal maneira que não poderiam ser reunidos para uma avaliação conjunta.

Os dados em formato digital, diferentemente, têm homogeneidade física e geralmente convergem para a internet, com grandes possibilidades de desaguar em grandes bancos de dados (*Big Data*), onde podem ser devidamente reunidos e cruzados para criar conhecimentos e, consequentemente, riqueza para quem os detém.

Convém notar que a digitalização tem vastos efeitos sobre a vida em sociedade, não apenas no seu aspecto econômico, senão também no político, no cultural e no social. Angela Maria Barreto[55] observa, com propriedade:

> (...) Entretanto, a relação com as informações digitalizadas apresenta-se com um caráter agressivo que substitui a atitude reflexiva, requerida na significação. Neste aspecto, é bom assinalar que a memória digital é universalizante, gerada num espaço e num tempo não experenciados, sem vínculo humano presencial. Trata-se de um vínculo coletivo, virtual, em rede, com informações produzidas por muitos sujeitos, distantes uns dos outros, em contextos socioculturais que lhes dão referências diversas. Mais ainda, sua difusão ocorre de forma desordenada, além de que a memória digital é diferente da memória da significação. Enquanto aquela é universalizante, esta é particularizante, estando ligada à vida dos seres, ao sensível, pois que implica afeto, emoção, vínculos, formas sociais de convívio num ambiente materializado pelos objetos e espaços criados diferenciadamente por cada grupo de pessoas.

A LGPD considera o tratamento de dados como todo o ciclo de manipulação dos dados, desde a sua coleta até o seu processamento e eventual transmissão. Essa posição está de acordo com a natureza desses dados, pois o simples fato de algo ser digitializado potencialmente coloca esse dado na linha de produção de conhecimento própria do ambiente virtual.

Com efeito, diz o art. 5º, X da LGPD que tratamento de dados é:

> (...) toda operação realizada com dados pessoais, como as que se referem a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração.

Merece especial destaque, no tratamento de dados, a operação de *processamento*. Nela está o cerne dos métodos de produção de conhecimento a partir de dados.

O processamento pode ser feito por muitas técnicas e com muitas finalidades. Despontam, porém, a Inteligência Artificial e seus subconjuntos como os meios mais poderosos de geração de valor a partir de dados.

É essencial, por isso, compreender como funciona em geral o que se chama de Inteligência Artificial. É por meio dela que se formam as decisões automatizadas.

1.4.2 Inteligência Artificial.

Não há uma definição precisa para o que seja Inteligência Artificial (AI, na sigla em inglês para *Artificial Intelligence*). Uma das primeiras tentativas de definir Inteligência

[55]BARRETO, Angela Maria. Informação e conhecimento na era digital. **Transinformação**, Campinas, v. 17, n. 2, p. 111-122, maio 2005. Disponível em: https://www.redalyc.org/pdf/3843/384334739002.pdf. Acesso em: 06 jun. 2020.

Artificial foi feita em 1955, por John McCarthy, um dos pioneiros dos estudos nessa área. Segundo ele, "o objetivo da Inteligência Artificial (AI) é desenvolver máquinas que se comportem como se fossem inteligentes"[56]. Trata-se, como se vê, de uma definição circular, pois justamente o que se quer saber é em que consiste o comportamento inteligente.

Elaine Rich[57] tem uma definição melhor, embora a ela também se possa opor objeções: Inteligência Artificial é o estudo de como fazer computadores fazerem coisas nas quais, no momento, as pessoas são melhores.

A definição apanha bem o caráter eminentemente comparativo das ações da AI em relação à prática de atividades por humanos, e não ao modo de funcionamento do cérebro humano, ou aos demais atributos de um ser humano. Outro ponto forte da definição é que ela acentua que a comparação se dá no campo da eficiência e não da natureza do trabalho, ou seja, a AI é uma ferramenta eminentemente econômica, e, como tal, requer regulamentação quando produz danos colaterais a direitos.

No fundo, a expressão Inteligência Artificial acaba glamorizando demais um procedimento que é tão-só resultado de regras do senso comum aplicadas em massa e na velocidade da luz.

Para além dos seus aspectos técnicos, a AI é um grande negócio. Por meio dela, muita riqueza tem sido gerada ao redor do mundo e é consenso entre os estudiosos que a economia 4.0 gravitará em torno da Inteligência Artificial e suas ferramentas associadas[58].

Mas a AI não é uma entidade homogênea: há diferentes capacidades cognitivas que são mimetizadas pelos computadores. Saber distinguir os vários tipos de AI tem importância porque essas variantes podem apresentar maior ou menor risco de danos a direitos. Além disso, ao compreender o processo de formação da manifestação da máquina, tem-se condição de exercer alguma crítica sobre o resultado do processo (o *output*), se for o caso.

Adiante são apresentadas as técnicas mais utilizadas em Inteligência Artificial para produzir conhecimento a partir de dados.

1.4.2.1 Aprendizado de Máquina (Machine Learning – ML)

---

[56] ERTEL, Wolfgang. Introduction. **Undergraduate Topics In Computer Science**, [s.l.], p. 1-21, 2017. Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-58487-4_1, p. 1.

[57] RICH, Elaine apud Ibidem.

[58] AGRAWAL, A; GANS, J.; GOLDFARB. **Máquinas preditivas***:* a simples economia da Inteligência Artificial. Rio de Janeiro: Alta Books, 2018. Tradução de Wendy Campos, p. 11.

Há divergência sobre se o *Machine Learning* é um subcampo da AI ou se é um conjunto paralelo de inteligência de máquina[59]. Isso, porém, não tem maior relevância para esta pesquisa, de modo que se admite, para todos os efeitos, que o ML é subconjunto da AI.

Arthur Samuel, em 1959, foi um dos primeiros a usar o termo *machine learning* num artigo em que discutia a possibilidade de um computador "aprender" a jogar damas melhor do que o seu programador. Ou seja, se seria possível que a máquina fizesse uma atividade sem ter sido explicitamente programada para tal.[60]

O ponto central do *machine learning* é a capacidade de autoaprendizado, com base em um conjunto de dados, sem uma programação específica para cada situação que é apresentada ao computador. Esse tipo de técnica confere certa imprevisibilidade aos resultados, de maneira que decisões automatizadas tomadas com base em ML não são preconizadas detalhadamente na programação prévia.

Em vez de o *input* (os dados de entrada) ligar-se diretamente a um comando para gerar o *output* pré-programado, no *machine learning* o *input* é o próprio conjunto de dados em bruto, que será tratado por um modelo que produzirá o *output*, segundo relações de inferência estatística e outros métodos de cálculo matemático. Isso quer dizer que o *output* não é previamente conhecido sequer do programador, pois depende do conjunto de dados apresentado.

Uma característica importante do *machine learning* é a sua capacidade de aperfeiçoamento com a "experiência". À medida que o modelo é mais e mais treinado com diferentes conjuntos de dados, os *outputs* tendem a apresentar mais acurácia. Nas rígidas programações pré-definidas, isso não ocorre.

Nessa característica do ML está a sua virtude e o seu vício. O fato de o resultado depender do conjunto de dados apresentado à máquina torna o programa suscetível ao enviesamento. Se o modelo é treinado com um conjunto de dados muito homogêneo, ele pode gerar distorções ao deparar com dados diversos (*underfitting*); por outro lado, se ele for treinado para abranger dados muito diversos, então ele pode gerar resultados errados, por excesso de adaptação aos dados (*overfitting*).

---

[59] SKANSI, Sandro. Introduction to Deep Learning: from logical calculus to artificial intelligence. **Undergraduate Topics In Computer Science**, [s.l.], v. 1, n. 1, p. 1-196, jan. 2018. Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-73004-2.

[60] BOWLING, Michael; FÜRNKRANZ, Johannes; GRAEPEL, Thore; MUSICK, Ron. Machine learning and games. **Machine Learning**, [s.l.], v. 63, n. 3, p. 211-215, 10 maio 2006. Springer Science and Business Media LLC. http://dx.doi.org/10.1007/s10994-006-8919-x.

Normalmente, um mecanismo de *machine learning* passa por uma fase de treinamento prévia, na qual um conjunto relativamente amplo de dados é apresentado ao programa e são feitos ajustes no modelo para excluir previsões errôneas e, assim, refinar o programa antes que ele entre em operação real. Esses ajustes, por seu turno, são também tarefas que podem ser assumidas por máquinas.

Está claro que o desenvolvedor do programa tem um poder enorme de conformação do modelo de aprendizado de máquina na fase de treinamento. Não é exato, portanto, imaginar que uma decisão automatizada gerada por um modelo de ML é infenso a ideologias ou a interesses escusos. Como diz Cathy O´Neal[61], "modelos são opiniões embutidas em matemática".

O ML tem muitos usos comerciais. Ele é apto para ser utilizado, por exemplo, na análise de risco de crédito, na detecção de fraude e no gerenciamento de portfólio em serviços financeiros campanhas de marketing direcionadas. Na verdade, todas as situações a respeito das quais há um longo histórico de dados passados, sendo o comportamento futuro constituído por reiterações de comportamento, podem ser proveitosamente abrangidas pelo ML.

O aprendizado de ML, por isso mesmo, pode melhorar na conclusão de tarefas ao longo do tempo com base nos dados rotulados que ele ingere, ou pode impulsionar a criação de modelos preditivos para melhorar uma infinidade de tarefas.

1.4.2.2 Aprendizado Supervisionado

No chamado aprendizado supervisionado, intenta-se encontrar padrões a partir de um conjunto de dados rotulado (estruturado), de modo a estabelecer relações entre *inputs* (variáveis independentes) e seu *output* conhecido (variável dependente). As variáveis independentes impactam a varável dependente.

Uma aplicação desse tipo de aprendizado pode ocorrer na criação de um modelo para avaliar preços de carros usados. Após a introdução das variáveis independentes conhecidas (ano do carro, modelo, quilometragem, revisões, etc.), o programador associa os *outputs* conhecidos (ou seja, os preços respectivos). Após um tempo de treinamento (experiência), o modelo estará pronto para entrar em operação em casos reais.

Nesse aprendizado, portanto, o programador "ensina" a máquina a "julgar" e, depois de apresentar muitos exemplos, sob sua supervisão, coloca em prática o modelo.

---

[61] O'NEIL, Cathy. **Weapons of math destruction:**: how big data increases inequality and threatens democracy. New York:: Crown Publishers, 2016. E-book., p. 21

É evidente que nesse tipo de técnica, como aliás em todas as de ML, o conjunto de dados da fase de treinamento é determinante para os resultados que serão apresentados pela máquina na fase de operação real.

2.4.2.3 Aprendizado não supervisionado

Aqui o *output* variável não é etiquetado e também não são previamente conhecidas as relações entre *inputs* e *outputs*. Pelo contrário, o programador quer que a máquina descubra os padrões ocultos.

O aprendizado não supervisionado concentra-se em analisar as variáveis de entrada e descobrir padrões ocultos que podem ser extraídos para criar novas rotinas relacionadas a possíveis saídas.

A vantagem do Aprendizado não Supervisionado é que ele habilita o usuário a descobrir padrões ocultos nos dados sobre os quais não tinha consciência, o que dá condições para que se avance na análise dos dados e novos grupos sejam descobertos. Ele pode, depois, gerar um modelo de Aprendizado Supervisionado.

O Aprendizado Não Supervisionado é especialmente interessante no domínio da detecção de fraudes, em que os ataques mais perigosos ainda não estão classificados.

2.4.2.4 Aprendizado por reforço (*reinforcement learning*)

É o terceiro tipo e o mais avançado dos aprendizados de máquina. Diferentemente do Aprendizado Supervisionado e do Não Supervisionado, o Aprendizado por Reforço constrói um modelo de predição por ganhos de *feedback* a partir de processos randômicos de tentativa e erro e alavancagens por *insights* decorrentes de interações anteriores.

A melhor analogia para compreender esse aprendizado é feita com um jogador de videogame. À medida que ele avança no jogo, ele aprende os valores positivos e negativos das ações sob diferentes condições. Esse aprendizado influencia seu comportamento futuro e assim a sua performance gradualmente melhora com base no aprendizado e na experiência. O Aprendizado por Reforço funciona assim também em carros autônomos e em jogo de xadrez, por exemplo.

2.4.2.5 Aprendizado Profundo (*Deep Learning*)

O aprendizado profundo tenta imitar mais de perto a mente humana. Os problemas que são apresentados ao computador são analisados em várias camadas, com o intuito de simular o

processo de pensamento humano. O cérebro humano, de fato, é capaz de sintetizar um vasto conjunto de dados de diferentes origens (sons, imagens, textos, etc) em uma só ou poucas informações.

Ao enfrentar uma situação em que uma solução está escondida em um grande conjunto de dados (*Big Data*), o aprendizado de máquina é uma ótima opção. A capacidade de cálculo do computador torna possível a tarefa de tratar rapidamente conjuntos enormes de dados, revelando padrões por vezes insuspeitos para o ser humano.

O aprendizado profundo normalmente utiliza redes neurais artificiais (*Artificial Neural Networks* – ANN), que absorvem padrões durante a fase de treinamento, mediante o oferecimento de dados rotulados e a chave de resposta pelo programador. As ANN então "aprendem" qual a saída correta a partir de certos tipos de entrada.

1.5 O ecossistema digital

Walter Ong, em trabalho postumamente publicado[62], observou que a Idade em que a existência humana está agora imersa, uma Idade em que a vida e a tecnologia humanas interagem de maneira massiva e íntima, não deve ser pensada apenas como uma Era da Informação, mas antes como uma Era Ecológica.

Isso porque a total interconexão entre as pessoas e as coisas e até a entre as coisas faz com tudo esteja relacionado com tudo, formando um autêntico ecossistema, no qual há influências recíprocas entre indivíduos, populações e o ambiente.

O estudo do chamado "comportamento de máquina" (*machine behaviour*) segue nessa direção.[63] Nessa linha de investigação, acredita-se que se pode usar, por analogia, os estudos biológicos sobre o comportamento animal e suas interações, para avaliar os impactos das máquinas sobre os humanos e vice-versa.

Seja como for, é essencial entender as características das máquinas inteligentes, bem como os efeitos delas sobre o comportamento humano, individual e coletivo. As decisões automatizadas podem trazer grandes benefícios e comodidades para a humanidade; no entanto, é preciso que haja salvaguardas contra possíveis efeitos negativos da introdução desses agentes na vida social. Apenas avaliando concretamente essas interações é possível construir uma teoria

---

[62] ONG, Walter. **Language as hermeneutic**: a primer on the word and digitization. Ithaca and London: Cornell University Press, 2017.

[63] RAHWAN, -Iyad et al. *Machine behaviour*. **Nature**, [s. l.], v. 568, p. 477–486, 24 abr. 2019. Disponível em: https://www.nature.com/articles/s41586-019-1138-y. Acesso em: 2 jun. 2020.

que possa antecipar virtudes e defeitos do relacionamento entre humanos e máquinas inteligentes.

2 DO DIREITO À PROTEÇÃO DE DADOS: EVOLUÇÃO E CARACTERÍSTICAS

2.1 A evolução do direito à privacidade

O direito à proteção de dados nasceu da evolução do direito à privacidade, como decorrência de transformações tecnológicas.[64] Se, por muitos séculos, os diários, as cartas escritas em papel e o respectivo sigilo postal representaram o máximo que se poderia pensar em termos de privacidade da expressão humana, no século XIX o surgimento da fotografia, do telégrafo, do telefone e a ampliação da influência dos jornais, trouxeram à luz problemas novos a respeito do conceito de privacidade.

O célebre artigo de Louis Brandeis e Samuel Warren[65] sobre o direito à privacidade (*right to privacy*) representa, nesse sentido, a tomada de consciência dos juristas sobre a necessidade de ampliação das proteções jurídicas à pessoa, em face de novas técnicas — que então surgiam — de produção, tratamento e difusão de informações. Embora focado na tradição da *common law*, o artigo apresentava *insights* que podiam ser generalizados para países filiados à *civil law*.

O tema que ocupava particularmente as preocupações de Brandeis e Warren era o da circulação de fotografias e notícias sobre a intimidade das pessoas por meio de jornais, sem a autorização dos indivíduos retratados ou referidos. Numa impressionante antecipação de vários dos problemas que a livre circulação da informação poderia trazer para a privacidade individual, Brandeis e Warren citam parcialmente o versículo bíblico que, de certo modo, resume a ameaça que pesa sobre o sujeito ao qual não é dado controlar o âmbito de repercussão de suas próprias informações: “Porquanto tudo o que em trevas dissestes, à luz será ouvido; e o que falastes ao ouvido no gabinete, sobre os telhados será apregoado.” (Lc 12,3).

Após analisarem vários casos que correram à época em tribunais ingleses e americanos sobre a questão, e mostrarem como o esquema do direito de propriedade não respondia mais satisfatoriamente às exigências reclamadas pela prática judiciária, Brandeis e Warren admitem que não é fácil estabelecer a linha exata entre o direito à privacidade e o bem-estar público, mas apontam que os *standards* interpretativos para casos assim deveriam ser buscados não no direito de propriedade, mas sim em institutos jurídicos como os crimes contra a honra e os direitos autorais. Nessa linha, eles apontaram alguns vetores interpretativos que deveriam ser levados

---

[64] MENDES, op. cit. p. 27.

[65] WARREN, Samuel D.; BRANDEIS, Louis D.. The Right to Privacy. **Harvard Law Review**, Cambridge, v. 4, n. 5, p. 193-220, 15 dez. 1890. Disponível em: https://www.cs.cornell.edu/~shmat/courses/cs5436/warren-brandeis.pd. Acesso em: 03 set. 2020.

em conta para avaliar se o direito à privacidade estaria ou não sendo violado: a) o direito à privacidade não impede a publicação de matéria que seja de interesse público; b) o direito à privacidade não veda a publicação de matéria sempre que haja circunstância legal que torne a publicação legítima, em situações como discursos em assembleias legislativas ou votos em tribunais; c) não há direito à reparação quando se trata de revelação verbal que não cause maiores danos; d) o direito à privacidade cessa quando a publicação é feita pelo próprio indivíduo afetado ou com o seu consentimento; e) a alegação de que é verdadeira a publicação não exclui a violação ao direito à privacidade; f) a ausência de dolo não exclui a violação ao direito à privacidade[66].

A doutrina acentua que a visão inicial sobre a privacidade, da qual o artigo de Brandeis e Warren é exemplo modelar, esteve essencialmente concentrada em pessoas de elevada projeção social, com o propósito de criar-lhes espaços de isolamento e tranquilidade[67]. Ou seja, espaços em que essas pessoas pudessem estar a sós (*right to be let alone*).

Porém, transformações econômicas, políticas e também da técnica informacional, ocorridas sobretudo a partir da década de 1960, implicaram a demanda cada vez maior, por pessoas comuns, pela proteção da sua privacidade. Como diz Mikhail Cancelier[68], "[na década de 1960] com velocidade considerável, o direito à privacidade vai expandindo suas fronteiras, alcançando novos sujeitos, englobando diferentes objetos e tornando-se presente em locais com ele antes incompatíveis."

É certo que, desde o século XIX (e até antes disso), havia já pelo menos dois direitos individuais associados à ideia de privacidade que já eram previstos em favor de todos os cidadãos: a inviolabilidade do domicílio; e o sigilo da correspondência. A Constituição do Império do Brasil, de 1824, por exemplo, já previa essas duas prerrogativas individuais (Constituição Imperial, art. 179, VII e XXVII). Assim também a Constituição brasileira de 1891 estipulava essas duas prerrogativas individuais na sua Declaração de Direitos (CR/1891, art. 72, §§11 e 18).

Como se vê, a privacidade pode ser ligada a diferentes elementos individuais e domésticos. O que conecta, por exemplo, a inviolabilidade do domicílio ao sigilo da correspondência? Decerto não é o "direito de estar só", que é mais próprio do domicílio, mas

---

[66] Ibidem, p. 214-218.
[67] DONEDA, op. cit., p. 32-33.
[68] CANCELIER, Mikhail Vieira de Lorenzi. O direito à privacidade hoje: perspectiva histórica e o cenário brasileiro. **Seqüência**: Estudos Jurídicos e Políticos, Florianópolis, v. 38, n. 76, p. 213-240, 20 set. 2017. Disponível em: https://periodicos.ufsc.br/index.php/sequencia/article/view/2177-7055.2017v38n76p213/34870. Acesso em: 04 set. 2020.

sim uma prerrogativa subjacente de ocultar do público certos aspectos da vida individual e familiar — o que, pensando em termos mais abstratos, é o mesmo que o direito de controlar quem deve ter acesso a informações pessoais do titular.

Assim é que, pela amplitude dos temas que abarca, a própria terminologia para melhor designar o direito à privacidade permanece ambígua. Para nomeá-lo, fala-se de "privacidade", "intimidade", "vida privada", "sigilo", "segredo", etc. Danilo Doneda, após mostrar que a imprecisão do conceito de privacidade não é exclusividade de nenhum tempo ou país, afirma que se deve tomar tal indeterminação "como característica ontológica da própria construção da esfera privada que pode ajudar a nortear o nosso campo de estudo [referindo-se ao direito à privacidade]"[69].

A Constituição brasileira de 1988 ilustra bem o problema. No art. 5º, X, a CF/88 usa quatro diferentes expressões para a privacidade (intimidade, vida privada, honra e imagem), quando afirma: "são invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou moral decorrente de sua violação."
Ao tratar da inviolabilidade do domicílio (art. 5º, XI), a Carta de 1988 usa expressão consagrada desde a Constituição de 1824, dizendo ser a casa "asilo inviolável", isto é, um refúgio bem protegido — e aqui há a mais forte aproximação com o clássico sentido do "direito de ser deixado só".

Cuidando das comunicações postais, de dados, telegráficas, ou telefônicas, a Constituição (art. 5º, XII) usa a palavra "sigilo" para se referir à privacidade respectiva: "é inviolável o sigilo da correspondência e das comunicações telegráficas, de dados e das comunicações telefônicas, salvo, no último caso, por ordem judicial, nas hipóteses e na forma que a lei estabelecer para fins de investigação criminal ou instrução processual penal". Talvez aqui esteja o maior deficit semântico, consideradas as atuais circunstâncias históricas, porque o direito á proteção de dados, que adiante será apresentado, está longe de esgotar-se apenas na ideia de "sigilo".

As palavras "intimidade", "sigilo" e "segredo" são usadas mais vezes pela Constituição de 1988 para traduzir a ideia de privacidade: art. 5º, LX; art. 5º, LXXII; art. 93, IX; art. 136, §1º, *b* e *c*; art. 139, III.

E, em outras ocasiões, a Constituição refere-se de modo elíptico à privacidade, como o oposto de algo que dependa de autorização estatal, usando o adjetivo "livre". Assim, por exemplo, no art. 226, §7º, diz-se que "o planejamento familiar é livre decisão do casal,

[69] DONEDA, op. cit., p. 102.

competindo ao Estado propiciar recursos educacionais e científicos para o exercício desse direito, vedada qualquer forma coercitiva por parte de instituições oficiais ou privadas."

Já o Código Civil brasileiro de 2002 prefere usar a expressão "vida privada" para designar toda a esfera das manifestações vitais do ser humano que não são apresentadas propositalmente no espaço público. No art. 21 do CC/2002 está dito: "A vida privada da pessoa natural é inviolável, e o juiz, a requerimento do interessado, adotará as providências necessárias para impedir ou fazer cessar ato contrário a esta norma."

A polissemia aponta a profunda complexidade do conceito de privacidade e as múltiplas formas de expressão sob as quais ele se apresenta. É vão e pernicioso buscar extrair uma definição apodíctica de privacidade. Na ausência de regras particulares legitimamente delineadas pelo legislador, o importante é ter presente que o livre desenvolvimento da personalidade humana, a garantia da autonomia individual e a prerrogativa de ocultar do público comportamentos idiossincráticos são os vetores interpretativos fundamentais para apreciar, sob a luz das nuances históricas concretas, em que medida alguma situação está ou não sob a proteção do direito à privacidade[70]. Como apontaram Warren e Brandeis no seu artigo seminal, o que se dever ter em mente são balizas interpretativas que aproximem a privacidade dos direitos da personalidade e não do direito de propriedade.

Um elemento que aumenta significativamente a dificuldade para definir o âmbito da privacidade é que o seu espectro varia não apenas de um lugar para outro, ou de um tempo histórico para outro, mas até mesmo de uma pessoa para outra, pois cada um tem um limiar de suscetibilidade à curiosidade alheia, e não são raros os casos em que a pessoa voluntariamente expõe-se ao público para obter algum ganho (econômico, político, artístico, etc), ou para simplesmente satisfazer a alguma aspiração sua. Uma concepção justa de privacidade, portanto, precisa ser elástica, admitindo modulações que permitam ajustá-la a cada contexto específico.

O Supremo Tribunal Federal já firmou, por exemplo, que o político não pode invocar, nos crimes contra a honra, proteção do mesmo nível daquela que é oferecida ao cidadão comum, que não oferece o seu nome ao escrutínio público. O Ministro Sepúlveda Pertence, em voto que proferiu no Inquérito nº 503-7-RJ[71], observou:

> É certo que, ao decidir-se pela militância política, o homem público aceita a inevitável ampliação do que a doutrina italiana costuma chamar a *zona di iluminabilitá*,

[70] SCHREIBER, Anderson. **Direitos da personalidade**. 2. ed. São Paulo: Editora Atlas, 2013, p. 13.
[71] BRASIL. Tribunal Pleno do Supremo Tribunal Federal. Inquérito nº 503-RJ. Relator: Ministro Sepúlveda Pertence. Brasília, DF, 24 de junho de 1992. **Diário de Justiça**. Brasília, 23 mar. 1993.

> resignando-se a uma maior exposição de sua vida e de sua personalidade aos comentários e à valoração do público, em particular, dos seus adversários.

A importância de contexto é tão acentuada, no tema da privacidade, que as mudanças tecnológicas ocorridas no século XX alteraram radicalmente o conceito que se tinha desse direito. Em vez da visão individualista que buscava desligar o indivíduo do grupo, passou-se a cogitar de uma nova concepção de privacidade, mais como o poder de governar as próprias informações em mãos de terceiros[72]. Tratava-se da aurora do direito à proteção de dados pessoais, embora ainda sem esse nome.

Bem entendido, essas mudanças não eliminam as anteriores garantias da privacidade dentro dos respectivos contextos, mas sim criaram novas demandas por direitos mais ajustados às novas condições de vida em sociedade, proporcionadas pela evolução técnica, cultural e política. Daí nasceria o direito à proteção de dados pessoais, como um corolário dos fundamentos estabelecidos para a ideia de privacidade dentro da atmosfera criada pelos computadores e pela internet.

Como aponta Stefano Rodotá[73], a histórica definição de privacidade de Warren e Brandeis, como "o direito de ser deixado em paz", foi sucedida por muitas outras, sem exclusão recíproca, que incorporaram circunstâncias novas, tais como: "o direito a controlar a maneira na qual os outros utilizam as informações a nosso respeito" (A. Westin); "a proteção de escolhas de vida contra qualquer forma de controle público e estigma social" (L. M. Friedman); "a reivindicação dos limites que protegem o direito de cada indivíduo a não ser simplificado, objetivado, e avaliado fora de contexto" (J. Rosen)[74]. O próprio Stefano Rodotá propõe uma definição: "o direito de manter o controle sobre suas próprias informações e de determinar a maneira de construir sua própria esfera particular"[75].

Assim como à época de Warren e Brandeis a propriedade já não respondia convenientemente às necessidades históricas específicas de proteção do indivíduo — embora continuasse a ser um direito importante para certas relações sociais —, a noção de privacidade estabelecida no século XIX já não atende às exigências do tempo presente, conquanto permaneça sendo um direito relevante em determinadas situações.

De fato, a partir dos anos 1970, a privacidade começa a associar-se ao problema dos bancos de dados públicos. Assim como o jornalismo fotográfico havia sensibilizado juristas e

---

[72] MENDES,op. cit., p. 29.
[73] RODOTÁ, Stefano. **A vida na sociedade da vigilância**: a privacidade hoje. Rio de Janeiro: Renovar, 2008. Tradução Danilo Doneda e Luciana Cabral Doneda, p. 15.
[74] Apud Ibidem, p.12.
[75] Ibidem, p. 15.

tribunais na virada do século XIX para o XX, a manipulação de grande volume de dados pessoais pelo Estado despertou a atenção de legisladores, juízes e estudiosos do direito em geral. É de 1970, por exemplo, a primeira lei americana a disciplinar diretamente a *privacy*[76]. Trata-se da *Fair Credit Reporting Act*, que disciplinava a atividade das empresas de cadastro de proteção ao crédito. Em 1974 foi aprovado, também nos Estados Unidos, o *Privacy Act*, para a disciplina da coleta, guarda, uso e disseminação de informações pessoais dos cidadãos pela administração pública.

Essa primeira geração de normas sobre a proteção de dados pessoais, como explica Laura Mendes, "surgiu como reação ao processamento eletrônico de dados nas administrações públicas e nas empresas privadas"[77], bem como à ideia de centralização do tratamento de dados em gigantescos bancos de dados.

Se o direito à privacidade começou a receber tratamento sistemático com o propósito de posicioná-lo ante a ampla liberdade de expressão da imprensa e o surgimento do fotojornalismo, o seu desenvolvimento posterior foi mais particularmente afetado pela ampliação do poder estatal e pelas possibilidades técnicas criadas pelos computadores para o tratamento automatizado de dados.

Assim, o direito à privacidade, que nascera como direito do indivíduo oponível a outros indivíduos ou corporações particulares, agora mostrava outra face: a da prerrogativa contra ameaças provenientes da vigilância estatal. Dessa vez, não foi principalmente obra da doutrina despertar a atenção para o problema, mas sim do Tribunal Constitucional Alemão, num julgamento pioneiro, ocorrido em 1983, sobre uma lei que disciplinava o censo demográfico[78].

Se Warren e Brandeis foram sensibilizados para o direito à privacidade pela bisbilhotice do que poderíamos chamar de primeiros colunistas sociais, o Tribunal Constitucional alemão estava diante de um problema bem mais amplo, de caráter político, que opunha a privacidade ao poder de ordenação do estado, mediante o uso massivo de informações pessoais dos indivíduos.

O caso chegou ao Tribunal Constitucional em razão de várias queixas constitucionais apresentadas por cidadãos, que contestavam a Lei do Censo Federal de 1983, alegando que ela violava diretamente os arts. 1, I e 2, I da Lei Fundamental de Bonn.

---

[76] DONEDA, op. cit., p.127-128.
[77] MENDES, op.cit., p. 29.
[78] Para um resumo do julgamento, Cf. https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/1983/12/rs19831215_1bvr020983en.html. Acessado em 08 set. 2020; Cf. MENDES, op. cit., p. 30-32.

A norma impugnada previa a coleta ampla de dados pelos agentes do censo. Além da tradicional contagem da população, o censo deveria coletar também dados sobre o nome, endereço, opção religiosa, entre outros, bem como informações sobre a formação educacional dos sujeitos do censo, ocupação profissional e situação de moradia. Ademais, foi estipulada multa para quem não respondesse ao censo e criou-se a possibilidade de compartilhamento e cruzamento dos dados coletados com outros que constassem dos bancos de dados já existentes, de modo a aferir a veracidade das informações.[79]

As queixas foram admitidas e o Tribunal declarou a lei constitucional, com exceção dos dispositivos que previam o cruzamento e compartilhamento dos dados. Considerou-se que o direito geral de personalidade[80] abrange, o conceito de "autodeterminação informativa", no que se inclui o poder de decidir em que âmbito divulgará as suas informações pessoais. O cruzamento e o compartilhamento dos dados pessoais, na visão do Tribunal, violavam essa autodeterminação porque os indivíduos não saberiam com segurança o alcance que teriam as informações por eles prestadas.

O fundamento principal do julgado, portanto, estava em que a falta de certeza dos indivíduos sobre o tipo de informação pessoal sua que seria conhecida de terceiros prejudicava de modo relevante a liberdade de autodeterminação. Desse modo, no contexto do processamento de dados, o livre desenvolvimento da personalidade de alguém pressupõe que o indivíduo seja protegido contra a coleta, armazenamento, uso e compartilhamento ilimitados de dados pessoais.

Ressaltou, porém, o Tribunal que o direito à "autodeterminação informativa" não é ilimitado. Nos termos do próprio art. 2, I da Lei Fundamental, o direito ao livre desenvolvimento da personalidade sofre limitações decorrentes do choque com os direitos dos outros, com a ordem constitucional ou com leis morais. Firmou-se que o legislador deve respeitar, na regulamentação, o princípio da proporcionalidade.

Como explica Laura Mendes[81]:

> A Corte afirmou que o moderno processamento de dados pessoais configura uma grave ameaça à personalidade do indivíduo, na medida em que possibilita o

[79] MENDES, op. cit., p. 31.

[80] O art. 2, I da Lei Fundamental de Bonn dispõe: "Todos têm o direito ao livre desenvolvimento da sua personalidade, desde que não violem os direitos de outros e não atentem contra a ordem constitucional ou a lei moral." Cf. DEUTSCHER BUNDESTAG (PARLAMENTO FEDERAL ALEMÃO). **Lei Fundamental da República Federal da Alemanha** de 23 de maio de 1949. Berlin, Tradutor: Aachen Assis Mendonça. Disponível em: https://www.btg-bestellservice.de/pdf/80208000.pdf. Acesso em: 02 nov. 2020.

[81] MENDES, op.cit., p.31.

> armazenamento ilimitado de dados, bem como permite a sua combinação de modo a formar um retrato completo da pessoa, sem a sua participação ou conhecimento.

É justamente a coleta, combinação e recombinação de dados que está na base das modernas técnicas de Inteligência Artificial e Aprendizado de Máquina, de modo que a decisão do Tribunal Constitucional antecipou em algumas décadas questões jurídicas que se tornariam agudas no século XXI.

2.2 O direito à proteção de dados pessoais

Nos anos 1990, o surgimento da internet traria problemas novos para o direito à privacidade, a ponto de ser necessária a criação de uma nova terminologia e de uma nova gramática para falar sobre o tema, visto como a expressão "direito à privacidade" já não respondia satisfatoriamente ao fenômeno que começava a se configurar na rede mundial de computadores. Em vez de direito à privacidade, passou-se a falar de um "direito à proteção de dados pessoais".

Para começar, não faz sentido pensar-se na privacidade na internet como algo isolacionista. A internet é uma rede que conecta um vasto número de dispositivos (móveis ou fixos), que são usados por pessoas ou operam automaticamente. Ao entrar na internet, por meio de um Provedor de Acesso (ou Provedor de Conexão, na linguagem da Lei 12.965/2014 – Marco Civil da Internet), o usuário já oferece uma série de dados sobre si mesmo (no mínimo, a data, o local e o horário em que se conectou). Ao pesquisar alguma coisa, ou navegar pelos *sites*, mais informações são oferecidas naturalmente pelo usuário. A internet funciona como uma grande teia global de conversação, entretenimento, negócios, administração, etc., na qual, para obter informação, o usuário tem de oferecer informação (no mínimo, a informação sobre qual informação está a procurar).

Num contexto assim, todos os usuários estão atrelados entre si e a produção de dados pessoais é ampla e transnacional. O fato de essas movimentações serem gravadas e passíveis de recuperação, combinação e rearranjo permite a reconstituição de ações passadas e até a construção de perfis pessoais detalhados. Ou seja, pode-se formar uma imagem do indivíduo a partir dos seus movimentos na internet. Isso pode afetar a privacidade no sentido clássico se, por exemplo, for divulgada, sem causa legítima e sem o consentimento do interessado, alguma informação íntima do usuário para o público, tais como fotos, escritos, locais visitados, etc. Mas, além disso, pode-se também utilizar os dados pessoais para fins de estratégia comercial, política, científica ou cultural, o que nem sempre será legítimo, e pode vir a deturpar o propósito

do titular dos dados pessoais. É nesse ponto especificamente que se fala da necessidade de um direito à proteção de dados pessoais, como algo além da simples proteção da privacidade.

Diferentemente do direito à privacidade, que está baseado na dicotomia esfera pública/esfera privada, funcionando como porta seletiva para separar os dois espaços, o direito à proteção de dados pessoais não está claramente posicionado em uma dessas esferas[82]. Como observa Bruno Bioni, "a dinâmica de proteção de dados pessoais foge à dicotomia do público e do privado, diferenciando-se substancialmente do direito à privacidade"[83].

Embora não se confundam, o direito à privacidade e o direito à proteção de dados têm largo campo de interseção e não se pode deixar de ver que historicamente o segundo está ligado de modo evidente ao primeiro. A circunstância de que o direito à proteção de dados envolva mais elementos e apresente maior complexidade, sobretudo por força das novas tecnologias da informação, não elide o fato de que ele, assim como o direito à privacidade, nasce da ideia de liberdade individual para governar a própria vida conforme as idiossincrasias respectivas, dentro de um campo reservado para o exercício dessa liberdade. Está claro que a privacidade aponta mais para o sentido de segregação, ao circunscrever uma extensão finita de espaço social em que a pessoa pode movimentar-se sem prestar contas ao público; ao passo que o direito à proteção de dados volta-se mais para o governo da coleção finita de dados e informações pessoais dispersas tanto no espaço privado quanto no espaço público, tutelando o processo de formação de conhecimento a partir dessas informações. Mas ambos (privacidade e proteção de dados pessoais) têm raiz na autonomia privada.

Como acentua Danilo Doneda, "mantém-se uma constante referência objetiva a uma disciplina para os dados pessoais, que manteve o nexo de continuidade com a disciplina da privacidade, da qual é uma espécie de herdeira, atualizando-a e impondo características próprias."[84]

As gerações de leis sobre a proteção de dados revelam como se deu a evolução a partir de uma ideia primitiva de privacidade. Na década de 1970, surgiram as primeiras leis sobre a proteção de dados (o *Privacy Act*, de 1974, nos EUA; e Lei do *Land* alemão de Hesse, de 1970). Nessa primeira geração de leis, a preocupação dos legisladores centrava-se na ideia de regular grandes centros estatais de processamento de dados[85]. A técnica de regulação consistia em prever mecanismos preventivos contra o vazamento ou mau uso dessas informações, com foco

---

[82] BIONI, Bruno R. **Proteção de dados pessoais**: a função e o limite do consentimento. 2.ed. Rio de Janeiro: Forense, 2020, p. 94.
[83] Ibidem, p. 95.
[84] DONEDA, op. cit., p. 173.
[85] Ibidem, p. 175.

na administração dos bancos de dados. Tratava-se de uma visão ainda próxima do conceito de privacidade como compartimentação da informação em espaços pouco comunicáveis.

A difusão espacial dos bancos de dados e o seu aumento exponencial tornaram essa disciplina obsoleta, pois ficou muito difícil exercer um controle, à la poder de polícia, com autorizações e fiscalizações *in loco* sobre cada banco de dado, e sobre as atividades que neles se faziam. Assim, no final dos anos 1970 começa a segunda geração de leis de proteção de dados, considerando-se a lei francesa de proteção de dados (*Informatique et Libertès*), de 1978, como o marco inicial dessa fase[86]. O centro de gravidade da proteção deixou de ser o funcionamento dos bancos de dados em si e passou a ser a privacidade dos titulares dos dados. Observou-se que a técnica de fiscalizar o funcionamento dos bancos de dados tornara-se impraticável e, ademais, impertinente, já que o problema não era a existência e o funcionamento dos bancos de dados, mas sim os danos que eles poderiam causar. Passou-se então a imaginar meios de dar ao próprio titular dos dados o direito de governar seus dados. Logo esse modelo mostrou-se também insuficiente: isso porque os cidadãos estavam cada vez mais oferecendo seus dados voluntariamente para terem acesso a serviços de produtos, de modo que dar-lhes esse poder de autocontrole dos dados tornou-se ilusório.

A terceira geração de direitos sobre a proteção de dados, surgida nos anos 1980, manteve o foco na liberdade do titular dos dados, mas agora considerava a sua fragilidade em certos contextos e criava mecanismos para tornar efetiva essa liberdade de autodeterminação informativa. A decisão do Tribunal Constitucional alemão sobre a Lei do Censo, referida anteriormente, é considerada o marco inicial dessa geração da proteção de dados pessoais.

Não obstante, os custos para o exercício individual da autodeterminação informativa eram altos e pressupunham uma consciência cívica muito elevada, de maneira que, ainda buscando real efetividade para as leis de proteção de dados, foram editadas leis de quarta geração, que são as que existem hoje em diferentes países. Nelas, duas características são preponderantes: a) o reconhecimento de que o indivíduo não tem condições práticas de, isoladamente, exercer o controle sobre os seus próprios dados, por várias razões (econômicas, técnicas, informacionais, etc.); b) a criação de autoridades independentes, vinculadas ao estado, com poder para fazer valer as leis de proteção de dados[87].

O direito à proteção de dados fica bem colocado entre os direitos da personalidade e, sem dúvida, vai muito além da privacidade. Isso, no entanto, não induz a necessidade de

[86] Ibidem, p. 177.
[87] Ibidem, p. 179.

negação do direito à privacidade como fonte de inspiração para a criação desse novo direito, como quer Bruno Bioni[88]. Assim como a ideia de propriedade, no século XIX, não respondia mais às necessidades da privacidade, esta agora é que se mostra insuficiente para oferecer respostas à miríade de problemas suscitada pelo processamento eletrônico de grande volume de dados pessoais. Mas há um fio evolutivo entre esses direitos (propriedade-privacidade-proteção de dados), tanto mais porque o pensamento jurídico, para responder a novas demandas, mesmo quando faz giros copernicanos, naturalmente recorre a analogias, a princípios, a generalizações, a métodos heurísticos que não excluem o passado, antes o absorvem por meio de releituras convenientes ao momento de que se trata.

Um aspecto, aliás, que merece toda a atenção é que a garantia do direito à proteção de dados tem sido associada, em algumas de suas aplicações, ao devido processo legal (*due processo of law*), uma antiga solução institucional proveniente da Magna Carta, de 1215, que nasceu ligada ao direito de propriedade, embora tenha evoluído, notadamente nos Estados Unidos, para tornar-se uma garantia muito geral contra toda forma de arbítrio. Dessa maneira, o direito à proteção de dados pessoais, posto tenha configuração própria e se apresente mais frequentemente no ambiente sofisticado das novas tecnologias da informação, guarda uma conexão profunda com as mais antigas ideias de tutela do indivíduo contra a opressão da coletividade (seja do estado ou do mercado).

O direito à privacidade continua a existir e encontra na internet um ambiente propício para a sua violação em massa; porém, além do tipo de transgressão característica da invasão da privacidade — como escândalos provocados pela divulgação de fotos e vídeos íntimos, por exemplo — o uso dos dados pessoais em mecanismos de aprendizado de máquina (*machine learning*) pode ir muito além disso, ao conceber avatares das pessoas para fins comerciais e políticos, ou ao induzir comportamentos de modo sub-reptício. Essa dimensão, por assim dizer pavloviana, do uso dos dados pessoais penetra mesmo nos aspectos neurológicos dos indivíduos e das comunidades humanas, sendo algo portanto de interesse público evidente, tanto em termos de tutela da personalidade individual, como em termos de tutela da própria sociedade política.

O art. 1º da Lei 13.709, de 14 de agosto de 2018, a chamada Lei Geral de Proteção de Dados Pessoais- LGPD, bem demonstra que o seu âmbito de tutela envolve a privacidade, mas vai além dela, ao estatuir: "Esta Lei dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre

[88] BIONI, op. cit., p. 96.

desenvolvimento da personalidade da pessoa natural." De resto, o parágrafo único do referido dispositivo não deixa dúvida sobre o interesse público envolvido na regulamentação da proteção de dados pessoais, ao acrescentar que as normas gerais da LGPD são de interesse nacional e devem ser observadas pela União, Estados, Distrito Federal e Municípios.

O art. 2º da LGPD também traz a proteção à privacidade e a inviolabilidade da intimidade, da honra e da imagem como fundamentos da disciplina de proteção de dados pessoais (art. 2º, I e IV), mas vai adiante e diz que também entre os seus fundamentos: a) a autodeterminação informativa (art. 2º, II); b) a liberdade de expressão, de informação, de comunicação e de opinião (art. 2º, III); c) o desenvolvimento econômico e tecnológico e a inovação (art. 2º, V); d) a livre iniciativa, a livre concorrência e a defesa do consumidor (art. 2º, VI); e) os direitos humanos, o livre desenvolvimento da personalidade, a dignidade e o exercício da cidadania pelas pessoas naturais.

2.2.1 A proteção de dados no ecossistema da internet

2.2.1.1 O dilema do indivíduo nas redes

O direito à proteção de dados, como visto, precede a internet. A LGPD, inclusive, é expressa em dizer que se aplica a todo tipo de dado pessoal, e não apenas àqueles que estão em formato digital[89]. No entanto, é inegável que foi o surgimento da internet e das tecnologias a ela agregadas que impulsionou os movimentos políticos pela regulamentação do direito à proteção de dados como o conhecemos atualmente. Ao reduzir todo tipo de dado (textos, imagens, arquivos de som, registros de transações comerciais, etc.) para um mesmo formato (o digital) e fazê-los convergir para uma rede comum, tornando factível a sua combinação e recombinação, a internet criou um ambiente de geração de comodidades e riquezas a partir dos dados jamais visto antes.

A adaptação a essa nova forma de geração de riqueza, que mistura a vida doméstica com entretenimento, trabalho, negócio, saúde, educação, é um desafio não apenas para os indivíduos, mas também para coletividade. Como observa Klaus Schwab[90], as mudanças tecnológicas atuais afetam o "eu" interior, que corre o risco de escravização por essas tecnologias, assim como levantam questões para as sociedades e para os Estados, que têm de lidar com os prós e

[89] Art. 1º da LGPD: "Art. 1º Esta Lei dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural."
[90] SCHWAB, op.cit., p. 105.

contras dessas inovações, buscando um equilíbrio interno e externo que assegure o progresso sem comprometer a soberania nacional e a dignidade do ser humano.

Em geral, o lado positivo da manipulação de dados pessoais está no oferecimento, para o titular, de inegáveis benefícios de ordem econômica ou de bem-estar, customizados para atenderem às suas necessidades de modo específico, eventualmente até antecipando essas necessidades por meio de modelos preditivos. O lado negativo é que, para obter essas comodidades, o usuário tem de fornecer um número cada vez maior de dados pessoais, alguns relativos até à sua intimidade corporal e psicológica. Qual o preço maior a pagar: a perda da privacidade ou a perda das comodidades que fazem parte do próprio estilo de vida atual?

Na perspectiva do indivíduo, esse é o dilema mais delicado que envolve a proteção de dados. Olhando para as revoluções tecnológicas anteriores, não é difícil prever que esse impasse tende a resolver-se em favor das tecnologias. A resistência pura e simples aos avanços tecnológicos, como o prova o exemplo do ludismo na Inglaterra, não é uma alternativa viável — o que não quer dizer que se deva aceitar passivamente as técnicas de coleta e tratamento de dados propostas pelos tecnólogos.

Klaus Schwab lembra, a esse propósito, do caso dos dispositivos "vestíveis" (*weareable*) para coleta de dados sobre a saúde do usuário, com o intuito de prevenir certas doenças. Ele pondera:

> Será que devemos dar boas-vindas a esse avanço porque ele nos motiva a viver mais saudáveis? Ou ele toma um rumo preocupante a um estilo de vida em que a vigilância — do governo e das empresas — irá tornar-se cada vez mais intrusiva? No momento, esse exemplo refere-se a uma escolha individual — a decisão de aceitar ou não um dispositivo de bem-estar.
> Mas insistindo nisso uma vez mais, vamos supor que agora o empregador peça que todos os seus funcionários usem um dispositivo que envia dados relativos à saúde para a seguradora, porque a empresa quer melhorar a produtividade e, possivelmente, diminuir seus custos com seguros de saúde. E se a empresa exigir que seus funcionários mais relutantes aceitem o pedido ou paguem uma multa? Então, o que anteriormente parecia ser uma escolha consciente individual — usar ou não um dispositivo — passa a ser uma questão de conformidade com as novas normas sociais, mesmo que alguém as considerem inaceitáveis.[91]

Desse dilema decorre outra questão fundamental: admitindo-se que a pressão econômica sobre os dados pessoais é inevitável, não seria mais efetivo simplesmente tratá-los como objeto de propriedade do seu titular, em vez de postular uma nova forma de tutela jurídica?

Laura Mendes[92] menciona que a concepção proprietária dos dados pessoais liga-se aos adeptos da corrente *Law and Economics*. O argumento central dessa linha de pensamento reside

[91] SCHWAB, op.cit., p. 106.
[92] MENDES, op.cit., p. 121.

no seguinte: se os dados pessoais passaram a ter alto valor econômico, naturalmente há incentivos para o seu uso pelo mercado; por outro lado, esse uso de fato traz riscos à privacidade do titular; assim, a privacidade é uma "externalidade negativa". A melhor forma de compatibilizar essa situação é a "internalização" do custo da privacidade, mediante o pagamento, pela empresa que trata os dados, de um valor econômico ao titular desses dados pessoais. Ora, isso nada mais seria do que considerar os dados pessoais como objeto de uma espécie de propriedade imaterial.

Argumenta-se que essa concepção apresenta, pelo menos, três dificuldades[93]. Em primeiro lugar, a adoção dessa tese implicaria que as desigualdades de renda e patrimônio se refletiriam na proteção de dados, pois as pessoas sem poder de barganha estariam sempre dispostas a abrir mão de seus dados pessoais, em favor de algum ganho econômico. Em segundo lugar, as vantagens econômicas de certo perfil de dados, que real ou supostamente tivessem mais valor para as empresas de tratamento de dados, fariam com que as pessoas buscassem adotar um comportamento nas redes, ou mesmo na vida *off-line*, de modo a fazer com o que o seu perfil tivesse um *upgrade* nos mecanismos de classificação de dados. Finalmente, a perda de controle sobre as próprias informações macularia a individualidade do cidadão, tornando-o inapto para vivificar a democracia e os processos de decisão coletiva, visto que não seria mais ele o titular da sua própria individualidade.

O certo é que, como acentuam Sacha Romanosky e Alessandro Acquisti[94], é muito difícil encontrar uma solução puramente econômica para o problema da proteção dos dados na atual conjuntura. Primeiro porque os danos causados nas redes são probabilísticos: podem ser catastróficos para uns e irrelevantes para outros. Depois, eles frequentemente se manifestam de modo indireto, como, por exemplo, na utilização de dados pessoais para formar o perfil de toda uma classe de pessoas, e não especificamente de alguém. Outro ponto é que, em certas circunstâncias, embora não haja dano econômico, pode haver dano psicológico para o titular dos dados.

É preciso admitir também que não raras vezes faltam informações suficientes sobre a causa, a gravidade e o volume das violações aos dados pessoais, de tal maneira que praticamente qualquer abordagem política atual mostra-se insuficiente — e eventualmente até perniciosa —

---

[93] MENDES, op. cit., p. 122-123.

[94] ROMANOSKY, Sacha; ACQUISTI, Alessandro. Privacy costs and persona data protection: economic and legal perspectives. **Berkeley Technology Law Journal**, Sacramento, v. 24, n. 3, p. 1063-1102, 2009. Disponível em: https://www.heinz.cmu.edu/~acquisti/papers/RomanoskyAcquisti-INFORMS-2009.pdf. Acesso em: 04 set. 2020., p. 1099.

para prevenir e reprimir as práticas de violação de dados[95]. Nesse contexto, o sistema político ainda está numa fase de experimentação de soluções para o problema, não sendo boa a prática de concentrar toda a solução no aspecto puramente econômico.

2.2.1.2 A dimensão política da proteção de dados pessoais

A internet tornou-se um ecossistema, um *locus* onde é mais relevante do que nunca o direito à proteção de dados pessoais, justamente porque ali os dados são a matéria-prima da economia dita de quarta geração ou "Indústria 4.0"[96]. O ambiente da internet apresenta características peculiares, que trazem alguns desafios inéditos para o processo de aplicação das normas jurídicas em geral e, em particular, das normas de proteção de dados e de privacidade em sentido mais estrito.

Em primeiro lugar, os atores estatais mais relevantes da internet estão ainda num jogo de acomodação de suas posições de poder em âmbito internacional, o que traz repercussões sobre o alcance e os meios práticos para proteger dados pessoais em todos os lugares do planeta. De fato, é notório que Estados Unidos e China (e, em segundo plano, a Rússia), os dois protagonistas dos avanços na produção de conhecimento por meio de dados e na infraestrutura da internet, estão numa luta encarniçada para impor os seus respectivos modelos de governança da internet em escala global, a tal ponto que já se fala de uma nova Guerra Fria. A União Europeia, por seu turno, tem se adiantado em relação a outras partes do mundo no tema da legislação protetiva de dados.

Um ponto central na disputa EUA-China reside no desenvolvimento da Inteligência Artificial, que está na fronteira de toda a evolução por que passam as novas tecnologias. Vladimir Putin, o presidente da Rússia, proferiu uma frase a esse respeito que se tornou célebre: "quem liderar a esfera da Inteligência Artificial ditará regras ao mundo"[97]. O perturbador dessa afirmação é que existem aplicações militares da Inteligência Artificial que podem, de fato,

---

[95] Ibidem, p. 1100-1101.

[96] Klaus Schwab (2016, p. 16) explica que a expressão "Indústria 4.0" teria sido cunhado em 2011 na Feira de Hanover, para descrever o modo de produção atual, que funde os meios físicos, digitais e biológicos para criar riquezas. A alusão à quarta ordem, se dá porque se considera que a Humanidade já teria passado por três etapas anteriores de industrialização: a primeira, entre 1760 e 1840, decorreu das ferrovias e da invenção da máquina a vapor; a segunda, no final do século XIX, veio com o domínio da eletricidade; a terceira, a partir dos anos 1960, decorreu do uso de computadores para processar dados; a quarta, atualmente, seria fruto das novas tecnologias digitais e da internet.

[97] VINCENT, James. Putin says the nation that leads in AI 'will be the ruler of the world'. *In*: VOXMEDIA. **The Verge**. 4 set. 2020. Disponível em: https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world. Acesso em: 23 set. 2020.

alterar drasticamente o equilíbrio de poder entre os países, e até provocar um conflito bélico de grandes proporções[98].

Como acentua Miachael C. Horowitz[99], a questão não está somente nas tecnologias em si, mas sobretudo nos usos que cada país fará delas. E o uso é diretamente influenciado, de um lado, pelos incentivos públicos, e, de outro, pelas regulamentações legais e exigências burocráticas que cada país faz para pesquisa e aplicações das tecnologias. Logo, o problema da regulamentação e da fiscalização estatal sobre o uso de dados pessoais para o desenvolvimento da Inteligência Artificial e tecnologias associadas é decisivo não apenas para a preservação de direitos individuais, mas igualmente para o tipo de estratégia política global que cada país traça para o seu desenvolvimento tecnológico.

Nesse ponto, tem se tornado comum a visão segundo a qual as leis de proteção de dados pessoais rigorosas tendem a frear os avanços da Inteligência Artificial e, consequentemente, a ameaçar a liderança dos países do Ocidente, quando comparados à China, por exemplo, que aparentemente não vê a proteção de dados pessoais como um problema relevante a ser resolvido no momento. Kai Fu Lee, referindo-se à coleta de dados pessoais em locais públicos na China, para aplicações em Inteligência Artificial, assim expressa a questão:

> (...)Esse tipo de coleta de dados pode não ser bem-visto para muitos norte-americanos. Eles não querem que o Big Brother ou a América corporativa saibam tanto sobre o que estão fazendo. Mas o povo chinês aceita que seu rosto, sua voz e escolhas de compra sejam capturados e digitalizados [...] As cidades chinesas já usam uma densa rede de câmeras e sensores para reforçar as leis de trânsito. Essa teia de imagens de vigilância agora está alimentando diretamente algoritmos de otimização para gerenciamento de tráfego, policiamento e serviços de emergência.
> Não há resposta certa para questões sobre que nível de vigilância social é um preço que vale a pena pagar para maior conveniência e segurança, ou que nível de anonimato devemos garantir em aeroportos ou estações de metrô. Mas, em termos de impacto imediato, a relativa abertura da China para a coleta de dados em locais públicos significa uma grande vantagem inicial na implementação da IA de percepção.[100]

Se os dados são o combustível que move a Inteligência Artificial, está claro que a proteção de dados coloca-se no epicentro de uma batalha não apenas por domínio econômico, mas sim por imposição política, visto como a Inteligência Artificial alimentada por grandes volumes de dados parece ser o núcleo de toda a revolução que se desenrola atualmente. Kai Fu Lee, com

---

[98] HOROWITZ, Michael C. Artificial Intelligence, International Competition, and the Balance of Power. **Texas National Security Review**, Austin, v. 1, n. 3, p. 37-57, maio 2018.
[99] Ibidem, p. 43.
[100] LEE, Kai-Fu. **Inteligência artificial**: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos. Rio de Janeiro: Globo Livros, 2019. Tradução de Marcelo Barbão, p. 152.

certo ufanismo, aponta que a corrida entre Estados e Unidos e China pela hegemonia sobre os insumos da Inteligência Artificial vai muito além dos aspectos meramente econômicos:

> (...) o Vale do Silício continua sendo o líder claro no desenvolvimento de chip de IA. Mas é uma liderança que o governo chinês e a comunidade de capital de risco do país estão tentando eliminar ao máximo. Isso porque quando a ruptura econômica ocorrer na escala prometida pela inteligência artificial, não será apenas uma questão de negócios — também será uma questão política importante.[101]

Constata-se, assim, que a proteção de dados pessoais no ecossistema da internet coloca-se numa esfera muito mais ampla do que aquela do direito à privacidade. Se, por um lado, assim como na proteção da privacidade, há relevância para o indivíduo em ter os seus dados pessoais protegidos, por razões eminentemente particulares; por outro, a circunstância de que esses dados pessoais, quando entrecruzados com muitos outros, pode ser uma fonte inesgotável de conhecimento para diferentes aplicações (comerciais, políticas, científicas, culturais, etc.), faz com que altos interesses políticos sejam despertados sobre o assunto, tanto para a governança interna de cada país como para fins de geopolítica internacional.

O exemplo da legislação brasileira é bem ilustrativo sobre as implicações políticas e administrativas da proteção de dados pessoais, para além da privacidade individual. Antes de tudo, convém observar que o Brasil, ao aprovar uma Lei Geral de Proteção de Dados fortemente inspirada no modelo europeu, já ensaia um alinhamento ideológico com aquele bloco comunitário, o que, no entanto, apenas se confirmará na prática da política internacional quando testes mais sérios se apresentarem. A aprovação da lei, no contexto internacional, funciona mais como indício simbólico sobre as intenções políticas do país, não sendo de fato ainda uma prova de completa adesão ao modelo europeu, tanto mais porque tal modelo, entre nós, não foi ainda colocado à prova.

Estão presentes na legislação de proteção de dados brasileira, como se verá adiante, tanto elementos de política externa, como de governança interna, notadamente por meio de regulação e fiscalização. Evidentemente, se o direito à proteção de dados fosse circunscrito à privacidade individual, as repercussões internas e externas desse direito seriam infinitamente menores.

Para começar, a Lei Geral de Proteção de Dados - LGPD, ao arrolar os fundamentos da disciplina de proteção de dados pessoais, no seu art. 2º, omite o tema da soberania nacional[102],

---

[101] Ibidem, p.120.

[102] Art. 2º A disciplina da proteção de dados pessoais tem como fundamentos:
I - o respeito à privacidade;
II - a autodeterminação informativa;
III - a liberdade de expressão, de informação, de comunicação e de opinião;

mencionando apenas valores ligados ao indivíduo e ao mercado. Porém, em outros pontos a LGPD revela que aspectos de política externa estão dentro do seu raio de ação.

De fato, o âmbito territorial e material de aplicação reclamado pela lei expõe o grau da pretensão da soberania nacional sobre os dados pessoais que são produzidos ou de algum modo tratados no país. No art. 3º, I, II e III, a LGPD ressalta que, independentemente do meio, do país de sua sede ou do país onde estejam localizados os dados, aplica-se a LGPD brasileira: a) à operação de tratamento seja realizada no território nacional; b) à atividade de tratamento tenha por objetivo a oferta ou o fornecimento de bens ou serviços ou o tratamento de dados de indivíduos localizados no território nacional; c) aos dados pessoais objeto do tratamento tenham sido coletados no território nacional.

Vê-se que é uma reivindicação de autoridade normativa bastante extensa, que dependerá, para ser efetiva, de cooperação internacional e quiçá de estrutura tecnológica apropriada para desvendar eventuais violações de dados pessoais nas circunstâncias descritas na lei, bem como para implementar as soluções encontradas.

Internamente, arvorando-se uma prerrogativa maior do que a dos atores privados, o Poder Público goza de posição mais confortável na LGPD. Com efeito, o art. 4º, III aponta para situações ligadas ao Poder Público que são excluídas da disciplina geral da proteção de dados, tais como o tratamento de dados para fins de: segurança pública; defesa nacional; e segurança do Estado ou atividades de investigação e repressão de infrações penais. Também estão excluídas aquelas situações em que o tratamento de dados pessoais implique apenas uma "passagem inocente" em território nacional (para usar expressão cara ao Direito do Mar), sem qualquer repercussão interna ou compartilhamento com agentes de tratamento brasileiros (LGPD, art. 4º, IV).

A questão da necessidade de consentimento do titular para que haja o tratamento de dados também mostra nuances políticas interessantes, agora em âmbito doméstico. É assim que, em geral, para o tratamento de dados pessoais, sensíveis ou não, há a necessidade do consentimento do titular (LGPD, arts. 7º, I e 11, I). Entretanto, se o tratamento for realizado pelo Poder Público com o objetivo de executar alguma política pública, então o consentimento não se mostra necessário (LGPD, arts. 23 a 30). Aqui desponta o velho princípio da supremacia do interesse público sobre o privado, tão caro aos publicistas.

IV - a inviolabilidade da intimidade, da honra e da imagem;
V - o desenvolvimento econômico e tecnológico e a inovação;
VI - a livre iniciativa, a livre concorrência e a defesa do consumidor; e
VII - os direitos humanos, o livre desenvolvimento da personalidade, a dignidade e o exercício da cidadania pelas pessoas naturais.

Voltando à política externa, merece destaque a questão da disciplina da Transferência Internacional de Dados (LGPD, arts. 33 a 36); ela ilustra as preocupações do estado brasileiro quanto à circulação de dados pessoais coletados ou tratados no país e mostra que, embora os dados pessoais sejam suscetíveis de circulação econômica, o direito à sua proteção é irrenunciável e intransmissível, como todo direito da personalidade (Código Civil, art. 11). De resto, a própria disciplina da proteção de dados em si confere ao país um *status* internacional que pode lhe franquear o acesso a mercados econômicos de outros países com igual nível de desenvolvimento jurídico[103].

Segundo a LGPD, a Transferência Internacional de Dados do Brasil para o exterior é possível, desde que se observem algumas condições (LGPD, art. 33 c/c art. 7º, II, V, e VI):

a) Se a transferência for para países ou organismos internacionais que propiciem grau de proteção aos dados igual ou superior ao do Brasil, cabendo à Autoridade Nacional, se for o caso, avaliar o grau de proteção do país de destino;
b) Quando o controlador[104] oferecer e comprovar garantias de cumprimento dos princípios, dos direitos do titular e do regime de proteção de dados previstos na LGPD;
c) Em caso de cooperação jurídica internacional entre órgãos públicos de investigação ou persecução penal;
d) Quanto a transferência for necessária para proteger o direito à vida ou a incolumidade física do titular ou de terceiro;
e) Quando a Autoridade Nacional autorizar a transferência;
f) Se a transferência estiver fundada em acordo internacional de cooperação;
g) Quando a transferência se der para cumprir alguma política pública;
h) Se o titular dos dados tiver dado o seu consentimento específico e em destaque para a transferência, com informação prévia sobre o caráter internacional da operação, distinguindo claramente esta de outras finalidades;
i) Quando necessário para fins de cumprimento de obrigação legal ou regulatória pelo controlador;
j) Quando necessário para a execução de contrato ou de procedimentos preliminares relacionados a contrato do qual seja parte o titular, a pedido do titular dos dados;

---

[103] TABACH, Danielle; LINHARES, Ludmila Anaquim. Transferência Internacional de dados. In: FEIGELSON, Bruno; SIQUEIRA, Antonio Henrique Albani. **Comentários À Lei Geral De Proteção De Dados**. São Paulo: Thomson Reuters Brasil Revista dos Tribunais, 2019. Cap. 5, p.149.

[104] Controlador, na linguagem da LGPD, é pessoa natural ou jurídica, de direito público ou privado, a quem competem as decisões referentes ao tratamento de dados pessoais. Cf. art. 5º, VI da Lei 13.709/2019.

k) para o exercício regular de direitos em processo judicial, administrativo ou arbitral.

As condições para transferência apoiam-se em quatro fundamentos: a) consentimento do titular; c) política pública nacional; c) cooperação internacional; d) legítimo interesse do controlador, desde que a transferência se dê para entidade estrangeira com nível de proteção igual ou superior ao do Brasil.

Esse tema exemplifica com clareza que o problema da proteção de dados pessoais é complexo e tem diferentes camadas de tutela jurídica. Por um lado, como direito da personalidade que é, o direito à proteção de dados não é transmissível, nem disponível, conquanto se admita que, temporariamente, pode-se licenciar o acesso a uma parte dos dados pessoais para fins econômicos[105]. Por outro, como um direito fundamental que também é, a tutela do direito à proteção de dados interessa à República Federativa do Brasil no âmbito das suas relações internacionais[106].

Como ativo econômico, o acesso aos dados pessoais pode ser flexibilizado, pelo próprio titular, mediante o seu consentimento, para fins de obtenção de alguma vantagem negocial. Como direito fundamental fortemente ligado a operações transnacionais, o direito à proteção de dados está subordinado a interesses políticos superiores do país e aos influxos das relações internacionais, nas suas mais diferentes expressões.

Um caso recente, ocorrido na Europa, torna manifesta a tensão política que há em torno do direito à proteção de dados na relação entre as soberanias. Referimo-nos ao caso da transferência de dados de usuários europeus do Facebook para arquivos nos Estados Unidos[107]. A Comissão Irlandesa de Proteção de Dados emitiu uma ordem preliminar para que o Facebook adote providências para impedir a transferência de dados de usuários europeus para os Estados Unidos. A ordem — cuja implementação não é tecnicamente simples, visto que o modelo de negócios do Facebook e das *Big Techs* em geral, envolve a livre circulação de dados — não deverá ser executada de imediato e muito provavelmente ainda sofrerá questionamentos perante as instâncias decisórias europeias.

A decisão da Comissão Irlandesa está baseada num julgamento da Corte de Justiça da União Europeia, conhecido como Schrems II (em referência ao ativista da privacidade

---

[105] BITTAR, Carlos Alberto. **Os direitos da personalidade**. 8. ed. São Paulo: Saraiva, 2015. *Ebook*, posição 547.

[106] MENDES, op.cit., p. 172.

[107] SATARIANO, Adam. Facebook May Be Ordered to Change Data Practices in Europe. *In*: NYTCO. **The New York Times.** 9 set. 2020. Disponível em: https://www.nytimes.com/2020/09/09/technology/facebook-european-union-data-privacy.html. Acesso em: 24 set. 2020.

Maximiliam Schrems, cuja queixa deu origem ao precedente)[108]. O julgado diz respeito ao choque de dois regimes jurídicos muito diferentes relacionados com os dados digitais das pessoas: por um lado, a lei de vigilância dos EUA e, por outro lado, a proteção de dados e a privacidade na União Europeia, notadamente com base no *General Data Protection Regulation* – GDPR.

Enquanto os EUA, por diversas razões, dão prioridade à vigilância digital — como revelado por Edward Snowden em 2013[109] —, conferindo poderes a autoridades estatais para capturar dados pessoais, conforme a Seção 702 do FISA (*Foreign Intelligence Surveillance Act*)[110] e a Ordem Executiva 12.333[111] (que permite a coleta de grandes volumes dados pessoais), a União Europeia, por outro lado, opta por dar maior proteção aos dados pessoais. Em tal contexto, a Corte de Justiça Europeia constatou, no julgamento do caso Schrems II[112], que o fluxo transatlântico de dados pessoais operacionalizado pelo Facebook, colide diretamente com os direitos fundamentais europeus que conferem aos cidadãos direitos à privacidade e à proteção de dados, tal como estabelecido na Carta dos Direitos Fundamentais da UE, na Convenção Europeia dos Direitos do Homem e em peças específicas da legislação europeia (especialmente o Regulamento Geral de Proteção de Dados). Isso porque, em solo americano, os dados pessoais não têm o mesmo nível de proteção oferecido pela União Europeia.

O caso Schrems II, embora envolva diretamente apenas o *Facebook*, tem implicações muito mais amplas sobre a forma como o processamento de dados em grande escala dos dados dos cidadãos da UE pode ser feito, visto como o tribunal emitiu interpretação abstrata de vários dispositivos da normativa europeia sobre a transferência internacional de dados. Isso prefigura o trabalho das instâncias decisórias mundo afora: dificilmente um tema de proteção de dados pode ser decidido apenas para o caso concreto; pela própria forma de operação das tecnologias,

---

[108] O mesmo Maximiliam Schrems houvera questionado perante a Corte de Justiça da União Europeia, em 2013, a transferência de seus dados pessoais pelo Facebook, da União Europeia para os Estados Unidos, com base no acordo bilateral (EUA e UE) chamado de *Safe Harbour*, no que posteriormente ficou conhecido como o caso Schrems I. Cf. DONEDA, op. cit., p. 255-256.

[109] BAMFORD, James. The Most Wanted Man in the World. *In*: CONDÉ NAST. **Wired**, 2014. Disponível em: https://www.wired.com/2014/08/edward-snowden/. Acesso em: 29 set. 2020.

[110] FOREIGN INTELLIGENCE SURVEILLANCE ACT. **Section 702 Overview**, 2008. Disponível em: https://www.dni.gov/files/icotr/Section702-Basics-Infographic.pdf. Acesso em: 29 set. 2020.

[111] THE U.S. NATIONAL ARCHIVES AND RECORDS ADMINISTRATION. **Executive Order 12333:** United States intelligence activities, 1981. Disponível em: https://www.archives.gov/federal-register/codification/executive-order/12333.html. Acesso em: 29 set. 2020.

[112] GRAND CHAMBER. Judgment of the court in Case C-311/18. *In*: CURIA. **O TRIBUNAL DE JUSTIÇA DA UNIÃO EUROPEIA**. [*S. l.*], 16 jul. 2020. Disponível em: http://curia.europa.eu/juris/document/document.jsf?text=&docid=228677&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=4956180. Acesso em: 29 set. 2020.

baseadas em algoritmos, decisões de casos singulares acabam tendo repercussão sobre todo o modelo da operação.

Por fim, vale a pena notar que a decisão da Corte Europeia não diz respeito às chamadas transferências de dados "necessárias", como, por exemplo, o envio de um *e-mail* para reservar um quarto de hotel. O que de fato foi objeto da decisão foram as transferências de dados em massa, que estão associadas a certos modelos de negócio na internet, em razão da redução de custos que esse tipo de transferência implica.

2.3 Natureza complexa do direito à proteção de dados: direito da personalidade, direito fundamental e direito básico do consumidor

A proteção de dados, como visto, provém historicamente da tutela da privacidade, mas é muito mais ampla do que ela. À medida que tecnologias de captação e registro de textos, imagens, sons e fatos sociais em geral foram se desenvolvendo, a personalidade foi projetando um espectro correspondente em dados, dispersos nos mais diferentes meios de registro (cartas, fotografias, filmagens, etc.). O advento dos computadores e da internet não apenas aumentou de maneira exponencial o número de registros das manifestações da personalidade, como também permitiu o entrecruzamento dos dados respectivos, recopilando-os para sintetizar ícones do indivíduo nas mais diferentes situações sociais. A tal ponto chegou esse processo quase demiúrgico, que estão se tornando borradas as fronteiras entre o *on-line* e o *off-line*, num processo de hipostasia eletrônica da vida.

A tutela jurídica dos dados pessoais, nesse contexto disperso e movediço, não pode por isso mesmo ser monolítica. É assim que o direito à proteção de dados pode ser encarado tanto como um direito da personalidade, quando se cogita de proteger aspectos da imagem e da vida privada de um indivíduo; como um direito fundamental, quando se trata sobretudo de salvaguardar, em contextos mais amplos, a liberdade individual ou coletiva e a cidadania; e como um direito básico do consumidor, na hipótese de relações de consumo, em que preponderam discussões sobre assimetrias econômicas e técnicas, consentimento e regulação.

O que une essas diferentes abordagens e está na raiz do direito à proteção de dados é o livre desenvolvimento da personalidade humana e a autodeterminação informativa. Sempre condutas de pessoas, empresas ou estados, direta ou indiretamente (por meio de técnicas automatizadas de manipulação de dados), sejam fonte de objetificação do ser humano, desprezando a sua dignidade, nas relações entre indivíduos, ou entre indivíduos e empresas, ou entre indivíduos e estados, o direito à proteção de dados terá a função de obstruir semelhantes condutas e restaurar o espaço de respeito à liberdade humana.

O art. 2º da LGPD bem expressa essa polivalência do direito à proteção de dados, ao enumerar os seus fundamentos. Ali se fala desde o "respeito à privacidade", no inciso I, até os direitos humanos, o livre desenvolvimento da personalidade, a dignidade e o exercício da cidadania pelas pessoas naturais, no inciso VII, passando pela a livre iniciativa, a livre concorrência e a defesa do consumidor, no inciso VI.

É assim que não há tanta importância em fixar um conteúdo exato para o direito à proteção de dados, tanto mais porque ele precisa ser flexível para abarcar o universo de questões novas que se apresentam com a evolução da técnica informática, à qual a proteção de dados precisa ser aderente para ser efetiva. A propósito disso, observa Danilo Doneda:

> O reconhecimento de um direito fundamental à proteção dos dados pessoais vem acompanhado por uma reflexão sobre sua eficácia. A legislação sobre a matéria, desde as primeiras manifestações, dedicou especial atenção à forma de atuação da proteção de dados, o que determinou a adaptação de criação de instrumentos para sua tutela. Essa não chega a ser uma característica particular da proteção de dados pessoais: Nicolò Lipari notou que nos novos direitos frequentemente ocorre uma superação dos modos de determinação do conteúdo, privilegiando o estabelecimento das técnicas de sua tutela.[113]

A falta de entendimento exato das causas ou da extensão de um problema frequentemente faz com que as políticas de abordagem normativa concentrem-se nas consequências, mediante o uso de conceitos jurídicos indeterminados capazes de apanhar fenômenos não inteiramente compreendidos. Isso está ocorrendo com o direito à proteção de dados. Para confirmar essa observação, basta verificar, por exemplo, a disciplina da "anonimização" de dados na LGPD. Primeiro a lei diz que a anonimização é a "utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo" (LGPD, art. 5º, XI). Em seguida, tentando esclarecer o que seria considerado "razoável", a LGPD dispõe: "A determinação do que seja razoável deve levar em consideração fatores objetivos, tais como custo e tempo necessários para reverter o processo de anonimização, de acordo com as tecnologias disponíveis, e a utilização exclusiva de meios próprios" (LGPD, art. 12, §1º).

Outro exemplo flagrante dessa técnica de abordagem está no art. 44 da LGPD. Ao definir o que seria um "tratamento irregular de dados pessoais", a lei afirma:

> Art. 44. O tratamento de dados pessoais será irregular quando deixar de observar a legislação ou quando não fornecer a segurança que o titular dele pode esperar, consideradas as circunstâncias relevantes, entre as quais:
> I - o modo pelo qual é realizado;
> II - o resultado e os riscos que razoavelmente dele se esperam;

[113] DONEDA, op. cit., p. 289.

> III - as técnicas de tratamento de dados pessoais disponíveis à época em que foi realizado.
> Parágrafo único. Responde pelos danos decorrentes da violação da segurança dos dados o controlador ou o operador que, ao deixar de adotar as medidas de segurança previstas no art. 46 desta Lei, der causa ao dano.

Vê-se que a lei não estipula detalhadamente a conduta do que seria um "tratamento irregular", mas apenas esboça uma cláusula genérica de proteção, com o uso de conceitos jurídicos indeterminados, que podem ser devidamente ajustados para técnicas novas que venham a ser desenvolvidas.

Essa política legislativa focada nos riscos e danos tem também o objetivo de redistribuir os ônus econômicos e sociais da absorção de novas tecnologias pela sociedade. Aqueles que se beneficiam economicamente do tratamento de dados pessoais em massa devem suportar os maiores ônus de eventual dano ou do mero risco provocado aos titulares dos dados. O foco normativo deixa de centrar-se na conduta em si do tratamento de dados e volta-se para possíveis prejuízos que ela é capaz de causar. Não por acaso, o parágrafo único do art. 44 estipula uma modalidade de responsabilidade objetiva em desfavor do agente de tratamento de dados (controlador ou operador) que deixar de adotar as medidas de segurança, com isso causando danos ao usuário. Nesse sentido, observa com propriedade Marion Albers[114]:

> Em contraposição aos conceitos originais de proteção de dados, de fato não é possível prever com facilidade o tratamento de dados e informações pessoais, o conhecimento gerado a partir deles e as decisões daí resultantes. A ideia de que esses processos pudessem ser quase completamente previstos, planejados e controlados por meios jurídicos mostrou ser demasiado simples. O processamento de dados e informações, a geração de informação e conhecimento, a tomada de decisões com base em informação e conhecimento incluem certa dinâmica e incerteza em muitos pontos. Isso se aplica com mais razão ainda com vistas ao uso de tecnologias. Consequentemente, é menos a ideia de controle que caracteriza ou deveria caracterizar o direito referente à proteção de dados do que, de modo semelhante ao direito ambiental, a ideia de regulamentação dos riscos.

2.3.1 Reconhecimento do direito fundamental à proteção de dados pelo STF

O Supremo Tribunal Federal proferiu julgamento recente sobre o direito à proteção de dados, ao julgar a Medida Cautelar, deferida pela Ministra Rosa Weber, nos autos das Ações Diretas de Inconstitucionalidade nºs 6.387, 6.388, 6.389, 6.390 e 6.393. Sobre as razões dos votos, é possível dizer que este julgamento representa, para o Brasil, aquilo que caso do Censo de 1983 representou para a Alemanha[115].

---

[114] ALBERS, Marion. A complexidade da proteção de dados. **Direitos Fundamentais & Justiça**, Belo Horizonte, a. 10, n. 35, jul./dez. 2016, p. 43.

[115] MENDES, Laura Schertel; FONSECA, Gabriel Campos Soares da. STF reconhece direito fundamental à proteção de dados: comentários sobre o referendo da Medida Cautelar nas ADIs 6387, 6388, 6389, 6390 e 6393. **Revista de Direito do Consumidor**, v. 130/2020, Jul./Ago, 2020, p.472

Coincidentemente, o julgamento do Supremo Tribunal Federal também diz respeito a um caso de Censo Demográfico. A Medida Provisória nº 954, de 17 de abril de 2020, editada em razão da pandemia de COVID-19, estabeleceu a possibilidade de compartilhamento de dados de todos os usuários (nomes, números de telefone e endereços), pelas companhias de telefonia fixa ou móvel, para fins de realização de censo demográfico, sem a necessidade de entrevista presencial (que aumentaria os riscos de contágio da doença COVID-19). Diversos partidos políticos, bem como a Ordem dos Advogados do Brasil, propuseram ações diretas de inconstitucionalidade contra a referida Medida Provisória, cujos números são aqueles já referidos acima.

Tendo sido deferida a medida cautelar monocraticamente, pela Ministra Rosa Weber, o processo foi ao Plenário do Tribunal para o referendo ou não da citada medida, ocasião em que o colegiado teve oportunidade de se manifestar sobre a questão da existência ou não do direito à proteção de dados no Brasil.

Em jogo estavam duas correntes principais: a) a que defendia a constitucionalidade da medida, dada a excepcionalidade da situação provocada pela crise sanitária; b) a que defendia a proteção dos dados pessoais dos cidadãos, mesmo no contexto de pandemia. O Supremo Tribunal Federal optou por privilegiar essa segunda corrente e, assim, consagrou de modo expresso o direito à proteção de dados como direito fundamental decorrente do texto constitucional brasileiro.

A Ministra Rosa Weber, na decisão monocrática que suspendeu os efeitos da Medida Provisória nº 954, lançou o núcleo da argumentação em favor da proteção de dados e contra toda a instrumentalização do indivíduo em face de políticas públicas de coleta sistemática de informações, nos seguintes termos:

> Não se subestima a gravidade do cenário de urgência decorrente da crise sanitária, nem a necessidade de formulação de políticas públicas que demandam dados específicos para o desenho dos diversos quadros de enfrentamento. O seu combate, todavia, não pode legitimar o atropelo de garantias fundamentais consagradas na Constituição.[116]

Tendo essa decisão sido homologada pelo Plenário do STF, por 10 votos favoráveis e apenas 1 contrário, e considerando os motivos do julgado, está claro que ficou reconhecido o caráter fundamental do direito à proteção de dados no Brasil, para muito além do direito à privacidade ou intimidade. Os dados cuja coleta em massa foi proibida pelo STF não revelam

[116] BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº 6.387. Relator: Ministra Rosa Weber. Brasília, DF, 24 de abril de 2020. **Diário de Justiça Eletrônico**. Brasília, 07 maio 2020, p. 12.

alto grau da intimidade do cidadão, sobretudo quando se sabe que seriam entregues apenas a uma instituição pública respeitável, como é o Instituto Brasileiro de Geografia e Estatística - IBGE, com especifica finalidade de permitir a realização do censo demográfico sem a necessidade de entrevistas presenciais, que poderiam contribuir para espraiar o novo coronavírus (Sars-Cov19). Mas a Corte considerou que quando esses dados são coletados em grande quantidade, estando sujeitos a processos automatizados de tratamento, são altas as probabilidades de se perder o controle do tipo de conhecimento que se consegue produzir através deles. Assim como no julgamento do Tribunal Constitucional alemão, de 1983, reconheceu-se que o problema não está exatamente nos dados em si, mas sim no seu agrupamento e entrecruzamento para produzir informações e conhecimentos.

2.3.2 Proteção de dados e direitos do consumidor

O direito à proteção de dados relaciona-se também aos direitos básicos do consumidor. Particularmente em relação aos provedores de conteúdo ou provedores de aplicações, o usuário consome os bens e serviços por meio das aplicações de internet, ao tempo em que eventualmente "paga" por eles apenas mediante a entrega de seus dados pessoais. Esses dados, por sua vez, quando associados aos de outros usuários por meio de mecanismos de *machine learning*, produzem conhecimento sobre a cadeia de consumo e influencia o processo de produção de bens e serviços, num ciclo infinito de recíproca estimulação. A alavancagem econômica provocada pelo uso dos dados pessoais gera riqueza na medida em que permite aos fornecedores compreenderem muito especificamente as necessidades de cada pessoa e, ao mesmo tempo, o comportamento de grupos de consumidores, de tal maneira que podem criar aquilo que se chama de "sistema flexível de produção"[117], com muito mais eficácia e menos desperdício. Podem os produtores também induzir o consumo, explorando desejos e aspirações que se tornam conhecidos pelo acesso aos dados pessoais.

Logo, os dados pessoais, especialmente quando digitalizados, tornaram-se um ativo de alto valor para o mundo dos negócios pelo predicado que têm de revelar nichos de consumo e oportunidades de negócio inusitadas, cujo volume pode crescer com rapidez e leveza ao mesmo tempo. As chamadas *startups*, graças a modelos de negócio repetíveis e escaláveis, embora não se restrinjam à internet, encontram nela certamente o ambiente propício para crescer. Ora, não é possível cogitar-se de uma *startup* sem o uso massivo de dados pessoais para explorar e compreender o comportamento e as necessidades dos clientes em situações individuais e no

[117] BIONI, op. cit., p. 13.

contexto coletivo. Um bom exemplo são as chamadas "plataformas digitais". Esse modelo de negócio apresenta aptidão para gerar riqueza nas mais diferentes formas de relacionamento social, do namoro à compra de bens e produtos, passando por aluguel de casas, carros, táxis, etc. E no seu núcleo está o tratamento de dados pessoais, a tal ponto que as plataformas estão transfigurando a venda de produtos em serviços, dado o vínculo constante que mantêm com o usuário por meios dos seus dados, mesmo depois da entrega de um bem adquirido. É o fenômeno que tem sido chamado de "servitização". Klaus Scwab[118] explica assim o seu funcionamento:

> As estratégias das plataformas, combinadas com a necessidade de concentrar-se mais no cliente e melhorar os produtos por meio de dados, estão alterando o foco de muitas indústrias, da venda de produtos para o fornecimento de serviços. Um número crescente de consumidores não mais compra e possui objetos físicos, mas preferem pagar pela entrega de um serviço subjacente que será acessado através de uma plataforma digital. É possível, por exemplo, obter acesso digital a bilhões de livros por meio da Kindle Store da Amazon, ouvir quase todas as músicas do mundo pelo Spotfy, ou juntar-se a uma empresa de compartilhamento de carros que fornece serviços de mobilidade sem a necessidade de possuir o veículo. Essa mudança é poderosa e permite o aparecimento de modelos econômicos mais transparentes e sustentáveis de troca de valores.

Em alguns casos, a plataforma não vende nada ao consumidor diretamente, mas lhe entrega "gratuitamente" algumas comodidades, em troca da obtenção de seus dados pessoais. É o caso das redes sociais. O *Facebook* ou o *Instagram*, por exemplo, não cobram nenhum tipo de assinatura do usuário e, não obstante, oferecem-lhe uma série de serviços. O preço a pagar, curiosamente, é usar esses serviços, pois isso gerará naturalmente dados pessoais que, devidamente tratados por mecanismos inteligentes, produzirão conhecimento sobre os usuários, com alto valor econômico. Esse conhecimento pode ser vendido para terceiros. Além do mais, o usuário da rede social é exposto a publicidade direcionada, que é outra importante fonte de ganhos das redes sociais.

O Superior Tribunal de Justiça muito cedo percebeu que a relação do usuário com os provedores de conteúdo é uma relação de consumo. Há vários precedentes daquele tribunal, antes mesmo do Marco Civil da Internet, em que se reconhece, por exemplo, que a relação do usuário com as chamadas *redes sociais* ou *sites de busca* é relação de consumo, a despeito de não haver nenhum tipo de inversão monetária direta do usuário para a rede. Assim, por exemplo, o seguinte julgado:

> DIREITO CIVIL E DO CONSUMIDOR. INTERNET. RELAÇÃO DE CONSUMO. INCIDÊNCIA DO CDC. GRATUIDADE DO SERVIÇO. INDIFERENÇA. PROVEDOR DE CONTEÚDO. FISCALIZAÇÃO PRÉVIA DO TEOR DAS

[118] SCHWAB, op.cit., p. 63.

> INFORMAÇÕES POSTADAS NO SITE PELOS USUÁRIOS. DESNECESSIDADE. MENSAGEM DE CONTEÚDO OFENSIVO. DANO MORAL. RISCO INERENTE AO NEGÓCIO. INEXISTÊNCIA. CIÊNCIA DA EXISTÊNCIA DE CONTEÚDO ILÍCITO. RETIRADA IMEDIATA DO AR. DEVER. DISPONIBILIZAÇÃO DE MEIOS PARA IDENTIFICAÇÃO DE CADA USUÁRIO. DEVER. REGISTRO DO NÚMERO DE IP. SUFICIÊNCIA.
> 1. A exploração comercial da internet sujeita as relações de consumo daí advindas à Lei nº 8.078/90.
> 2. O fato de o serviço prestado pelo provedor de serviço de internet ser gratuito não desvirtua a relação de consumo, pois o termo "mediante remuneração" contido no art. 3º, § 2º, do CDC deve ser interpretado de forma ampla, de modo a incluir o ganho indireto do fornecedor.
> 3. A fiscalização prévia, pelo provedor de conteúdo, do teor das informações postadas na web por cada usuário não é atividade intrínseca ao serviço prestado, de modo que não se pode reputar defeituoso, nos termos do art. 14 do CDC, o site que não examina e filtra os dados e imagens nele inseridos.
> 4. O dano moral decorrente de mensagens com conteúdo ofensivo inseridas no site pelo usuário não constitui risco inerente à atividade dos provedores de conteúdo, de modo que não se lhes aplica a responsabilidade objetiva prevista no art. 927, parágrafo único, do CC/02.
> 5. Ao ser comunicado de que determinado texto ou imagem possui conteúdo ilícito, deve o provedor agir de forma enérgica, retirando o material do ar imediatamente, sob pena de responder solidariamente com o autor direto do dano, em virtude da omissão praticada.
> 6. Ao oferecer um serviço por meio do qual se possibilita que os usuários externem livremente sua opinião, deve o provedor de conteúdo ter o cuidado de propiciar meios para que se possa identificar cada um desses usuários, coibindo o anonimato e atribuindo a cada manifestação uma autoria certa e determinada. Sob a ótica da diligência média que se espera do provedor, deve este adotar as providências que, conforme as circunstâncias específicas de cada caso, estiverem ao seu alcance para a individualização dos usuários do site, sob pena de responsabilização subjetiva por culpa in omittendo.
> 7. Ainda que não exija os dados pessoais dos seus usuários, o provedor de conteúdo, que registra o número de protocolo na internet (IP) dos computadores utilizados para o cadastramento de cada conta, mantém um meio razoavelmente eficiente de rastreamento dos seus usuários, medida de segurança que corresponde à diligência média esperada dessa modalidade de provedor de serviço de internet.
> 8. Recurso especial a que se nega provimento.[119]

É pelos efeitos que produz sobre a publicidade e o *marketing* que a coleta e tratamento de dados pessoais liga-se ao direito do consumidor. De fato, os modelos de negócio da internet baseiam-se na chamada publicidade direcionada.[120] Como explica Bruno Bioni, "a publicidade direcionada é uma prática que procura personalizar, ainda que parcialmente, tal comunicação social, correlacionando-a a um determinado fator que incrementa a possibilidade de êxito da indução do consumo"[121]. Ela se subdivide em três tipos: a) publicidade contextual; b) publicidade segmentada; e c) publicidade comportamental on-line.

---

[119] BRASIL. Terceira Turma do Superior Tribunal de Justiça. Recurso Especial nº 1193764 SP. Relator: Ministra Nancy Andrighi. Brasília, DF, 10 de dezembro de 2010. **Diário de Justiça Eletrônico**. Brasília, 8 ago. 2011.
[120] BIONI, op. cit., p. 15.
[121] Ibidem., p. 15.

A publicidade contextual é aquela que se dá mediante a veiculação da mensagem publicitária dentro de um ambiente por si mesmo ligado ao tema em questão. Assim, a propaganda de um livro no *site* de uma livraria, ou a divulgação de um filme em um canal de *streaming* são exemplos de publicidade contextual. Essa forma de publicidade está ancorada no objeto, no assunto.

A publicidade segmentada, diferentemente, tem o seu foco num grupo predeterminado de pessoas: o público-alvo. Evidentemente, para implementar esse tipo de publicidade é preciso que se tenha acesso aos dados pessoais do público-alvo, de modo a fazer a sua "clusterização" (agrupamento), com itens como idade, gênero, local de residência, hábitos de compra, renda, etc. O vetor publicitário, neste caso, é o aspecto subjetivo, de tal maneira que a estratégia para atingir o público-alvo é ligar a publicidade aos diferentes espaços que esse público frequenta, nas suas mais diversas movimentações sociais. "Segmenta-se, portanto, a publicidade a uma determinada camada da massa de consumidores, independentemente de qual seja o contexto da plataforma em que a publicidade está sendo veiculada", diz Bioni.

Um dos problemas mais delicados da publicidade segmentada é que ela combina dados pessoais de diferentes indivíduos para inferir comportamentos de manada. Os indivíduos são enquadrados em conjuntos a partir de inferências estatísticas extraídas da manipulação de dados pessoais que apontam padrões de comportamento grupal. Presume-se, com base em modelos matemáticos, que nesses grupos todos terão sempre o mesmo interesse e a mesma resposta a certos estímulos, de tal maneira que essa forma de publicidade é um campo fértil para a criação ou o reforço de estereótipos e vieses.[122]

Finalmente, a publicidade comportamental *online*. Enquanto a publicidade segmentada concentra-se em grupos, a publicidade comportamental on-line vai além e busca explorar mais a fundo os dados pessoais, com o intuito de conceber um retrato detalhado de cada indivíduo, um perfil que o defina na internet. A técnica principal usada aqui é a "perfilização" a partir de dados de navegação *online*. Esse tipo de publicidade tende a ser obsedante, porquanto a construção do perfil jamais se completa, havendo permanente interesse em espreitar os movimentos *online* do indivíduo, para refinar o seu perfil a partir de novas informações, inclusive os *feedbacks* das estimulações baseadas no perfil em construção.

Essa é, sem dúvida, a forma mais agressiva de publicidade, que apenas se tornou factível por meio da internet e das técnicas de tratamento de dados pessoais em meio digital. Se é certo que há algumas vantagens para o consumidor em tal forma de propaganda, não se pode negar,

[122] As consequências dessa técnica são amplamente analisadas em: O'NEIL, op.cit., 2016.

por outro lado, que há nela também riscos ponderáveis aos direitos do consumidor, em especial pelo seu caráter muitas vezes sub-reptício e até manipulativo.

O Código de Defesa do Consumidor — que é de 1990, quando a internet sequer havia sido implantada para o uso de leigos — não tem previsão expressa sobre esse tipo de prática comercial. No entanto, há vários dispositivos que proíbem e punem a publicidade enganosa ou abusiva (CDC, art. 6º, IV; arts. 36-38; art. 60; art. 68). O art. 36, § 2º do CDC chega a tocar no ponto da publicidade manipulativa, ao definir como abusiva qualquer publicidade que seja discriminatória, incite à violência, explore o medo ou a superstição, se aproveite da deficiência de julgamento e experiência da criança, desrespeite valores ambientais, ou que seja capaz de induzir o consumidor a se comportar de forma prejudicial ou perigosa à sua saúde ou segurança. Mas não são mencionados na norma os prejuízos e perigos à intimidade e à personalidade como índices de abusividade da publicidade.

Laura Mendes propõe um catálogo mínimo de direitos do consumidor em face da publicidade comportamental[123]. Segundo ela, a primeira exigência que tem de ser atendida para a prática desse tipo de publicidade é o consentimento do consumidor. Não um consentimento genérico e definitivo, mas sim um consentimento "informado, expresso, específico e anterior nos moldes do sistema *opt in*"[124]. Ademais, procedimentos adicionais têm de ser adotados, tais como: a) anonimização e pseudonimização dos dados pessoais; b) impedimento de coleta de dados sensíveis; c) vedação à coleta de dados pessoais em sites seguros; d) impossibilidade de reidentificação do consumidor; d) proibição da prática desse tipo de publicidade pelos provedores de acesso à internet, visto que estes têm em mãos todas as movimentações do consumidor, podendo assim exercer uma vigilância pervasiva sobre a personalidade do consumidor.[125]

Avaliando essas propostas à luz da LGPD e do Marco Civil da Internet - MCI, verifica-se que a legislação brasileira atual, em alguma medida, contempla a maior parte delas. Com efeito: a) o consentimento, que foi definido como "manifestação livre, informada e inequívoca pela qual o titular concorda com o tratamento de seus dados pessoais para uma finalidade determinada" (LGPD, art. 5º, XII), é normalmente exigível para o tratamento de dados pessoais (LGPD, art. 7º, I e V), com exceção de algumas situações de interesse público (LGPD, art. 7º, II, III, IV, VI), ou de interesse presumível do próprio titular dos dados (LGPD, art. 7º, VII e VIII), ou no "legítimo interesse" do controlador ou de terceiro, em especial para a proteção do

[123] DONEDA, op. cit., p. 225-228.
[124] Ibidem, p. 225.
[125] Ibidem, p. 226-227;

crédito (LGPD, art. 7º, IX e X)[126]; b) A anonimização, definida na LGPD como a "utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo" (LGPD, art. 5º, XI), foi prevista como um direito do titular, mas apenas em relação a dados desnecessários, excessivos ou tratados em desconformidade com a lei (LGPD, art. 18, IV); c) acessos não autorizados aos dados pessoais foram vedados expressamente (LGPD, art. 46); d) A inviolabilidade do sigilo e do fluxo das comunicações via internet, garantida pelo Marco Civil da Internet (MCI, art. 7º, II), virtualmente proíbe a publicidade comportamental pelo provedor de acesso.

2.4 A configuração técnica da internet e o direito à proteção de dados

Todas as técnicas têm uma origem mágica, como diz Milton Vargas[127]. Isso porque os conhecimentos a respeito de práticas sobre o como fazer algo, para além das leis da natureza, normalmente são adquiridos por um grupo pequeno que, de geração em geração, passam em segredo os seus saberes para os legatários da tradição.

Toda técnica tem uma dimensão ética subjacente. Se há mais de uma forma de praticar alguma ação humana, então aquela que foi escolhida diz algo a respeito das opções políticas vigentes. Essas opções, por sua vez, acabam moldando o futuro, num processo de retroalimentação permanente. A esse respeito, Winston Churchill, que antes de tudo era um exímio frasista, disse: "Nós moldamos nossas construções e depois nossas construções nos moldam."[128] A frase tem perfeita aplicabilidade ao ambiente da internet. As soluções técnicas que foram adotadas, desde o nascedouro da rede, moldaram muito do que hoje se vive.

A internet é um sistema sociotécnico muito complexo. Há um amontoado de sistemas, protocolos, padrões, *hardwares* e organizações por detrás do funcionamento da rede. As decisões sobre as tecnologias, por isso mesmo, têm efeito vasto e duradouro, não sendo nunca politicamente neutras. Como dizem O'Hara e Hall, "cada decisão sobre *design* reflete e impõe (eventualmente de forma inconsciente) um equilíbrio de poder, enquanto tensões culturais, econômicas e políticas atuam em todos os problemas da ação coletiva gerados pela modernidade digital"[129].

---

[126] O art. 10 da LGPD tenta delimitar o sentido da expressão "legitimo interesse", mas o tema é bastante complexo e será retomado mais adiante neste trabalho.

[127] VARGAS, Milton. **Para uma Filosofia da Tecnologia**. São Paulo: Editora Alfa Omega, 1994, p. 19.

[128] FRIEDMAN; HENDRY, op.cit., p. 3, tradução nossa. Texto original: "We shape our buildings and afterwards our buildings shape us".

[129] O'HARA, Kieron; HALL, Wendy. Four Internets: The Geopolitics of Digital Governance. **CIGI Papers**, n. 206, p. 128, dez, 2018. Disponível em:

O ponto central no embate entre as diversas visões sobre o futuro da internet é a questão da preservação da privacidade em face das oportunidades de negócio geradas pela rede mundial de computadores. Todos os modelos de negócio que hoje crescem exponencialmente na internet têm por base a infraestrutura de identificação dos usuários. É por ela que as empresas podem atingir cada indivíduo e oferecer-lhe experiências personalizadas e focadas em suas necessidades específicas, que por sua vez são conhecidas pelas empresas a partir de dados gerados pelo próprio usuário por meio de sua navegação na rede, isto é, dos dados pessoais (LGDP, art. 5º, I). A associação dos dados à pessoa é possível graças à infraestrutura da internet, especialmente ao endereço de protocolo de internet (endereço IP), a que alude o art. 5º, III da Lei 12.965/2014 (Marco Civil da Internet).

Isso quer dizer que, quanto mais usuários identificáveis tiver um sistema, mais dados sobre eles serão coletados e mais valor econômico pode ser gerado a partir dessas informações. Robert Metcalfe chegou a formular uma lei matemática para expressar essa relação. Segundo a Lei de Metcalfe, o valor de um sistema de comunicação é diretamente proporcional ao quadrado do número de usuários[130]. O aumento é exponencial porque os dados pessoais de cada usuário, ao serem cruzados com os dos demais, gera conhecimento não apenas sobre ele próprio mas sobre o grupo todo, ao revelar padrões coletivos que podem ser explorados economicamente.

Existe alguma controvérsia sobre se o valor do sistema, de fato, é proporcional ao quadrado do número de usuários, ou a um número menor do que esse. Andrew Odlyzco e Benjamin Tilly propõem um número bem menor[131], mas em todo caso é certo que a adição de novos usuários a uma rede gera considerável grau de riqueza nova. Isso se dá principalmente por causa de um fenômeno que foi batizado de "cauda longa" (*long tail*)[132] por Chris Anderson[133], num artigo de mesmo nome, em que ele demonstrou que grandes volumes de

---

https://www.cigionline.org/sites/default/files/documents/Paper%20no.206web.pdf. Acesso em 07 out, 2020, p. 1, tradução nossa. Texto original: "Every design decision reflects, and imposes (perhaps unconsciously), a balance of power, while cultural, economic and political tensions play out across the collective-action problems generated by digital modernity."

[130] METCALFE'S Law. In: JONATHAN LAW. **A Dictionary of Business and Management**. 5. ed. Milford: Oxford University Press, 2009. Disponível em: https://www.oxfordreference.com/view/10.1093/oi/authority.20110810105406240. Acesso em: 10 out. 2020.

[131] ODLYZKO, Andrew; TILLY, Benjamin. **A refutation of Metcalfe's Law and a better estimate for the value of networks and network interconnections**. 2005. Versão preliminar. Disponível em: https://www.researchgate.net/profile/Benjamin_Tilly/publication/228829389_A_refutation_of_Metcalfe's_Law_and_a_better_estimate_for_the_value_of_networks_and_network_interconnections/links/547f49960cf2ccc7f8b91b2b.pdf. Acesso em: 05 out. 2020.

[132] A referência à "cauda longa" diz respeito ao formato que a curva de vendas apresenta num gráfico cartesiano em que o volume de vendas seja representado no eixo das ordenadas e a quantidade de produtos no eixo das abcissas.

[133] ANDERSON, Chris. The Long Tail. *In*: CONDÉ NAST. **Wired**, 10 jan. 2004. Disponível em: https://www.wired.com/2004/10/tail/. Acesso em: 7 out. 2020.

vendas de algumas mercadorias não tão vendáveis, produzem mais riqueza do que a venda de uma só mercadoria muito vendável. Em outras palavras, atualmente se gera mais riqueza no mercado de nicho que no mercado de massa. Mas, é claro, isso apenas é possível em caso de o varejista ter amplo conhecimento sobre os consumidores e seus hábitos, além de estar em permanente contato com ele.

No mercado de massa, buscava-se encontrar um produto que causasse sensação entre os consumidores (um *hit*), de modo que todo foco estava no produto; ao passo que no mercado de nicho, procura-se conhecer o consumidor e suas preferências individuais para personalizar ao máximo a venda. Quanto mais dados pessoais estão em poder dos varejistas, mais nichos dispersos são descobertos e agregados, mais negócios são realizados e consequentemente mais riqueza é gerada.

Como explica Chris Anderson, no artigo referido, há também um outro aspecto importante: a economia digital supera as limitações próprias do mundo off-line, ao não depender de grandes estruturas físicas espalhadas em diversos lugares para vender um produto ou serviço. Enquanto um filme *blockbuster* dependia de muitas, espaçosas e caras salas de cinema para ser exibido, o mercado de nicho foca no *streaming* direcionado a cada consumidor no conforto da sua própria casa (cujas despesas são pagas, naturalmente, pelo próprio consumidor). Isso barateia o valor de cada produto e permite a distribuição dos ganhos por uma base de produtos muitas vezes maior (a "longa cauda", a que se refere Chris Anderson). Logicamente, isso só é possível porque o consumidor está em permanente conexão com o fornecedor, oferecendo-lhe dados pessoais — por meio da sua navegação pelo canal de *streaming* —, que alimentam os mecanismos de predição, os quais irão permitir a apropriada sugestão de novos títulos personalizados, fechando um ciclo virtuoso de abundância jamais visto.

2.5 Internet de Quinta Geração (5G), Economia da Abundância, Internet das Coisas e Privacidade

A internet 5G promete potencializar a economia da abundância ainda mais, ao preço de maior invasão à privacidade. A Agência de Proteção de Dados da Espanha apresenta um catálogo de oportunidades de negócios e riscos à privacidade trazidos pela tecnologia 5G[134], cujo resumo apresenta-se adiante.

[134] ESPANHA. AGENCIA ESPAÑOLA DE PROTECCIÓN DE DADOS. . **Introduction to 5G technologies and their risks in terms of privacy**. 2020. Disponível em: https://www.aepd.es/sites/default/files/2020-06/nota-tecnica-privacidad-5G-en.pdf. Acesso em: 05 out. 2020.

À medida que conteúdo mais rico, como *streaming* de vídeo, jogos, músicas, etc., é fornecido pela tecnologia 4G, a demanda por largura de banda cresce a ponto de estrangular os meios atuais de transmissão de dados; mas o 5G tem como atender essa demanda.

Com uma banda larga de dados, os consumidores podem desfrutar de experiências imersivas, não apenas usando os tradicionais portais de entrada da internet, que são celulares, computadores pessoais e televisões, mas sim interagindo com a internet o tempo todo, a partir de infinitos pontos de entrada, espalhados em casa, nos objetos circundantes, nas cidades, em toda parte enfim. Assim, os anunciantes podem usar efetivamente aplicativos que consomem muita largura de banda, como Realidade Virtual e Realidade Aumentada[135].

Redes de alta velocidade e baixa latência[136] podem ajudar os desenvolvedores a oferecer suporte para tecnologias emergentes, tais como como dispositivos de Internet das Coisas (IoT, na sigla em inglês para *Internet of Things*), jogos ao vivo, telemedicina sensível ao tempo, computação em nuvem, veículos e aparelhos autônomos, etc.

Admite-se que a Internet 5G irá ampliar as experiências móveis com transmissão de conteúdo de tela para tela, marca para marca em velocidades extremamente rápidas. A 5G pode também desencadear a explosão de dispositivos conectados, incluindo objeto vestíveis (*weareable*) para aparelhos conectados em praticamente todos os locais de convivência humana — casa, escritórios, mercados, etc. — proporcionando a experiência de ubiquidade completa da internet. Tais dispositivos conectados irão gerar enorme quantidade de dados (a maioria pessoais, porque suscetíveis de serem ligados a pessoas naturais), que serão a chave para fornecer experiências hiper-realista para usuários.

Nesse contexto, a privacidade de dados encontra ainda mais riscos do que na atmosfera 4G. De fato, dispositivos IoT, tais como automóveis, geladeiras, lâmpadas, semáforos, etc., constantemente conectados e circundando os movimentos de cada pessoa, transferem grandes volumes de dados pessoais em redes 5G, mediante coleta automatizada. Isso amplia os desafios para proteger a privacidade dos dados dos usuários.

---

[135] Realidade Virtual é um ambiente artificial que é experimentado por meio de estímulos sensoriais (como imagens e sons) fornecidos por um computador e no qual as ações de uma pessoa determinam parcialmente o que acontece no ambiente. Realidade Aumentada é uma versão aprimorada da realidade criada pelo uso de tecnologia para sobrepor informações digitais em uma imagem de algo sendo visualizado por meio de um dispositivo (como uma câmera de smartphone). Cf. https://www.merriam-webster.com/dictionary/ .Acesso em: 10 out. 2020.

[136] Latência é o tempo que um pacote de dados leva para percorrer um trajeto de um ponto de origem ao seu destino, dentro de uma rede de computadores. Cf. GOGONI, Ronaldo. O que é Ping (latência)?: O Ping é uma ferramenta essencial e não serve apenas para levar a culpa quando você perde uma partida online. *In*: MOBILON MÍDIA. **Tecnoblog**. Disponível em: https://tecnoblog.net/263177/o-que-e-ping-latencia/. Acesso em: 10 out. 2020.

No 5G, à medida que mais dispositivos vestíveis e dispositivos inteligentes se conectam à rede, eles transmitem informações cada vez mais confidenciais, como hábitos domésticos ou dados dos sinais vitais do usuário, por exemplo. A privacidade dos dados de localização do usuário, que já é uma preocupação sob a Internet 4G, torna-se alarmante com a 5G. Isso porque a 5G tem uma área de cobertura mais restrita, por isso são necessárias muito mais torres de celular em um raio menor. Tal circunstância permite às operadoras móveis rastrear a localização com alto grau de precisão e até mesmo a trilha de movimento de cada usuário.

Ademais, dispositivos IoT, por questão de eficiência energética, estão sendo concebidos de modo a serem deixados funcionando ininterruptamente. Esses dispositivos podem se tornar um incômodo de vigilância no futuro. Por meio deles, as empresas (e também os estados) podem manter o indivíduo num estado de vigilância total, que vai dos hábitos até os pensamentos.

Não é por outra razão que a disputa geopolítica EUA—China, da qual já se falou acima, encontra na tecnologia 5G um campo de batalha perfeito. Os EUA têm trabalhado firmemente, inclusive no Brasil[137], para que a infraestrutura de 5G não seja implementada em território de países aliados por empresas chinesas, em particular a Huawei, sob o argumento de que é possível, por meio dessas tecnologias, espionar o tráfego de dados e obter informações pessoais virtualmente de toda a população de um país. Confirma-se, assim, que a proteção dos dados pessoais é complexa e toca não apenas temas de interesse individual, senão também de ordem política.

2.6 O consentimento do titular dos dados

O consentimento tem muitas funções no âmbito do direito, a maioria delas estando associada à ideia de abertura para a realização de um ato que, sem ele ou contra ele, não poderia ser realizado. Assim, o fundamento do consentimento está na liberdade individual para autorregulação dos próprios interesses, quando estes são confrontados com os de outrem em espaços não ocupados por legislação cogente.

As obrigações contraídas sem a necessidade do consentimento são apenas aquelas a todos impostas por lei validamente aprovada e dentro das balizas constitucionais — princípio da legalidade. Logo, o consentimento é que permite a expressão da individualidade de cada um, por meio da assunção de obrigações específicas, em troca de algum benefício correspondente.

---

[137] BBC NEWS (Brasil). **Huawei, Trump, Bolsonaro e China**: o que o brasil tem a ganhar e perder se ceder aos eua no 5g?. o que o Brasil tem a ganhar e perder se ceder aos EUA no 5G?. 2020. Disponível em: https://www.bbc.com/portuguese/brasil-54634201. Acesso em: 10 out. 2020.

É pressuposto do consentimento juridicamente válido, portanto, que exista de direito e de fato a independência necessária para a assunção de compromisso a partir da vontade livre. As principais dificuldades práticas para a aplicação do consentimento como critério de validação estão em: a) delimitar as situações nas quais o consentimento não pode surtir efeito, porquanto a regulamentação cogente deve prevalecer; e, não ocorrendo a primeira hipótese, b) avaliar em concreto o grau de independência e entendimento necessários para que se considere legítima e vinculante a expressão da vontade.

Opções políticas de cada país traçarão diferentes espaços de ação para o consentimento no que diz respeito à proteção de dados pessoais. Assim, por exemplo, enquanto a União Europeia tende a um paternalismo maior, limitando bastante a função do consentimento por meio de regulação abundante, os Estados Unidos tendem a conferir maior papel à autogestão da privacidade.[138]

A articulação entre consentimento e regulação está no cerne das discussões que envolvem a institucionalização das políticas de proteção de dados. As controvérsias subjacentes ao papel do consentimento na economia dos dados, por isso mesmo, são universais. Antes de tudo, discute-se se o consentimento efetivamente confere ao titular um autêntico controle sobre os dados. Daniel Solove menciona que há diversos estudos empíricos que demonstram a debilidade cognitiva dos titulares para usarem lucidamente o consentimento como instrumento de autoproteção[139]. Adicionalmente, é preciso considerar que as técnicas de tratamento de dados digitais são profusas e estão em constante evolução, de tal maneira que nem mesmo especialistas têm condições de prever o quanto de conhecimento se pode produzir a partir de um conjunto determinado de dados. O consentimento, nesse contexto, nunca é totalmente informado.

As fragilidades do consentimento como âncora da proteção de dados têm duas origens principais: a) problema cognitivos, que têm a ver com a forma como seres humanos tomam decisões concretamente; e b) problemas estruturais, que têm a ver com a forma como são desenhadas as tecnologias da informação.[140]

---

[138] Cf. SOLOVE, Daniel. Autogestión de la privacidad y el dilema del consentimiento. **Revista Chilena de Derecho y Tecnología**, n. 3, p. 11-47, 23 jan. 2014. Universidad de Chile. http://dx.doi.org/10.5354/0719-2584.2013.30308. Disponível em: https://rchdt.uchile.cl/index.php/RCHDT/article/view/30308/32095. Acesso em: 01 out. 2020.

[139] SOLOVE, Daniel. Autogestión de la privacidad y el dilema del consentimiento. **Revista Chilena de Derecho y Tecnología**, n. 3, p. 11-47, 23 jan. 2014. Universidad de Chile. http://dx.doi.org/10.5354/0719-2584.2013.30308. Disponível em: https://rchdt.uchile.cl/index.php/RCHDT/article/view/30308/32095. Acesso em: 01 out. 2020, p.13.

[140] Ibidem, p.17.

2.6.1 Problemas cognitivos

Os problemas cognitivos aqui referidos são tanto aqueles ligados à capacidade jurídica em geral, que são abordados pelo direito civil como causas de incapacidade (CC, arts. 1º a 10) ou de vício de consentimento ou sociais (CC, arts. 138 a 165; LGPD, art. 8º, §3º), como os mais propriamente ligados às heurísticas específicas do *ethos* da internet. Neste trabalho, apenas os últimos serão objeto de análise, haja vista o propósito da pesquisa.

Como já referido, o primeiro grande obstáculo oposto ao titular de dados para que ele expresse o seu consentimento é a falta de informação ou a desinformação. Os estudiosos do tema apontam que é muito alto o nível de informação necessário para a manifestação de um consentimento bem formulado. É virtualmente impossível para o indivíduo leigo ter exato conhecimento das potencialidades dos seus dados para gerar informações e conhecimentos em mecanismos de aprendizado de máquina, de tal maneira que qualquer consentimento por ele manifestado terá uma zona sombreada em que não haverá segurança sobre as consequências desse consentimento. Por isso mesmo, a lei exige que o agente de tratamento use de boa-fé e se atenha à finalidade específica considerada no momento do consentimento (LGPD, arts. 6º, I), à adequação entre o tratamento de dados e o contexto (LGPD, art. 6º, II), à eficiência do tratamento, com o uso mínimo possível dos dados (LGPD, art. 6º, III), à qualidade dos dados considerados (LGPD, art. 6º, V). E, mesmo em dados pessoais tornados públicos, para os quais se dispensa a manifestação de consentimento para o tratamento, a finalidade deve ser considerada (LGPD, art. 7º, §§3º, 4º e 7º). Autorizações genéricas de tratamento, por sua vez, são tidas como nulas (LGPD, art. 7, §4º).

Acresce que, diferentemente do consentimento negocial, que cria um vínculo insuscetível de rompimento unilateral, a revogação imotivada do consentimento para o tratamento de dados é um direito do titular, que lhe assegura o permanente controle sobre seus dados, mesmo depois de ter inicialmente consentido em seu uso para alguma finalidade (LGPD, art. 8º, §5º). Em se tratando a proteção de dados de um direito da personalidade, embora se possa aceitar a licença temporária e voluntária da utilização de dados pessoais para finalidades específicas, mediante retribuição econômica ou não, não pode o titular ficar permanentemente preso a um eventual consentimento anterior. É assim que a revogação imotivada do consentimento não se assemelha a um distrato, que depende de recíproca concordância das partes; a revogação é direito subjetivo potestativo, que assegura ao titular dos dados a sua autodeterminação informativa. Nesse sentido, observa Laura Mendes:

> A possibilidade de revogação do consentimento sem justificativa parece mais adequada dogmaticamente, tendo em vista a natureza da proteção de dados como uma espécie dos direitos da personalidade. O pressuposto de descumprimento de obrigações contratuais para o consentimento é típico de um modelo negocial. Dada a natureza de direito à personalidade, em que o exercício do direito à proteção de dados se realiza pelo consentimento, a possibilidade de revogação é inerente ao próprio direito.[141]

Nas situações em que o consentimento é requerido, se houver mudanças da finalidade para o tratamento de dados pessoais não compatíveis com o consentimento original, o controlador deverá informar previamente o titular, podendo este revogar o consentimento, caso discorde das alterações (LGPD, art. 9, §2º). Vale ressaltar que o tratamento de dados sensíveis obviamente também se subordina à finalidade associada ao consentimento (LGPD, art. 11, I) e, em alguns casos estritos, podem ser tratados até mesmo sem o consentimento do titular (LGPD, art. 11, II). Igual disciplina segue o tratamento de dados pelo Poder Público (LGPD, arts. 23 e 26).

A finalidade também é importante porque representa um termo resolutivo para o encerramento do tratamento de dados pessoais (LGPD, art. 15, I). Após o cumprimento da finalidade considerada pelo titular ao emitir o consentimento, os dados devem ser eliminados, autorizada a conservação apenas para algumas finalidades específicas (LGPD, art. 16).

A fidelidade ao propósito do tratamento de dados, portanto, é que garante a eficácia do consentimento. Esta circunstância produz aquilo que já se chamou de "paradoxo da privacidade"[142], pois apenas quando a finalidade é deturpada é que o titular poderá obter a tutela com base no consentimento.

Outro ponto relevante acerca da fragilidade do consentimento como esteio da política de proteção de dados está em que as decisões humanas, mesmo quando o indivíduo tem bons elementos para avaliar a situação, não são resultado de um processo racional, mas sim de heurísticas que dependem muito da forma e do contexto em que questão é apresentada para ser decidida.

No âmbito da internet, as decisões sobre consentir ou não ao processamento de dados são tomadas em contextos muito nebulosos e hostis à negativa. Para começar, normalmente as perguntas sobre o tema são feitas ao titular dos dados em termos maniqueístas: ou ele consente no processamento dos seus dados, ou não tem acesso ao bem ou serviço que deseja. É, portanto, uma escolha entre caminhos inconciliáveis. Naturalmente, mesmo para um usuário atento e bem informado, isso gera uma pressão extra no processo decisório.

---

[141] MENDES, op. cit., p. 64.
[142] Cf.: DONEDA, op. cit., p. 299.

Para além disso, a velocidade e a simplicidade características do processo decisório a respeito dos dados na internet, que se dá apenas por um clique numa caixa de diálogo, favorece a utilização de heurísticas, que são atalhos mentais para resolver problemas triviais. Assim, por exemplo, podemos referir as seguintes heurísticas como perfeitamente aplicáveis ao processo decisório do usuário comum no dia a dia do usuário da internet:

a) Heurística da Substituição de Atributo – Esse tipo de pensamento rápido substitui a questão original mais complexa por outra mais simples: assim, em vez de pensar sobre a questão "Você autoriza o uso dos seus dados pessoais para personalizar a sua navegação?", é bem mais simples responder à pergunta "Você não acha que vale a pena dar logo essa autorização, já que ela não vai lhe causar dano imediato?". Essa última pergunta simplesmente não foi feita, além de ser tendenciosa, mas o cérebro sub-repticiamente substitui a primeira pela segunda, como forma de facilitar o processo decisório. Daniel Kahneman, que estudou a fundo o problema das heurísticas e dos vieses nas decisões rápidas, afirma que a essência das chamadas "heurísticas intuitivas" está nisto: "quando confrontados com uma questão difícil, muitas vezes respondemos a uma mais fácil em lugar dela, normalmente sem perceber a substituição."[143] Ora, é evidente que, mesmo para um especialista, é muito complexo avaliar todos os custos e benefícios de licenciar a autorização do uso dos dados pessoais para a personalização da navegação, porque na verdade não se tem a ciência exata da extensão e dos propósitos dessa personalização. Então, torna-se mais simples, para decidir, substituir essa pergunta por algo mais palpável, como confrontar perdas e ganhos imediatos da autorização. De logo, o titular ganha acesso a uma informação, a um serviço ou a um bem do seu interesse; remotamente, ele imagina, pode até advir alguma importunação, mas isso se resolve depois. Então, é natural que a decisão pelo "sim" praticamente se impõe por meio da substituição do atributo.

b) Heurística da Disponibilidade – Esse é um atalho mental que funciona mediante o uso do que vem à cabeça em primeiro lugar, ou seja, é uma forma de decidir com o que está à mão, desprezando tudo mais que não se sabe. Como explica Kahneman, é possível resistir a essa heurística, mas é muito cansativo[144]. No cotidiano, portanto, ela tende a se impor na grande maioria das situações. Imagine-se o exemplo de alguém que precisa

[143] KAHNEMAN, Daniel. **Rápido e devagar**: duas formas de pensar. Rio de Janeiro: Objetiva, 2012. Tradução de Cássio Arantes de Leite, p. 22.
[144] Ibidem, p. 167.

decidir se dá ou não o consentimento para o tratamento dos seus dados pessoais a um mercado que vende produtos com desconto, desde que o cliente adira a um "plano de fidelidade". Uma decisão como essa, para ser bem fundamentada, precisaria avaliar o que outros mercados oferecem, qual o valor dos descontos, quanto valem os dados pessoais, etc.. É muito mais fácil e fluente lembrar-se de uma experiência positiva recente (por exemplo, alguém conhecido que foi sorteado no programa de fidelidade desse mercado), e decidir com base nisso, do que esquadrinhar todos os fatores positivos e negativos que estão em jogo. Um traço perigoso da heurística da disponibilidade é que ela pode ser manipulada pela publicidade com técnicas de *priming*, adiante explicado. À medida que imagens positivas são associadas sistematicamente a um produto, a um serviço, a uma marca, elas se tornam disponíveis como modelos para uma decisão rápida, por isso aumentam as probabilidades de que o titular tenha contato com essas associações e decida com base nelas.

c) Heurística da Ancoragem – A ancoragem normalmente decorre da sugestão ou *priming*. O *priming* se refere a uma descoberta científica segundo a qual a exposição a uma palavra ou a uma ideia faz o cérebro evocar imediata e naturalmente várias palavras ou ideias associadas àquela que foi apresentada, sem que o indivíduo tenha controle sobre esse processo cognitivo[145]. Kahneman[146] dá vários exemplos — alguns, bizarros — de como a psicologia experimental avançou na compreensão desse fenômeno. Eis alguns deles:

> Estudos sobre efeitos de *priming* renderam descobertas que ameaçam nossa autoimagem como autores conscientes e autônomos de nossos julgamentos e nossas escolhas. Por exemplo, a maioria de nós pensa no ato de votar como um gesto deliberado que reflete nossos valores e nossas avaliações da política e não influenciado por questões irrelevantes. Nosso voto não deveria ser afetado pelo local onde está a urna, por exemplo, mas é. Um estudo sobre padrões de voto no Arizona em 2000 revelou que o apoio a propostas de aumentar a verba para escolas era significativamente maior quando o prédio de votação era uma escola, em vez de qualquer outro lugar nas redondezas. Um experimento separado mostrou que expor as pessoas a imagens de salas de aula e armários escolares também aumentava a tendência dos votantes a apoiar a iniciativa pró-escola. O efeito das imagens foi maior do que a diferença entre os pais e os demais eleitores! O estudo do *priming* tem ido um pouco além das demonstrações iniciais de que lembrar as pessoas sobre a velhice faz com que caminhem mais devagar. Agora sabemos que os efeitos do *priming* podem atingir cada recesso de nossas vidas.

A ancoragem ocorre em cima da imagem, palavra ou ideia sugestiva. O indivíduo, colocado proposital ou casualmente diante de um desses signos, automaticamente faz

[145] Ibidem, p. 69.
[146] Ibidem, p. 72-77.

associações a temas conexos, como se fosse um autopreenchimento da mente para o contexto. Está claro que estratégias de *marketing* que usam esse tipo de conhecimento invadem a esfera mais íntima da pessoa, e isso tem sido objeto de preocupação mundo afora, principalmente considerando os avanços das neurotecnologias, capazes de "ler" pensamentos. No Chile, recentemente foi apresentado um projeto reforma da Constituição e um projeto de lei para regulamentar os chamados "neurodireitos"[147]. A ideia é proteger a "privacidade mental" por meio de um direito de não ser manipulado. Intenta-se também evitar que novas tecnologias que aumentam a capacidade cerebral possam se constituir em bens privados com valor econômico, porquanto isso poderia implicar a criação de castas de super-humanos.

Há muitas outras heurísticas cuja aplicação pode, em certos contextos, explicar a fragilidade do consentimento como um critério de validação de tratamento de dados. Embora não se possa negar a importância do consentimento como um dos pilares do processo de legitimação do tratamento de dados, dado o regime de liberdade individual que vigora nos estados de direito, não se pode sobrecarregar esse caminho, confiando a ele a solução de toda a extensa e complexa cadeia de consequências da manipulação dos dados pessoais.

2.6.2 Problemas estruturais

Há problemas estruturais que debilitam o consentimento como instrumento fundamental de legitimação do tratamento de dados. Entre eles, destacam-se três: a) o efeito da escala; b) o efeito de agregação; e c) o problema da valoração dos danos.

Para começar, há o problema da escala: a quantidade de *sites*, aplicativos, redes sociais, etc., que um indivíduo frequenta nos dias de hoje é enorme. Estima-se que um cidadão americano, por exemplo, visita em média cem páginas de internet diferentes todo mês[148]. Supondo-se que cada um dos *sites* tenha uma política de privacidade ligeiramente diversa, seria preciso avaliar todas elas para manifestar o consentimento informado. Ocorre que é virtualmente impossível, para qualquer ser humano, lidar de modo eficaz com tamanha quantidade de informações de forma manual. Assim, mesmo que haja *websites* com política de privacidade respeitosa aos direitos individuais, o usuário teria de esquadrinhar toda a sua vasta navegação para separar os *sites* com boa política de privacidade daqueles que não têm essa característica.

---

[147] Cf. SENADO DO CHILE. **Boletín 13828-19**: sobre protección de los neuroderechos y la integridad mental, y el desarrollo de la investigación y las neurotecnologías.2020. Disponível em: https://www.senado.cl/appsenado/templates/tramitacion/index.php?boletin_ini=13828-19. Acesso em: 02 out. 2020.

[148] Cf. SOLOVE, op.cit., p.36.

O efeito de agregação é outro grave problema estrutural que demonstra como o consentimento nunca alcança todos os aspectos do problema da privacidade na internet. Tal efeito se refere ao fato de que os dados têm grande poder de gerar conhecimento quando agregados. Assim, o indivíduo é convidado a consentir em entregar o dado A, totalmente inofensivo; em outra oportunidade, ele entrega o dado B, também aparentemente sem maior relevância; e assim sucessivamente. O agente de tratamento de dados, que tem todos os dados A, B, C..., poderá combiná-los convenientemente, tanto entre si quanto com os dados de outros usuários, por modelos de aprendizado de máquina, para inferir conexões e *insights* que jamais poderiam ter sido imaginados ou processados por um ser humano, de tal modo que o consentimento, nesse contexto, opera em campo incógnito, mesmo para pessoa de diligência normal e em circunstâncias não excepcionais.

O efeito de agregação fragiliza inclusive as bases da própria legislação de proteção de dados. É que essa legislação está calcada substancialmente sobre a definição de dado pessoal como sendo "informação relacionada a pessoa natural identificada ou identificável" (LGPD, art. 5º, I). Sucede que, muitas vezes, uma informação é dada de forma anônima, portanto não seria dado pessoal e consequentemente não estaria no raio de proteção da lei, mas quando ela é combinada com muitas outras, inclusive de outras pessoas, ela pode gerar a identificação *a posteriori* do indivíduo[149]. Só então ela terá se transformado em dado pessoal? E o tratamento anterior, que se deu sobre dados anônimos, teria a proteção da LGPD?[150] Essas são questões que se colocam em razão da complexidade estrutural que o efeito de agregação tem sobre todo o ecossistema de proteção de dados. Elas mostram que o problema da privacidade na internet não é apenas individual, senão também coletivo.

O problema da valoração dos danos decorre da forma como o consentimento é requerido no contexto da internet. Enquanto normalmente se requer o consentimento para o tratamento dos dados no momento da sua coleta, os danos apenas se produzem algum tempo depois e mediante arranjos, em parte previsíveis, em parte imprevisíveis, de muitos fragmentos que são espalhados pelo usuário na rede (cada qual com um consentimento em separado). Assim, os danos se produzem mais pela montagem que se engendra com os dados, que propriamente por cada conjunto de dados a respeito dos quais foi autorizado o tratamento. Numa metáfora muito

[149] SOLOVE, op.cit., p. 26-27

[150] O art. 12 da LGPD toca no problema, mas apenas para falar de dados pessoais que foram "anonimizados". Não há menção ao fato de que certos dados, realmente anônimos na origem, podem ser matéria-prima para inferências sofisticadas, que podem chegar a identificar o titular dos dados.

expressiva, Daniel Solove diz que a picada de uma abelha pode ser inofensiva, mas centenas ou milhares delas são letais[151].

A transformação dos dados em informação e, depois, em conhecimento, é um processo muito complexo, que articula diferentes elementos técnicos (tanto de *hardware* como de *software*), econômicos (modelos de negócio) e políticos (regulação e fiscalização). Os danos provocados aos usuários, por isso mesmo, são gerados nessa longa e pouco compreensível cadeia causal, e normalmente aparecem muito depois da coleta dos dados. É difícil, por isso, valorá-los, bem como imputar a responsabilidade a quem de direito.

Daí porque é evidente que o enquadramento de todos os problemas de privacidade na internet no raio de alcance do consentimento é um equívoco. Também parece equívoco e ineficaz buscar todas as soluções no polo oposto, do paternalismo estatal, tanto mais em regimes democráticos, que consideram a liberdade individual um valor fundamental a ser preservado.

Embora consentimento e regulação tenham um papel importante nessas questões, eles devem ser combinados com elementos orgânicos do próprio *ethos* da internet, em especial com o chamado Design Sensível ao Valor (V*alue Sensitive Design*), uma abordagem interacional e orgânica que considera que as tecnologias devem ser desenvolvidas e aplicadas com ênfase no bem-estar humano, com foco na ética e moralidade humanas, e não apenas na eficácia de um processo produtivo qualquer[152].

O Design Sensível ao Valor tem uma série de compromissos para ajustas processos e técnicas a valores humanos. Como explicam Firdeman & Hendry, esses compromissos incluem: a) a proposição-chave de que a relação entre tecnologia e valores humanos é fundamentalmente interacional; b) a análise das partes interessadas, direta ou indiretamente; c) as distinções entre os valores do designer, valores explicitamente apoiados pelo projeto e valores das partes interessadas; d) os níveis de análise individual, do grupo e da sociedade; e) as investigações conceituais integrativas e iterativas, técnicas e empíricas; f) a coevolução da tecnologia e da estrutura social; e g) um compromisso com o progresso, mas não com a perfeição[153].

Evidentemente, essa é uma abordagem dinâmica que está, ela mesma, em evolução, para incorporar e digerir as complexidades de cada campo de avaliação das técnicas e de suas consequências. A inclusão de critérios de avaliação ampliados serve não apenas para julgar as consequências ou os danos de uma tecnologia, mas igualmente para influenciar o design de

---

[151] SOLOVE, op.cit., p.27.
[152] FRIEDMAN, Batya; HENDRY, David G. **Value Sensitive Design**: shaping technology with moral imagination. Cambridge (Ma): Mit Press, 2019, p. 4.
[153] Ibidem, p. 4-5.

novas tecnologias, durante todo o processo de sua concepção. A LGPD, aliás, adota essa visão ampliada, no art. 46, §2º, ao afirmar que as medidas de segurança, técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou qualquer forma de tratamento inadequado ou ilícito, devem ser adotadas "desde a concepção do produto ou serviço até a sua execução."

A orientação do Design Sensível ao Valor compromete-se também com a crítica e avaliação dos próprios valores humanos, em sua articulação recíproca, bem como no seu contato com elementos da filosofia, da antropologia, da psicologia, sociologia, dos estudos da interação homem-máquina, entre outros. Para o direito, a riqueza dessa análise permite uma avaliação mais lúcida de riscos e danos, facilitando a tomadas de decisão[154] e serve como poderoso vetor interpretativo que deve ser levado em conta tanto no momento da formulação como da aplicação das normas jurídicas pertinentes a esse campo do interesse.

As chamadas Tecnologias de Facilitação da Privacidade (*Privacy Enhancing Technologies* – PETs) podem ser consideradas um bom exemplo de Design Sensível ao Valor, embora numa fase ainda embrionária porque focadas apenas na tecnologia em si. Nessas tecnologias, como explica Bruno Bioni, está embutida "a ideia de que a proteção de dados pessoais deve orientar a concepção de um produto ou serviço"[155]. Denomina-se esse tipo de solução como *privacy by design*-PbD.

Há inúmeros exemplos de medidas de *privacy by design*[156]. A criptografia é um deles, e tem como objetivo assegurar a confidencialidade das comunicações. A anonimização dos dados pessoais também é uma técnica de *privacy by design*. Por ela, evita-se a associação dos dados ao seu titular, ou, como diz a LGPD (Art. 5º, XI), anonimização consiste na "utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo". A navegação anônima é outro mecanismo importante de proteção da privacidade, com grande semelhança com o processo de anonimização.

Ann Cavoukian, a criadora do conceito de *privacy by design*, observa que a necessidade desse tipo de solução decorre da observação de que respostas meramente regulatórias não têm a eficácia necessária para proteger a privacidade no ambiente da internet. Segundo ela, essa abordagem busca uma funcionalidade total das tecnologias, em que elas assegurem a

[154] Ibidem.
[155] BIONI, op. cit., p.167.
[156] Ibidem, p.168.

privacidade do usuário e também a sustentabilidade das organizações que tratam os dados. A *privacy bv design* estende-se sobre três aplicações principais: 1) sistemas de Tecnologia da Informação; 2) modelos de negócio responsáveis; e 3) projetos físicos e de infraestrutura de rede. A articulação desses elementos, respeitando tanto a dignidade dos usuários como a sustentabilidade das empresas, é que pode assegurar a implementação de uma atmosfera sadia na rede, para além do consentimento e da regulação.[157]

Ainda segundo Ann Cavoukian[158], há sete princípios que de certo modo resumem toda a ideia de *privacy by design*:

a) Proatividade e não reatividade (prevenção e não correção) – A PbD volta-se para evitar que riscos à privacidade se materializem; ela não oferece respostas para infrações de privacidade já ocorridas (estas são objeto de outros meios de tutela);
b) Privacidade como configuração padrão (*Privacy by Default*) – A PbD deve ser incorporada às tecnologias como padrão, não devendo incumbir ao usuário nenhuma ação para aumentar a sua privacidade;
c) Privacidade incorporada ao design – A PbD deve estar está embebida no design e na arquitetura dos sistemas de TI e nas práticas de negócios, sem diminuir-lhes a funcionalidade; não deve ser um aposto acrescentado após alguma violação;
d) Funcionalidade completa – A PbD objetiva acomodar todos os interesses e objetivos legítimos envolvidos no processo de comunicação, em forma de soma positiva e "ganha-ganha", não por meio de uma abordagem datada de soma zero, onde compensações desnecessárias são feitas. Evita-se, assim, as falsas dicotomias, como entre segurança e privacidade, por exemplo;
e) Segurança de ponta a ponta – A PbD intenta obter a proteção total do ciclo de vida dos dados, desde a coleta até a sua eliminação;
f) Visibilidade e transparência – A PbD visa a garantir a todos os interessados (*stakeholders*) que, seja qual for a prática de negócios ou tecnologia envolvida, ela está agindo de acordo com as promessas e objetivos declarados nos termos de privacidade, de modo que haja a possibilidade real de verificação independente, dentro da ideia de confiança, mas com averiguação factível;

---

[157] PRIVACY BY DESIGN (Canadá). **The 7 Foundational Principles**. 2011. Disponível em: https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf. Acesso em: 23 out. 2020.
[158] Ibidem.

g) Respeito pela privacidade do usuário – A PbD exige que tantos os desenvolvedores como os operadores coloquem os interesses do usuário em primeiro lugar, oferecendo-lhes alternativas com fortes padrões de privacidade, avisos apropriados e capacitação amigável.

3 DECISÕES AUTOMATIZADAS: DEFINIÇÃO, BENEFÍCIOS E RISCOS

3.1 Conceito

A concepção de uma decisão automatizada envolve vários elementos, e somente se tornou possível com o uso de computadores eletrônicos. Na verdade, apenas em sentido metafórico se pode falar em "decisão" aqui, porque a máquina não age de modo consciente com algum propósito, mas apenas efetua cálculos aritméticos, segundo um programa (algoritmo) e conforme os dados que a alimentam. Logo, as máquinas apenas podem emular a parcela calculável da inteligência humana, não o livre-arbítrio, nem sentimentos, nem emoções.[159]

Como explicam Ajay Agrawal, Joshua Gans e Avi Goldfarb[160], quando a máquina toma uma decisão, ela usa dados de entrada (imagens, textos, sons, etc., que têm de ser reduzidos a um formato digital legível pela máquina), para fazer uma predição. A predição está baseada no "conhecimento" que o algoritmo ou modelo adquiriu na fase de treinamento, com os chamados dados de treinamento. Combinando a predição com o julgamento (escolha da solução, segundo o interesse do programador/desenvolvedor, expresso no algoritmo ou modelo), a máquina de decisão automática indica uma ação a ser efetivada (por humano ou outra máquina) e essa ação leva a um resultado (eventualmente com uma recompensa associada pelo programador). O resultado fornece ao modelo um *feedback* (positivo ou negativo), que assim realimenta todo o processo para decisões futuras.

A diferença entre algoritmo e modelo é fundamental para entender posteriores desdobramentos jurídicos relacionados às decisões automatizadas. Michael Kearns e Aaron Roth[161] explicam que a distinção entre algoritmo e modelo está em que o segundo é o resultado da aplicação do primeiro sobre uma vasta coleção de dados. Enquanto o algoritmo é o conjunto de regras que, aplicadas a um conjunto finito de dados, pode solucionar problemas semelhantes em tempo finito, o modelo é, por assim dizer, um algoritmo com experiência prática anterior em avaliar dados. O modelo tem, portanto, um *background* que condiciona o seu modo de tratar dados novos, encaixando-os na sua "pré-compreensão". Dizem os referidos autores:

---

[159] É certo que existem debates filosóficos, filmes e livros sobre a possível ascensão das máquinas inteligentes ao nível da autoconsciência, quando então ocorreria a singularidade, isto é, um crescimento teoricamente infinito da inteligência das máquinas sem a intervenção humana. Porém, tais discussões estão fora do propósito desta pesquisa. Para um bom panorama do tema, Cf.: CHACE, Calum. **Surviving AI**: the promise and peril of artificial intelligence. Oxford: Three Cs, 2015. Kindle Edition.

[160] AGRAWAL, Ajay; GANS, Joshua; GOLDFARB, Avi. **Máquinas Preditivas**: a simples economia da inteligência artificial. Rio de Janeiro: Editora Alta Books, 2018. Tradução de Wendy Campos, p. 74.

[161] KEARNS, Michael; ROTH, Aaron. **The Ethical Algorithm**: the science of socially aware algorithm design. New York: Oxford University Press, 2019. Edição Kindle, p.9.

> As we've suggested, many of the algorithms we discuss in this book would more accurately be called models. These models, which make the actual decisions of interest, are the result of powerful machine learning (meta-) algorithms being applied to large, complex datasets. A crude but useful sketch of the pipeline is that the data is fed to an algorithm, which then searches a very large space of models for one that provides a good fit to the data. Think of being given a cloud of 100 points on a piece of paper, each labeled either "positive" or "negative," and being asked to draw a curve that does a good but perhaps imperfect job of separating positives from negatives. The positive and negative points are the data, and you are the algorithm—trying out different curves until you settle on what you think is the best separator. The curve you pick is the model, and it will be used to predict whether future points are positive or negative But now imagine that instead of 100 points, there are 10 million; and instead of the points being on a 2-dimensional sheet of paper, they lie in a 10,000-dimensional space.[162]

A "experiência" do algoritmo com os dados de treinamento é aprimorada por meta-algoritmos que otimizam o trabalho de construção do modelo, mediante a revisão sistemática dos dados de saída, segundo o resultado desejado pelo programador, para melhor agrupá-los e interrelacioná-los. O meta-algoritmo mais conhecido e usado é de *backpropagation*, que resumidamente pode ser descrito como um conjunto de instruções para reanalisar várias vezes os dados de saída e corrigir erros de avaliação porventura verificados, mediante um processamento inverso, melhorando o desempenho do modelo e reequilibrando os pesos dos fatores em jogo para a tomada de decisão[163].

Assim, por exemplo, uma máquina de reconhecimento facial para fins de localização de possíveis foragidos da justiça que estejam circulando em áreas públicas funciona da seguinte maneira: 1º) ela coleta os dados automaticamente (imagens), por meio de câmeras apontadas para os transeuntes em vias públicas (dados de entrada); 2º) o modelo utilizado para analisar esses dados, comparando-os com as imagens dos foragidos armazenadas em seus arquivos, foi previamente treinado com dados de muitos prisioneiros (dados de treinamento), de modo a fazer a associação tida como "correta" pelo programador; 3º) feito o cruzamento, se for encontrado

---

[162] Como sugerimos, muitos dos algoritmos que discutimos neste livro seriam chamados de modelos com mais precisão. Esses modelos, que tomam as decisões reais de interesse, são o resultado de poderosos algoritmos de aprendizado de máquina (meta-) aplicados a conjuntos de dados grandes e complexos. Um esboço rudimentar, mas útil, do pipeline é que os dados são alimentados para um algoritmo, que então procura um espaço muito grande de modelos por um que forneça um bom ajuste aos dados. Pense em receber uma nuvem de 100 pontos em um pedaço de papel, cada um rotulado como "positivo" ou "negativo", e ser solicitado a desenhar uma curva que faz um bom, mas talvez imperfeito trabalho de separar os positivos dos negativos. Os pontos positivos e negativos são os dados, e você é o algoritmo - experimentando curvas diferentes até chegar ao que você acha que é o melhor separador. A curva que você escolhe é o modelo e será usado para prever se os pontos futuros são positivos ou negativos. Mas agora imagine que em vez de 100 pontos, há 10 milhões; e em vez de os pontos estarem em uma folha de papel bidimensional, eles ficam em um espaço de 10.000 dimensões (tradução nossa)

[163] Para descrição dos aspectos matemáticos da questão, Cf AGGARWAL, Charu C. **Neural Networks and Deep Learning**: a textbook. New York: Springer International Publishing, 2018, p. 21 e ss. Esse tipo de meta-algoritmo é usado para otimizar mecanismos de *deep learning* que, como se explicará adiante, são os mais utilizados atualmente em aplicações práticas da chamada Inteligência Artificial.

uma correspondência (*match*), a máquina faz a *predição* de que ali está um foragido, com base no alto nível de probabilidade de a imagem coincidir com a do foragido X, por exemplo; 4º) em seguida, a máquina "julga" e aponta aquele suspeito para o operador; 5º) com base nesse julgamento, adota-se uma ação, que são os atos posteriores (que podem ser humanos ou automatizados também — no caso, a detenção do sujeito) que levarão ao resultado (no caso, prisão correta ou incorreta). Conforme esse resultado tenha sido correto ou incorreto, a depender de uma análise humana posterior, a máquina é informada, por *feedback*, para reforçar ou não aquele julgamento.

As regras de julgamento terão sido dadas pelo programador[164] com base em níveis estatísticos de confiabilidade em ambiente de incerteza, daí a semelhança desse processo automatizado com o funcionamento da mente humana. A grande capacidade de adaptação ao ambiente é o ponto forte da inteligência humana; o cérebro humano é capaz de reconhecer padrões, generalizá-los e de ajustar a decisão tendo em conta pequenas mudanças nesses padrões. Os modelos que trabalham com *machine learning*, em particular os de *deep learning*, buscam reproduzir artificialmente essa capacidade adaptativa do funcionamento orgânico do cérebro e, por isso, estão no centro das mais importantes e avançadas aplicações práticas do que se convencionou chamar da Inteligência Artificial[165].

Observa-se que a decisão automatizada, para além do algoritmo, é fortemente influenciada pelos dados, mais especificamente por três tipos de dados: a) os *dados de treinamento*; b) os *dados de entrada*; e c) os *dados de feedback*. Os dados de treinamento criam o *background* do modelo, numa fase anterior à colocação dele em funcionamento; os dados de entrada, já na fase de aplicação, sinalizam para o modelo o que está no ambiente externo, e os dados de saída são o resultado do processo decisório artificial. Os dados de saída poderão voltar à máquina, como *feedback* positivo ou negativo, para que ela possa se autoajustar ou ser ajustada pelo desenvolvedor.

Somente quando os dados de entrada são *dados pessoais* ou quando o julgamento diz respeito a alguma pessoa natural (caso em que os dados de saída são dados pessoais), é que se pode falar em "decisão automatizada", no direito brasileiro, conforme se extrai do art. 20 da LGPD:

---

[164] Como se verá adiante, existem métodos de aprendizado de máquina em que, embora as regras iniciais sejam dadas pelo programador, o modelo pode autonomamente ponderar os pesos dos dados, a partir de exemplos que lhe são apresentados, alterando a programação inicial.

[165] ERTEL, Wolfgang. **Introduction to Artificial Intelligence**. Cham (Switzerland): Springer, 2017. Tradução de Nathanael T. Black, p.3

> Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade (...)

Ora, se em toda decisão automatizada o titular dos dados (de entrada ou de saída) tem direito de solicitar a revisão, então sempre haverá um titular em tais casos; logo, sempre estão em jogo dados pessoais nas decisões automatizadas, pois o titular é a "pessoa natural a quem se referem os dados pessoais que são objeto de tratamento" (LGPD, art. 5º, V).

De fato, há muitos processos automatizados na indústria ou na pesquisa científica que, no entanto, não produzem "decisões", no sentido empregado pela legislação brasileira. Em uma pesquisa científica sobre uma bactéria, por exemplo, pode-se usar processos automatizados para avaliar e prever aspectos ou comportamentos dessa forma de vida, sem que se possa falar, no entanto, em "decisão automatizada", na acepção jurídica da expressão. O mesmo pode ocorrer numa fábrica de parafusos que automatize os processos de avaliação da qualidade de seus produtos: isso não gera decisões automatizadas, no sentido empregado pela LGPD.

A LGPD não chega a definir o que seja decisão automatizada, mas a ela se refere para assegurar ao titular de dados pessoais o *direito à revisão* dessa decisão, bem como o *direito à explicação* sobre os processos e os dados utilizados na formulação da decisão, nos seguintes termos:

> Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade. (Redação dada pela Lei nº 13.853, de 2019)
> § 1º O controlador deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, observados os segredos comercial e industrial.
> § 2º Em caso de não oferecimento de informações de que trata o § 1º deste artigo baseado na observância de segredo comercial e industrial, a autoridade nacional poderá realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais.

Adiante analisa-se cada um dos elementos normativos utilizados para a composição de uma definição de decisão automatizada.

3.1.1 Uso de dados pessoais

O dispositivo legal referido estipula alguns elementos que permitem inferir o conceito de decisão automatizada, para os fins da LGPD. Em primeiro lugar, é preciso que a decisão tenha sido tomada mediante o uso de dados pessoais, visto que a lei fala de direitos do "titular" a respeito dessa decisão. E "titular" tem uma definição precisa na LGPD, a saber: é a "pessoa

natural a quem se referem os dados pessoais que são objeto de tratamento" (LGPD, art. 5º, V). Dados pessoais, por sua vez, são aqueles que produzam informações relacionadas a pessoa natural identificada ou identificável (LGPD, art. 5º, I).

Logo, como referido acima, processos de automatização adotados em atividades que não envolvam dados pessoais, não estão abrangidos pela disciplina da LGPD. Um caso particularmente interessante é o da pessoa jurídica. Os dados relativos a pessoas jurídicas não são dados pessoais, de modo que o tratamento automatizado de dados relacionados às pessoas jurídicas não estão no raio de incidência da norma da LGPD que assegura os direitos de revisão e de explicação — embora não fique excluída a hipótese de se buscar tais direitos, sobretudo em casos de assimetria negocial, por aplicação analógica do Código de Defesa do Consumidor, ou de alguma normativa protetiva específica, ou até mesmo por aplicação direta da Constituição, com base na ideia mais geral de proteção de dados como direito fundamental extensível também às pessoas jurídicas.

Outra questão que pode ser levantada aqui é dos dados anonimizados. Eles não são considerados dados pessoais pela LGPD (art. 12), salvo quando o processo de anonimização ao qual foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser revertido. Todavia, a agregação de dados pessoais com posterior anonimização para a criação de modelos preditivos de comportamento humano individual parece estar dentro da disciplina das decisões automatizadas, sobretudo quando venham a afetar algum interesse individual ou coletivo, pois nesses casos os dados de saída serão pessoais.

Assim é que, por exemplo, o autopreenchimento dos *sites* de busca é modelado a partir de um grande número de pesquisas individuais. Mesmo que a anonimização dos dados que deram base para a formulação do modelo retire o caráter pessoal desses dados, é certo que a decisão automatizada de preenchimento pode vir a trazer danos individuais ou coletivos e, por isso, está sujeito à disciplina do art. 20 da LGPD. É o que ocorre, por exemplo, quando o autopreenchimento se refere a alguma pessoa natural específica. Neste caso, o nome da pessoa é um dado pessoal e a decisão automatizada, ao ligar esse nome a um fato, a uma característica, a uma imagem, enfim a uma informação, produz conhecimento com dados pessoais do interessado.

Há inúmeros exemplos de precedentes, em vários países, sobre a questão do autopreenchimento pelos motores de busca na internet, notadamente o *Google*. Um tribunal em Milão obrigou o *Google* a rever o autopreenchimento de pesquisa que associava

automaticamente o nome de uma pessoa, quando pesquisada, à palavra "vigarista"[166]. No Japão, a mesma empresa foi obrigada a excluir um autopreenchimento que associava o nome de um indivíduo a crimes cometidos por um homônimo[167]. Em 2013, na Alemanha, um tribunal federal foi mais longe e obrigou o *Google* a eliminar todos os autopreenchimentos difamatórios, quando provocado pelo respectivo interessado[168].

3.1.2 Tratamento automatizado

O tratamento de dados por mecanismos eletrônicos (digitais) está no cerne da concepção de decisões automatizadas. É por meio do Aprendizado de Máquina (*Machine Learning*), o tipo de programação mais usado em aplicações práticas, que dados pessoais podem ser transformados em informações e em conhecimento, por dispositivos que funcionam de forma autônoma, mediante associações, agregações e desagregações, arranjos e rearranjos de dados; análises de padrões em vastos conjuntos de dados; inferências estatísticas e estimativas probabilísticas — enfim, técnicas matemáticas convenientes para extrair conhecimentos de dados, mimetizando o funcionamento da inteligência humana, ou, pelo menos, a parte computável da inteligência humana.

Com efeito, a LGPD, para esboçar a ideia de decisão automatizada, estabelece que tal é aquela que tenha sido tomada "unicamente com base em tratamento automatizado" (art.20, LGPD). Assim, a lei parece buscar excluir de seu raio de eficácia tanto as decisões decorrentes diretamente da inteligência humana, como as decisões humanas assistidas por processos automatizados, que não devem ser consideradas decisões automatizadas, segundo a lei brasileira.

Nesta altura, vale lembrar a interessante discussão travada nos Estados Unidos caso Loomis *x* Winsconsin. Em fevereiro de 2013, Eric Loomis foi preso por dirigir um carro roubado e por fugir de uma barreira policial em La Crosse (Wisconsin). Após o regular processamento da acusação, ele foi condenado pelo juiz local a uma pena de 6 anos de prisão. A sentença, ademais, negou a liberação condicional do condenado, sob o argumento, entre outras coisas, de que o COMPAS (*Correctional Offender Management Profiling for Alternative*

---

[166] A íntegra de decisão pode ser lida em: MONTI, Andrea. Tribunale di Milano: Ord. 24 marzo 2011. *In*: MONTI, Andrea. **ICT LEX**: Diritto, politica, cultura della Rete. [S. l.], 24 mar. 2011. Disponível em: https://www.ictlex.net/?p=1285. Acesso em: 7 jan. 2021.

[167] Cf.: https://www.bbc.com/news/technology-17510651, Acesso em: 7 jan. 2021.

[168] AMBROSE, Meg Leta; AMBROSE, Ben M.. When robots lie a comparison of auto-defamation law. **2014 Ieee International Workshop On Advanced Robotics And Its Social Impacts**, [S.L.], p. 56-61, set. 2014. IEEE. http://dx.doi.org/10.1109/arso.2014.7020980.

Sanctions), um modelo utilizado pelo Judiciário de Wisconsin para calcular o risco de reincidência dos acusados, apontava alto grau de periculosidade em Eric Loomis.

A defesa de Loomis apresentou recurso contra essa condenação, alegando que não se sabia exatamente de que maneira o COMPAS funcionava, e que os seus fabricantes naturalmente não iriam revelar, porque nesse sigilo residiria justamente o valor econômico do produto. Assim, o uso desse tipo de ferramenta, segundo a defesa, violaria o devido processo legal, especialmente o direito de ser sentenciado de forma fundamentada e sem o uso de fatores inverificáveis.

A Suprema Corte de Wisconsin rejeitou a apelação[169], sob o argumento de que o juiz não decidira unicamente com base no tratamento automatizado de dados, mas sim também com base em todo o contexto probatório. A Suprema Corte Americana, para a qual posteriormente foi dirigido um pedido de *writ of certiorari*, rejeitou o julgamento do mérito da questão[170].

Observa-se que a posição do Judiciário americano, nesse caso, tolerando o uso do tratamento automatizado de dados em um tema tão sensível como é a decisão sobre a liberdade de locomoção, apoiou-se no fato de que a deliberação, em última análise, não foi da máquina, mas sim do humano (o juiz) que apreciou o pedido de liberdade condicional, embora ele possa ter levado em conta a predição do modelo, que indicava alto risco de reincidência.

O problema do grau de contribuição humana para a decisão tende a ser geralmente o de mais difícil abordagem, quando se trata de delimitar o alcance da LGPD na questão das decisões automatizadas. O uso do advérbio "unicamente" parece sugerir que qualquer mínima intervenção humana no processo decisório descaracteriza a decisão como sendo automatizada. Isso porque se a decisão tem intervenção humana, qualquer que seja ela, não é possível calcular quanto dessa decisão decorreu de contribuição da máquina, de modo que, pela lei brasileira, tal decisão não é automatizada.

É certo que a decisão automatizada apenas se torna possível mediante ações humanas anteriores, de programadores, investidores, cientistas de dados, engenheiros, matemáticos, etc.. No entanto, chega um ponto em que o modelo pode funcionar autonomamente, produzindo deliberações de acordo com o seu modo de funcionamento ordinário, mediante a combinação de dados de entrada segundo um procedimento criado total ou parcialmente por programadores. É neste ponto que a intervenção humana pode descaracterizar a decisão como automatizada.

---

[169] Cf.: https://www.wicourts.gov/sc/opinion/DisplayDocument.pdf?content=pdf&seqNo=171690. Acesso em: 8 jan. 2021.

[170] Loomis v. Wisconsin, 881 N.W.2d 749 (Wis. 2016), cert. negado, 137 S.Ct. 2290 (2017).

Se a máquina apenas assiste o humano, fornecendo-lhe elementos para avaliar as melhores alternativas, cabendo a escolha do resultado ao humano, isso não pode ser definido como decisão automatizada, segundo a LGPD. Por outro lado, se o humano apenas ratifica a decisão da máquina, sem possibilidade de criticá-la ou descartá-la, então a decisão é automatizada, apesar de eventualmente ser assinada por um ser humano. Neste último caso, ocorre aquilo que se se chama de *rubber-stamping*[171], ou seja, um mero carimbo do ser humano.

A Autoridade Independente de Dados do Reino Unido[172] e o Conselho Europeu de Proteção de Dados[173] publicaram algumas orientações elucidativas sobre a questão da intervenção humana como causa da descaracterização da decisão como automatizada. Tais orientações podem ser resumidas no seguinte:

a) Os revisores humanos devem estar envolvidos na verificação da recomendação do sistema e não devem apenas "rotineiramente" aplicar a decisão automatizada (o envolvimento dos revisores deve ser ativo e não apenas simbólico);

b) Os revisores humanos devem ter uma influência "significativa" (*meaningful*) na decisão automatizada, inclusive com autoridade e competência para ir contra ela;

c) Os revisores humanos devem "pesar" e "interpretar" a predição da máquina, considerando todos os dados de entrada disponíveis e outros fatores adicionais.

Dois exemplos, citados nas orientações da Autoridade de Proteção de Dados do Reino Unido[174], podem esclarecer a diferença entre decisão automatizada e decisão humana assistida por processos automatizados: 1º) Pense-se numa fábrica que calcula e paga o valor de uma gratificação dos empregados conforme a sua produtividade, apurada por mecanismos automatizados e sem qualquer intervenção humana ou com intervenção humana meramente homologatória — isso é uma decisão automatizada; 2º) agora pense-se numa fábrica que use mecanismos automatizados para avaliar a pontualidade dos empregados, disparando um aviso

---

[171] BINNS, Reuben; GALLO, Valeria. **Automated Decision Making**: the role of meaningful human reviews. In: ICO. Information Commissioner's Office. 12 abr. 2019. Disponível em: https://ico.org.uk/about-the-ico/news-and-events/ai-blog-automated-decision-making-the-role-of-meaningful-human-reviews/. Acesso em: 11 jan. 2021.

[172] ICO. What does the UK GDPR say about automated decision-making and profiling?. *In*: ICO. **Information Commissioner's Office**. Disponível em: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-does-the-uk-gdpr-say-about-automated-decision-making-and-profiling/. Acesso em: 11 jan. 2021.

[173] JUSTICE AND CONSUMERS (Europea Union). Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. *In*: JUSTICE AND CONSUMERS (Europea Union). **European Commission**. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 11 jan. 2021

[174] ICO, op. cit.

a um gerente de recursos humanos sempre que algum empregado, segundo apuração automatizada de dados, chega atrasado mais de tantas vezes — isso não é decisão automatizada, pois a máquina apenas prediz a situação (a falta de pontualidade) e comunica ao ser humano responsável, para que decida e adote a ação adequada.

4.1.2.1 Tratamentos automatizados excluídos do alcance da LGPD (tratamentos domésticos, jornalísticos, artísticos e acadêmicos)

Conforme o art. 4º, I, II e III da LGPD, para além dos casos de extraterritorialidade, não estão sob a proteção da lei especial brasileira os tratamentos de dados que sejam realizados: a) por pessoa natural para fins exclusivamente particulares e não econômicos; b) para fins exclusivamente jornalísticos ou artísticos; c) para fins acadêmicos, observado o disposto nos arts. 7º a 11 da LGPD.

Significa isso dizer que eventual decisão automatizada tomada com os objetivos acima expostos não está sujeita às restrições da LGPD, notadamente as previstas no art. 20. Tal conclusão decorre logicamente do fato de não ser o tratamento de dados, nesses casos, protegido pela LGPD. Está claro, todavia, que eventual violação a direito individual em tais circunstâncias, especialmente à privacidade ou à imagem do titular de dados pessoais, não deve ficar sem meios de reparação, podendo ser corrigida por instrumentos atípicos mediante a aplicação direta da Constituição, sobretudo por força da cláusula do devido processo legal (CF, art. 5º, LIV).

Afinal, o pressuposto da lei para excluir essas decisões da sua disciplina é de que elas são presumivelmente inofensivas a direitos de terceiros, ou estão albergadas pela liberdade de expressão, ou pela liberdade de investigação científica, de modo que, se for alegado e comprovado dano, ameaça de dano por abuso dessas liberdades, há de existir proteção legal contra a violação, ainda que apenas judicial (CF, art. 5º, XXXV).

4.1.2.2 Tratamento automatizado regulado subsidiariamente pela LGPD (segurança pública, defesa nacional, segurança do Estado e atividades de investigação e repressão de infrações penais)

Os tratamentos de dados para fins de exclusivos de segurança pública, defesa nacional, segurança do Estado e atividades de investigação e repressão de infrações penais (LGPD, art. 4º, III), embora sujeitos a futura legislação específica (LGPD, art. 4º, §1º), deverão até lá prever medidas proporcionais e estritamente necessárias ao atendimento do interesse público, observados o devido processo legal (CF, art. 5º, LIV), os princípios gerais de proteção (LGPD, arts. 2º e 6º) e os direitos do titular previstos na própria LGPD (arts. 17 a 22).

Cabe à Agência Nacional de Proteção de Dados – ANPD um papel preponderante de regulamentação e fiscalização, na falta de lei específica, dos tratamentos de dados não inteiramente sujeitos à LGPD, como é a hipótese daqueles relacionados à segurança pública e à defesa nacional. Nesses casos, deverá a ANPD emitir opiniões técnicas ou recomendações, e deverá solicitar aos responsáveis relatórios de impacto à proteção de dados pessoais (LGPD, art. 4º, §3º).

Um ponto de grande interesse na disciplina do tratamento de dados para fins exclusivos de segurança pública, defesa nacional, segurança do Estado e atividades de investigação e repressão de infrações penais (LGPD, art. 4º, III), é que pessoas de direito privado não podem realizar esses tratamentos (LGPD, art. 4º, §2º) — exceto em procedimentos sob tutela de pessoa jurídica de direito público, que serão objeto de informe específico à autoridade nacional —, o que está em conformidade com o disposto nos arts. 142 e 144 da Constituição Federal, que atribuem com exclusividade às Forças Armadas e às Polícias Federal, Rodoviária Federal, Ferroviária Federal, Civis, Militares e Penais, a competência para as atividades de segurança externa e interna do país.

Enquanto não advém a legislação específica disciplinando o tratamento de dados para fins de segurança pública e atividades de investigação, o que se observa, pela remissão ampla feita pelo art. 4º, §1º da LGPD, é que eventuais decisões automatizadas tomadas nesse campo estarão sujeitas juridicamente ao disposto no art. 20 da LGPD, além de também deverem atender aos princípios gerais de proteção e ao devido processo legal.

A questão peculiar no tratamento de dados pessoais para fins de segurança e apuração criminal é que, como se sabe, há uma larga tradição jurídica, tanto legislativa quanto jurisprudencial, construída na era analógica, que abria exceções importantes à privacidade quando se cuidava de medidas investigativas requeridas judicialmente por autoridades policiais ou de segurança em geral, independentemente do consentimento do titular e, em alguns casos, até mesmo de sua ciência.

Assim, a proteção à privacidade se dava pela oposição de obstáculos ao acesso às informações íntimas do cidadão (por exemplo: sigilo bancário, sigilo fiscal, sigilo profissional, inviolabilidade de domicílio [175]); obstáculos esses que, excepcionalmente, poderiam ser afastados, com certas reservas procedimentais, a pedido de autoridades policiais. Entretanto, na era digital, esse tipo de garantia torna-se em certos aspectos anacrônica, porque o indivíduo já

[175] Lei Complementar 105/2001; Lei Complementar 104/2001; Lei 5.172/1966 (Código Tributário Nacional); Decreto-lei 2.848/1940 (Código Penal); Constituição Federal, art. 5º, XI.

não governa seus dados, que estão dispersos e profusos em muitos bancos de dados espalhados pela internet; e esses dados podem ser entrecruzados, por mecanismos de inferência apropriados, permitindo a prospecção indireta de informações sobre o indivíduo sem a necessidade de quebra de sigilos.

Nesse contexto, sem prejuízo dos sigilos tradicionais, é fundamental regulamentar a forma como a autoridade policial pode coletar dados pessoais ou reorientar dados já coletados para outros propósitos; como pode tratar esses dados; como pode correlacioná-los com outros, partindo já do pressuposto de que o acesso aos dados não sigilosos pode, indiretamente, levar ao conhecimento de informações sigilosas.

Jacqueline de Sousa Abreu, a esse propósito, faz as seguintes considerações:

> Se o direito à privacidade servia à proteção de escolhas e espaços individuais para realização de intimidade, o direito à proteção de dados pessoais emerge como uma ampla estrutura de proteção regulatória, em atenção a novas formas de danos e riscos a que cidadãos estão expostos. Está assentado na constatação de que a sociedade da informação expõe o indivíduo a diversos riscos de dano físico, material ou moral que comprometem o exercício de sua autonomia, a níveis individual e coletivo. Tais riscos são decorrentes de práticas e/ou estruturas institucionais que se desviam de noções básicas de justiça: ter uma expectativa legítima de respeito e consideração frustrada em suas relações sociais com empresas e com o Estado (pelo uso inesperado de suas informações, pela falta de segurança razoável dispensada a suas informações, pelo uso discriminatório, para dar alguns exemplos), e não possuir instrumentos de remediação, por exemplo."[176]

Constata-se aqui uma premissa que é constante na proteção de dados no ecossistema digital: as possibilidades de produção de informação a partir de dados são tantas e tão diversificadas que é mais conveniente regulamentar a forma como elas podem ser legitimamente implementadas do que tentar usar critérios materiais proibitivos, que sempre foi a técnica mais usada no mundo analógico.

Pequenos fragmentos de informação sobre o investigado, indícios quase desprezíveis, quando devidamente tratados e colocados em contato com grandes volumes de dados pessoais até mesmo de outras pessoas, colhidos muitas vezes para fins inocentes, podem ter um poder revelador insuspeitado. A importância desses fragmentos e indícios acaba se revelando *a posteriori*, em razão das ferramentas de mineração de dados (*data mining*), e não exatamente da matéria de que tratam. Por isso a técnica de isolar e tutelar mais fortemente certos tipos de

[176] ABREU, Jacqueline de Souza. Tratado de Proteção deTratamento de Dados Pessoais para Segurança Pública: contornos do regime jurídico pós-LGPD. In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle, p. 592-593.

dados pode não ser suficiente para a adequada proteção de dados pessoais no campo das investigações criminais e da segurança pública em geral.

3.1.3 Ameaça ou lesão a interesse juridicamente tutelado

Outro elemento integrante do conceito legal brasileiro de decisão automatizada, constante do art. 20 da LGPD, com forte inspiração na GDPR, está na necessidade de que a deliberação de máquina ameace ou atinja um interesse juridicamente protegido.

Assim, qualquer demanda, judicial ou extrajudicial, contra o controlador que produza decisões automatizadas está na dependência de que o titular dos dados pessoais alegue e prove a ameaça ou violação, pela decisão automatizada, de algum interesse próprio que tenha a tutela do direito. Depreende-se a contrario sensu que decisões automatizadas inofensivas a direitos individuais ou coletivos não estão sob a tutela da lei — como, de resto, ocorre em qualquer área do direito em relação a atos abnóxios, que recaem no campo da licitude.

Os interesses violáveis por decisões automatizadas são os mais diversos, tais como, por exemplo: liberdade de expressão, numa rede social que use algoritmos para moderar publicações ou filtros de *upload*[177]; imagem, num site de busca que associe automaticamente o nome de uma pessoa natural a uma notícia falsa; direitos autorais, numa rede que publique livremente os conteúdos carregados pelos usuários[178]; patrimônio e imagem, num site de compras que manipule automaticamente preços e ofertas, discriminando pessoas pela localização de sua residência, etc.

Todos esses interesses, quando violados dentro do contexto de processos automáticos de tratamento de dados, podem ser reconduzidos à esfera tutelada pelo direito à proteção de dados e suas manifestações especiais e instrumentais previstas na LGPD.

Embora já existam questões relativamente conhecidas nesse campo das decisões automatizadas, tais como aquelas associadas à discriminação algorítmica, não é possível antecipar todas as possíveis ofensas a interesses protegidos que são suscetíveis de ocorrer por força do tratamento automático de dados pessoais. A maior parte da responsabilidade nesse campo é atípica e centrada mais nos danos que nas condutas, como acentuado no capítulo anterior.

---

[177] SCHILLER, Arnold; WEISKOPF, Tobias. Automated Censorship in the Digital Space. *In*: YOUNG EUROPEAN FEDERALISTS (Europe). **The New Federalist**, 1 maio 2019. Tradução de Nora Teuma. Disponível em: https://www.thenewfederalist.eu/automated-censorship-in-the-digital-space?lang=fr. Acesso em: 9 dez. 2020.

[178] BREEN, Jason. YouTube or YouLose? Can YouTube Survive a Copyright Infringement Lawsuit. **Bepress Legal Series.Working Paper 1950**, Los Angeles, p. 1-37, 18 jan. 2007. Disponível em: https://law.bepress.com/cgi/viewcontent.cgi?article=9209&context=expresso. Acesso em: 10 dez. 2020.

O art. 44, parágrafo único, da LGPD, bem enfatiza que a responsabilidade por tratamento irregular de dados nasce do dano, quando o agente de tratamento não observa as normas de segurança previstas no art. 46 da LGPD.

Sem dano ou ameaça de dano, não há responsabilidade, porquanto não há o que reparar ou assegurar. Assim, o critério inicial para avaliar a presença de uma situação em que a decisão automatizada pode ser questionada ou mesmo anulada, com base em algum direito do titular, é o dano ou o potencial de dano que ela pode causar a interesse juridicamente protegido. Este é um critério de ordem pragmática que está na essência da própria ideia de direito subjetivo. Como explica Manuel A. Domingues de Andrade,

> De toda maneira, onde há um direito subjectivo, ele foi concedido para que através dele fosse obtido o predomínio de certo interesse; tal como a correspondente obrigação ou sujeição foi imposta para que um outro interesse oposto resultasse subordinado àquele.
> Mas uma coisa é o direito subjectivo em si mesmo e outra coisa é a razão por que, ou o fim em vista do qual, a lei atribui esse direito, ou seja o interesse para cuja prevalência tal direito foi concedido.
> O interesse constitui o substrato do direito subjectivo. É-lhe subjacente; está antes dele. Ou então — se assim se prefere — está para além dele. Em todo caso, está fora dele. Não diz respeito à sua estrutura, mas só à sua função. Não tem que entrar, portanto, na definição do respectivo conceito.[179]

O interesse está, consequentemente, no cerne da função de proteção jurídica. É por meio do interesse que, antes de tudo, se pode avaliar a necessidade e a utilidade de mecanismos jurídicos de tutela contra as decisões automatizadas. Se algum interesse juridicamente protegido for violado ou ameaçado pela decisão automatizada, há, quando menos, o direito de questionar em juízo o ato da máquina, por força da garantia do direito de ação (CF, art. 5º, XXXV). Adicionalmente, pode-se invocar os direitos consagrados na LGPD, e, conforme o caso, no CDC, no CC, na Lei do Cadastro Positivo (Lei 12.414/2011), na Lei de Acesso à Informação (Lei 12.527/2011) e em qualquer outra legislação, inclusive tratados, que, mesmo pensados para relações do mundo analógico, possam ser aplicados por semelhança ao contexto digital, conforme determina o art. 64 da LGPD.

3.1.4 Definição

Baseado nas premissas acima apresentadas, pode-se construir uma definição que expresse objetivamente em que consiste uma decisão automatizada no contexto da Lei Geral de Proteção de Dados – LGPD.

---

[179] ANDRADE, Manuel A. Domingues. **Teoria geral da relação jurídica**. Coimbra: Almedina, 1992. v. 1, p. 8.

Decisão automatizada é todo julgamento feito exclusivamente por máquina, com base em predição decorrente de tratamento automatizado de dados pessoais de entrada, segundo um modelo ou algoritmo condicionado por dados de treinamento, que afete imediatamente interesse juridicamente tutelado de pessoa natural, excetuados aqueles que tenham fins particulares e não econômicos, jornalísticos ou científicos.

Em adição, há dois tipos de julgamento que podem ser classificados como decisões automatizadas por equiparação: 1º) aqueles que, satisfazendo as condições referidas acima, recebam intervenção humana meramente homologatória (*rubber-stamping*); e 2º) as perfilizações automáticas, conforme se verá adiante.

No núcleo do processo de decisão automatizada estão técnicas estatísticas que permitem a extrapolação de informações, a partir de amostras de dados de uma população. Evidentemente, essas técnicas estão sujeitas a erros e desvios típicos do campo estatístico, embora no geral sejam confiáveis como processo de inferência e predição[180]. Como dados pessoais são indispensáveis para qualquer tipo de decisão automatizada, dentro do contexto da legislação brasileira, a construção de perfis individuais aparece sempre associada a qualquer julgamento feito por máquina e foi equiparada, por lei, à decisão automatizada.

3.2 Perfilização

A perfilização está tão intimamente ligada às decisões automatizadas que a LGPD (art. 20) a inclui no próprio conceito destas:

> Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, *incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade.*

Na verdade, porém, a perfilização está mais associada à predição e somente pode ser considerada a decisão automatizada se ela mesma for o objetivo do modelo ou algoritmo. Caso se queira, por exemplo, avaliar a capacidade de pagamento de alguém para efeito de concessão de um empréstimo, a perfilização será parte do tratamento de dados e da predição, mas a decisão não estará nisso, e sim na concessão ou não do empréstimo. A decisão é sempre uma tomada de posição diante dos dados, e não apenas uma inferência estatística. A predição, que decorre das inferências estatísticas, apontará o provável resultado da operação de empréstimo (digamos, há 80% de chance de o indivíduo pagar o empréstimo dentro do prazo); já a decisão estará em definir o titular dos dados como apto ou não para o empréstimo. Por exemplo, certa instituição

[180]KUBAT, Miroslav. **An Introduction to Machine Learning**. 2. ed. Coral Gables: Springer, 2017, p. 231.

financeira pode decidir pelo sim, com 80% de chance de pagamento, mas outra pode exigir um limiar de predição maior (digamos, 90%) para contratar o empréstimo. Portanto, a predição não é ainda a decisão; ela é o prenúncio do que provavelmente ocorrerá, caso a decisão seja tomada em um ou outro sentido, à luz dos dados tratados pelo modelo. A preferência por acolher essa probabilidade como um "sim" ou um "não" é que a decisão.

É difícil pensar a decisão automatizada sem algum grau de perfilização. Visto como os dados pessoais, por definição, sempre estão associados a alguma pessoa natural e devem fazer parte do processo de formação da decisão automatizada, como exposto acima; considerando também que o objetivo prático dessas decisões sempre está de algum modo associado à compreensão de características ou do comportamento pretérito de pessoas naturais, para avaliar as suas características ou seus comportamentos futuros, então algum grau de perfilização quase sempre está na base das decisões automatizadas.

Em certos casos, todavia, pode ocorrer decisão automatizada sem perfilização. A Autoridade Independente do Reino Unido menciona, a esse respeito, o caso de uma correção de prova automatizada[181]. Uma banca examinadora pode usar um sistema automatizado para marcar as folhas de respostas de um exame de múltipla escolha. O sistema é pré-programado com o número de respostas corretas necessárias para alcançar marcas de aprovação e distinção. As pontuações são automaticamente atribuídas aos candidatos com base no número de respostas corretas de cada um e os resultados estão disponíveis *online*. Trata-se de um processo automatizado de tomada de decisão que não envolve criação de perfil. Mas isso apenas ocorre em situações pontuais, que não busquem utilizar o modelo reiteradamente para o futuro, como essa cogitada, e não representa o coração das aplicações de processos automatizados nos processos produtivos.

O Regulamento Europeu para a Proteção de Dados (GDPR) define a perfilização (ou "definição de perfil", na tradução portuguesa) como algo diferente da decisão automatizada, embora não seja totalmente fiel a essa distinção em outros pontos. Com efeito, o art. 4º, n. 4 do GDPR associa a perfilização com a análise e a predição, que são anteriores à decisão, *verbis*:

> «Definição de perfis», qualquer forma de tratamento automatizado de dados pessoais que consista em utilizar esses dados pessoais para avaliar certos aspetos pessoais de uma pessoa singular, nomeadamente para analisar ou prever aspetos

[181] ICO. What is automated individual decision-making and profiling?. *In*: ICO. **Information Commissioner's Office**. Disponível em: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-is-automated-individual-decision-making-and-profiling/#:~:text=Automated%20decision%2Dmaking%20is%20the,to%20award%20a%20loan%3B%20and. Acesso em: 27 jan. 2021

> relacionados com o seu desempenho profissional, a sua situação econômica, saúde, preferências pessoais, interesses, fiabilidade, comportamento, localização ou deslocações.

Já em relação à decisão automatizada, o art. 22, n. 1 do GDPR (que, no ponto, foi praticamente copiado pela LGPD) estipula, *verbis*:

> O titular dos dados tem o direito de não ficar sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar.

A associação da perfilização com as decisões automatizadas decorre da circunstância de que, como visto, somente são consideradas automatizadas decisões que utilizem dados pessoais em seu processo de concepção. Como os dados pessoais, por definição, somente são aqueles referentes a uma pessoa natural, então o modelo capaz de produzir decisão automatizada sempre terá dados referentes a alguma pessoa natural como dados de entrada, daí porque a predição que ele fará resultará no prognóstico sobre alguma característica ou comportamento humano, baseado em características ou comportamentos anteriores da mesma ou de outras pessoas naturais que apresentem certo padrão reconhecido pela máquina. A decisão automatizada será baseada nessa predição, por isso ela de alguma maneira está conectada ao perfil decorrente dos dados de entrada.

Assim, um modelo que crie decisões automatizadas para admitir ou negar a entrada de pessoas numa universidade será previamente alimentado com um vasto conjunto de dados anteriores, sobre a admissão e a rejeição de candidatos (dados pessoais, portanto). Matematicamente, o modelo inferirá padrões desse conjunto de dados pessoais: tanto padrões para os que devem ser admitidos, como para os que devem ser rejeitados. Tão logo sejam inseridos os dados de interesse de um novo candidato (local de residência, notas, renda mensal, idade, enfim o conjunto de dados pessoais do postulante à vaga), o modelo predirá se o caso, à luz dos anteriores, é de admissão ou de rejeição; e a decisão de admitir ou rejeitar será tomada com base no grau da predição. É evidente que, em tal contexto, o novo candidato estará sendo perfilizado pelo modelo, embora não seja a perfilização propriamente o objetivo do tratamento de dados; ela é, na verdade, uma etapa para a construção da decisão — seguramente uma etapa muito relevante.

O mesmo ocorrerá em um modelo de previsão de fraudes bancárias, ou de cotação de preços de mercadorias com base nos dados do pretenso comprador, ou em um mecanismo policial ou alfandegário que decida automaticamente quem deve ser fiscalizado preferencialmente. Sempre haverá a concepção de tipos genéricos que serão comparados aos

dados pessoais dos sujeitos de interesse, para a solução do problema de negócio. Logo, a perfilização é um passo necessário para a tomada de decisões automatizadas, mas não é a própria decisão automatizada.

As decisões automatizadas podem ser realizadas com ou sem definição de perfis; a definição de perfis pode ocorrer sem que dela decorra uma decisão automatizada. Todavia, a definição de perfis e as decisões automatizadas não constituem necessariamente atividades separadas. Um procedimento iniciado como um processo de decisão automatizada pode tornar-se um procedimento de definição de perfis, dependendo da forma como os dados sejam utilizados.[182]

Em outras circunstâncias, a decisão automatizada pode ou não depender de perfilização, segundo o interesse do desenvolvedor na concepção do modelo. Assim, um sistema automatizado de imposição de multas de trânsito, a partir de imagens de câmeras de monitoramento espalhadas nas vias públicas, pode não levar em conta nenhum fator particular do infrator — nesse caso, portanto, desprezando a perfilização. Mas o mesmo modelo pode ser incrementado, para incluir características específicas do infrator (tempo de habilitação, multas anteriores, profissão, etc.), de modo a calibrar o valor da multa. Nesse caso, a perfilização estaria presente na composição da decisão automatizada.[183]

Pela redação da LGPD, no entanto, deve-se admitir que a perfilização, mesmo quando não seja seguida de uma decisão automatizada, mas sim de uma decisão humana assistida por máquina, deve ser considerada em si mesma uma decisão automatizada por equiparação, já que a lei afirma expressamente que estão incluídas entre as decisões automatizadas aquelas "decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade" (artigo 20, LGPD).

Historicamente, a perfilização antecede o uso massivo de processos automatizados de coleta e tratamento de dados. Já nos anos 1980, falava-se do processo de crescente perfilização em várias áreas, especialmente no campo criminal e no âmbito do marketing direcionado[184]. O

---

[182] Cf.: JUSTICE AND CONSUMERS (Europea Union). Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. In: **JUSTICE AND CONSUMERS** (Europea Union). European Commission. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 11 jan. 2021

[183] O exemplo é dado, com pequenas alterações, na página 7 do Guia de Orientações já citado: Cf.: JUSTICE AND CONSUMERS (Europea Union). Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. In: **JUSTICE AND CONSUMERS** (Europea Union). European Commission. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 11 jan. 2021

[184] CLARKE, Roger. Profiling: a hidden challenge to the regulation of data surveillance. **Journal Of Law, Information And Science**, v. 2, n. 4, jan. 1993. Disponível em: http://www.austlii.edu.au/cgi-bin/sinodisp/au/journals/JlLawInfoSci/1993/26.html?query=. Acesso em: 10 dez. 2020.

método de construção de perfis é notoriamente suscetível às técnicas que estão na base dos processos de aprendizado de máquina, daí porque a coleta massiva e o tratamento automatizado de dados pessoais naturalmente implicaram um processo exponencial de perfilização.

De fato, a perfilização, conforme Roger Clarke[185], é uma técnica por meio da qual um conjunto de características de um grupo particular de pessoas é inferido a partir de experiências passadas (das mesmas pessoas ou de pessoas com comportamento assemelhado), de modo tal a formar acervos que podem ser comparados com indivíduos no futuro, para avaliar o quanto estes se ajustam às características típicas do grupo. Bem analisada, a ideia de perfilização, em termos de método para conhecer objetivamente a mecânica dos comportamentos humanos, pode mesmo remontar ao conceito de "tipo ideal", de Max Weber[186], pois na sociologia o estudo de padrões de comportamentos sociais a partir da junção e organização de fragmentos esparsos de condutas individuais e de grupo é uma ferramenta há muito utilizada.

A metodologia matemática, na qual está a essência das técnicas de aprendizado de máquina, usa frequentemente o processo de reunião de objetos por características comuns (conjuntos), para inferir as relações de pertinências ou não de outros objetos. A perfilização por mecanismos automatizados é fundamentalmente um procedimento matemático de coleta, seleção, agrupamento e comparação de dados pessoais.

3.3 Benefícios

O que leva as empresas e os governos a automatizarem os seus processos decisórios é, sem dúvida, o aumento da capacidade e da velocidade de resposta a demandas repetitivas e a redução de custos que isso proporciona. Por isso mesmo, decisões políticas ou que contenham elementos discricionários ou de estratégia negocial normalmente permanecem sob a governança estritamente humana, embora possam ser subsidiadas por tratamentos automatizados de dados.

Decisões tomadas em massa, com certo padrão, traduzíveis em termos matemáticos, tais como preços de mercadorias, contratos de empréstimos e análises de risco, são particularmente suscetíveis ao processo de automatização, desde que se tenha um conjunto relevante de dados que permita construir um modelo replicador das decisões anteriores.

---

[185] Op.cit.

[186] WEBER, Max. A objetividade do conhecimento nas ciências sociais. In: FERNANDES, Florestan (org.). **Weber**: sociologia. São Paulo: Ática, 1999. Coleção Grandes Cientistas Sociais, p. 79-123.

O aprendizado de máquina busca imitar a racionalidade humana, a qual, por sua vez, está baseada na observação e organização intelectual do mundo, segundo padrões prévios, para predizer o futuro.

A automatização é um processo fundamentalmente estatístico-matemático: desvendam-se padrões nos dados e, a partir disso, a máquina "aprende" a reconhecê-los e a associá-los às "decisões corretas" respectivas. Cria-se, em suma, uma conexão lógica entre os dados de entrada e a decisão desejável para um futuro presumível. A máquina "aprende" a fazer essa imputação e, a partir de então, pode trabalhar de forma autônoma à vista da entrada de novos dados.

Durante o funcionamento do processo de decisão automatizada, as saídas podem ser otimizadas pelos programadores, mediante um processo ajuste fino do modelo por meio dos dados de *feedback*, ou mesmo por meta-algoritmos, como os de *backpropagation*; assim, o modelo pode criar decisões automatizadas ainda melhores, num processo teoricamente infinito de autoaprendizagem e autocorreção coadjuvado ou não por seres humanos, chamado de programação dinâmica (*dynamic programming*)[187].

No processo de autoaprendizagem são muito relevantes também as exceções, ou seja, aquelas situações que parecem se encaixar em certo padrão, mas na verdade são diferentes. É justamente nesse ponto que o modelo pode produzir decisões enviesadas ou iníquas, por generalizar demais ou de menos o padrão que lhe foi ensinado. Em tese, quanto mais dados são apresentados ao modelo, mais chance de ele encontrar exceções que precisam de um tratamento diferente. Em contraste, o modelo pode ser pobre em dados de treinamento, não atinando para padrões que seriam perceptíveis num conjunto maior de dados. Os dados de treinamento são determinantes para a acurácia de qualquer modelo de aprendizado de máquina atual.

Os modelos podem assumir grande número de processos decisórios em empresas, governos e organizações em geral, liberando recursos humanos e materiais para a assunção das exceções, normalmente ligadas a processos não quantificáveis. Assim, os benefícios da automatização para as empresas e organizações em geral são, antes de tudo, econômicos. No caso dos governos, a automatização pode trazer maior eficiência em serviços e políticas públicas, além de ser também fator de aperfeiçoamento econômico e administrativo.

Para os consumidores e usuários de serviços públicos, as vantagens dos processos automatizados residem na criação de comodidades cada vez mais personalizadas e, consequentemente, mais adequadas às necessidades específicas de cada indivíduo ou família.

[187] KUBAT, Miroslav, op.cit.,p. 338.

Desde a indicação de filmes, livros e produtos em geral, conforme os hábitos de consumo demonstrados em operações anteriores, até o relacionamento com o Fisco ou o deferimento de benefícios sociais ou outras prestações do Poder Público, conforme o perfil do contribuinte ou do grupo familiar, os mecanismos automatizados criam uma sinergia profunda que proporciona altos graus de eficiência em grande escala nos mais diferentes processos produtivos.

Numa visão mais radical e mais otimista, o processo de automatização levará a humanidade a uma Sociedade 5.0, de grande abundância e conforto proporcionado pelas máquinas, mediante a integração total de vários sistemas inteligentes, com a fusão quase completa do mundo *off-line* com o mundo *on-line*.

Embora alguns processos de automatização já tragam benefícios palpáveis para consumidores e usuários de serviços públicos, a ideia de uma sociedade 5.0 é muito mais ampla e profunda, porque imagina toda a vida social imersa no crisol da Inteligência Artificial, sem que haja a necessidade de "acessar" nada, uma vez que a realidade física estará ela mesma envolta e hibridizada com os mecanismos inteligentes, a tal ponto que não será possível perceber qualquer diferença entre estar *on-line* ou *off-line*. Nesse sentido, observou-se:

> In summary, Society 5.0 will feature an iterative cycle in which data are gathered, analyzed, and then converted into meaningful information, which is then applied in the real world; moreover, this cycle operates at a society-wide level.[188]

Seria, assim, um passo à frente da Indústria 4.0, que diz respeito apenas aos processos produtivos da indústria e do comércio, mas não de outros aspectos da vida individual e coletiva. Na Sociedade 5.0 as pessoas individual e coletivamente seriam o centro do processo tecnológico de disseminação da inteligência sobre objetos e sobre todo o ambiente circundante, ou seja, a culminância da perfilização.

3.3.1 Ciclo Virtuoso da Inteligência Artificial

Pelo visto, as vantagens do processo de automatização resultam da disseminação de "inteligência" sobre objetos inanimados, de tal maneira a "cognificar"[189] o mundo, fazendo com que objetos, tais como eletrodomésticos, automóveis, móveis, e até a infraestrutura das cidades,

---

[188] HITACHI-UTOKYO LABORATORY (H-UTOKYO LAB). **Society 5.0**: a people-centric super-smart society. Tokyo: Springer, 2018. Edição do Kindle, p.24.
Em resumo, o Sociedade 5.0 apresentará um ciclo iterativo no qual os dados são coletados, analisados e, em seguida, convertidos em informações significativas, que são então aplicadas no mundo real; além disso, este ciclo opera em um nível de toda a sociedade (tradução nossa).

[189] A expressão é de Kevin Kelly. Cf.:KELLY, Kevin. **Inevitável**: as 12 forças tecnológicas que mudarão nosso mundo. Rio de Janeio:Alta Books, 2019, Tradução de Cristina Yamagami, p. 31-65.

colaborem ativamente para ganhos de produção das empresas, melhoria de serviços públicos e aumento da qualidade de vida das pessoas.

Andrew Ng, uma das maiores autoridades no tema do aprendizado de máquina, diz, por essa razão, que a Inteligência Artificial é a nova eletricidade[190]. O caráter transversal dessa tecnologia tende a repetir o que ocorreu com a eletricidade, na virada do século XIX para o XX, isto é, tende a exercer influência sobre todas as áreas da vida humana, assim como se deu com a eletricidade. Desde tarefas domésticas, passando pela agricultura, pela indústria, pelo comércio, pelo entretenimento, enfim, tudo será de algum modo afetado pelo processo de espalhamento da inteligência artificial.

Logicamente, a aposta de que a IA será amplamente incorporada em objetos decorre do fato de que há atrativos muito claros na adoção dos mecanismos inteligentes, de modo que se pode presumir razoavelmente que a implementação dessas tecnologias ocorrerá sem a necessidade de qualquer incentivo adicional.

Adriano Mussa fala de um "Ciclo Virtuoso da Inteligência Artificial" para aqueles que implantarem a IA em seus negócios:

> Em linhas gerais, o ciclo funciona da seguinte forma: se a organização desenvolver um produto ou serviço de qualidade satisfatória, ela conseguirá alguns usuários iniciais. Os usuários iniciais, ao utilizarem o produto ou serviço, gerarão dados que serão coletados e armazenados pela organização. Esses dados, se bem tratados por técnicas de Inteligência Artificial, principalmente *Machine Learning*, possibilitarão a melhoria do produto ou serviço. O produto ou serviço aperfeiçoado levará à aquisição de mais usuários. Mais usuários gerarão mais dados; mais dados levarão à melhoria do produto ou serviço e esse ciclo seguirá continuamente.[191]

Ao contrário do que se pode pensar a partir do imaginário criado especialmente pela indústria cinematográfica, a IA não é uma poderosa e maligna ferramenta capaz até de se rebelar contra os seus criadores. A maioria das aplicações de IA hoje são estreitas (*narrow*), isto é, são direcionadas a finalidades bem específicas e limitadas. Não existe ainda, e provavelmente nunca existirá, uma Inteligência Artificial Geral (AGI, na sigla em inglês para *Artificial General Intelligence*), unificada e com aptidão para quaisquer propósitos. Repetindo a analogia com a eletricidade, Kevin Kelly especula sobre o futuro da implementação da IA nos negócios:

> Em meio a toda essa atividade, já podemos vislumbrar o futuro da IA. Esse cenário não envolve nem o HAL 9000 — a discreta máquina autônoma do filme 2001, Uma Odisseia no Espaço, animada por uma carismática (e potencialmente homicida) consciência similar à humana — nem o arrebatamento da singularidade tecnológica.

---

[190] ANDREW Ng: Artificial Intelligence is the New Electricity. Stanford: Stanford Graduate School of Business, 2 fev. 2017. 1 vídeo (1h 27 min). Publicado por Stanford Graduate School of Business. Disponível em: https://www.youtube.com/watch?v=21EiKfQYZXc. Acesso em 30 dez. 2020.

[191] MUSSA, Adriano. **Inteligência Artificial - Mitos e Verdades:**: as reais oportunidades de criação de valor nos negócios e os impactos no futuro do trabalho. São Paulo: Saint Paul, 2020. Edição Kindle, p.105.

> A inteligência artificial que já podemos ver despontando no horizonte é mais parecida com a Amazon Web Services — barata, confiável, com sua perspicácia em escala industrial por trás de tudo e quase invisível, exceto quando pisca ao ser desligada. Esse serviço público vai oferecer todo o QI que desejarmos, mas não mais do que precisaremos. Bastará conectar-se à rede para receber a IA como se fosse energia elétrica. Ela vai dar vida a objetos inertes, da mesma forma que a eletricidade fez mais de um século atrás. Há três gerações, muitos cientistas malucos fizeram fortunas criando versões elétricas para ferramentas e equipamentos comuns — por exemplo, a bomba manual e a máquina de lavar mecânica. Os empreendedores daquela época não precisavam gerar eletricidade. Eles a compravam na rede elétrica para automatizar o que antes era manual. Agora vamos cognificar aquilo que eletrificamos no passado. Praticamente tudo o que imaginarmos pode se tornar algo novo, diferente ou mais valioso com uma injeção de QI. Com efeito, é fácil prever como serão os planos de negócios das próximas 10 mil sturtups: 'Pegue X e adicione IA. Encontre algo que pode ser melhorado e o transforme por meio de inteligência online.'

As aplicações de IA atualmente estabelecem uma relação simples do tipo: A → B, em que "A" representa os dados de entrada (*Input*), "→" indica uma relação de implicação condicional, e "B", os dados de saída (*Output*). Os dados de saída resultam, portanto, do tratamento dos dados de entrada pelo modelo. O modelo cria uma conexão estatístico-matemática entre o *Input* e o *Output*, que emula a conexão semântica estabelecida pela inteligência humana, só que numa escala, precisão e velocidade muito maiores e, em compensação, infinitamente mais estreita e descontextualizada também.

Para que o modelo funcione adequadamente, os programadores "ensinam", com dados de treinamento, qual a conexão "correta" a ser estabelecida. Em seguida, o próprio modelo "aprende" o padrão da conexão e passa replicá-la. Quando já na fase de aplicação, os usuários também acabam ajudando o modelo a melhorar, por meio de suas interações, que nada mais são do que rotulações para o modelo. Por exemplo, quando o usuário dá um *like* num produto, ele rotula aquele produto — e todos os que a ele estão ligados — como um *output* desejável para si, caso posteriormente ele faça uma pesquisa de compra. O mesmo ocorre também quando o usuário, por exemplo, marca um *e-mail* como *spam*: o modelo incorpora esse rótulo como negativo, posteriormente qualificando *e-mails* com o mesmo padrão como *spams*.

Observa-se, assim, que à medida que o modelo entra em contato com os usuários e suas rotulações, salvo interferências propositais do programador, ele vai se amoldando às preferências e repulsões que estes manifestam, potencializando os comportamentos tidos como normais. Isso vale para o indivíduo e para o grupo. Há um processo de *perfilização* constante, individual e grupal.

Vê-se também que o modelo carece de muitos dados para ter acurácia e robustez, pois ele só prediz algo com que já tenha tido contato anterior. Por isso Inteligência Artificial e *Big Data* (grandes conjuntos de dados) andam juntos.

Se os dados de entrada e os dados de saída são conhecidos do programador, a máquina será programada para aprender de modo supervisionado (*supervised learning*-*SL*). Neste caso, o programador, na fase de treinamento, alimenta a máquina com os dados de entrada e também com os dados de saída, de modo a estabelecer o vínculo estatístico-matemático.

Aqui, porém, há uma subdivisão importante: a) o SL pode se dar por meio de *Statistical Machine Learning*, isto é, uma forma em que o algoritmo contém previamente fórmulas para calcular probabilidades e com base nelas gerar o *output*; ou b) por meio de *Deep Learning*-DL, em que o programador não cria totalmente as fórmulas de cálculo, mas apenas esboça um modelo em camadas aparentes de uma Rede Neural Artificial e depois alimenta essa rede com vastos volumes de dados de entrada, associando-os aos dados de saída "corretos" (rotulados), deixando que o próprio algoritmo, por tentativas e erros, encontre os pesos adequados para cada variável de modo tal que essas associações se encaixem de forma correta. Essas tentativas e erros, quando encerradas, geram camadas profundas e ocultas na Rede Neural Artificial, que o próprio modelo cria e que sequer é do conhecimento do próprio programador.

Grosso modo, no *Statistical Machine Learning* o programador ensina a pergunta, a resposta certa e a forma de chegar a ela; no *Deep Learning*, o programador mostra a pergunta e a resposta certa, mas não diz como chegar a ela, cabendo ao modelo criar esse caminho. E o caminho criado pelo modelo pode ser extremamente eficaz — os modelos de DL têm atingido 95% de acurácia de predição —, embora ele estabeleça conexões que, para nós, humanos, não fazem sentido algum, em termos de relação de causa e efeito.

Adriano Mussa, após explicar como funciona um modelo preditivo baseado em *Statistical Machine Learning*, no qual o programador escolhe as variáveis relevantes (área do imóvel, localização, tempo de construção, etc.) e ensina a máquina qual peso dar a cada uma delas, estima como seria o processo de *Deep Learning* na mesma situação:

> Na prática, alimentamos os algoritmos de DL com a camada de *Input* – A e com os dados de resultado, *Output* – B, e são os algoritmos que buscam todas as combinações possíveis de variáveis, testando a criação de inúmeras camadas e neurônios para buscar, matematicamente, a melhor combinação e pesos, que expliquem os preços dos imóveis com a maior acurácia possível, com base em suas características.
> Em outras palavras, os algoritmos buscam aumentar a acurácia do modelo utilizando as inúmeras combinações de variáveis, criando neurônios e utilizando pesos que otimizem a sua performance, independentemente de elas fazerem ou não sentido para nós, seres humanos."[192]

[192] MUSSA, op.cit, p.86.

Os modelos de DL, portanto, buscam extrair diretamente dos dados de entrada a combinação mais eficiente para chegar aos dados de saída, que por sua vez são rotulados conforme o objetivo do programador. Na concepção desse caminho lógico-matemático, o modelo acaba organizando camadas escondidas (*hidden layers*) de "neurônios artificiais" que atribuem pesos às diferentes combinações, preferindo aquelas cuja soma mais se aproxime do resultado de saída desejado. Em outras palavras, as camadas ocultas trabalham otimizando funções matemáticas que sejam capazes de transformar os dados de entrada na resposta informada pelo desenvolvedor do modelo na fase de treinamento.

Mesmo o programador original do modelo não saberá completamente como a Rede Neural Artificial chegou àquela combinação, tal a quantidade de cálculos e de arranjos testados pela máquina. Essas camadas intermediárias, assim, formam uma verdadeira "caixa preta" que oculta a maior parte do processo decisório automático. Assim, se por um lado elas tornam o modelo extremamente robusto para obter as respostas desejadas, por outro elas tornam opaco o processo decisório. Nas camadas escondidas dos modelos está a virtude e o vício do DL.

Quanto mais complexo for o problema a ser resolvido pela Rede Neural Artificial, mais camadas ocultas de combinações e pesos podem ser criadas pelo modelo para aumentar a acurácia. Em compensação, mais obscuros se tornam os critérios de cálculo, ou seja, mais densa a caixa-preta.

Nos modelos de *Statistical Machine Learning*, em que o programador escolhe as variáveis que o modelo deve levar em conta, o que ocorre é que o modelo ficará limitado à visão humana de causalidade, que apenas leva em conta os vínculos fortes entre entrada e saída. Se o mesmo problema de negócio for apresentado a um modelo de *Deep Learning*, ele encontrará correlações que não ocorreriam à mente humana, por aparentemente não terem vínculo de causalidade com o resultado.

Um exemplo impressionante, lembrado por Kai Fu Lee[193], é aquele do modelo criado por uma empresa chinesa para decidir automaticamente sobre a concessão de pequenos empréstimos com base em dados do celular do interessado. Uma Rede Neural Artificial, devidamente treinada com milhões de dados históricos de pequenos empréstimos, descobriu que o nível de bateria médio do celular ao longo do dia, a data de nascimento ou a velocidade de digitação do pedido de empréstimo pelo celular do interessado, tinham correlação com a classificação dele como bom ou mau pagador (os bons pagadores geralmente tinham a bateria do celular mais carregada, por exemplo).

[193] Op. cit., p.139.

Como isso se dá? Após ter acesso a um vasto conjunto de dados de bons e maus pagadores, o modelo de DL, na fase de treinamento, é apresentado a esses dados, tendo o programador previamente informado (rotulado ou etiquetado) os dados de saída (bons ou maus pagadores). A rede neural então, ante os dados de entrada (os mais diversos dados extraídos dos celulares, tais como tempo de uso diário, nível médio de bateria, sites que navega comumente, etc.), não procura "entender" o porquê de aquele ser um bom ou mau pagador — como faria um ser humano, que pensa em termos de causa e efeito — mas sim criar uma função matemática que ligue de maneira ótima os dados de entrada dos bons pagadores aos dados de saída respectivos, rotulados pelo programador. E o mesmo processo é feito com o telefone de muitos maus pagadores. Ao final desse treinamento, o modelo terá encontrado padrões nos bons e nos maus pagadores, levando em conta elementos que, para um ser humano, seriam completamente irrelevantes, tal como a carga média da bateria do celular, referida acima. Por isso que é apenas metafórica a comparação dos processos decisórios automatizados com a inteligência humana. O que a máquina faz é algo muito diferente do pensamento humano, embora chegue a resultados parecidos e eventualmente com maior acurácia. Edsger Dijkstra, a esse propósito, afirmou: "A questão de saber se um computador pode pensar não é mais interessante que a questão de saber se um submarino pode nadar."[194]

Uma descoberta como essa (que a carga média da bateria do celular influencia na probabilidade de que o contrato seja cumprido) pode representar um *insight* comercial que dá ao operador do modelo uma vantagem relevante, em relação aos concorrentes, sobretudo quando se pensa em grande escala. Mas pode também representar, a depender de qual seja o elemento diferenciador revelado pelos dados, uma fonte involuntária de discriminação de pessoas, grupos ou ideias.

A grande revolução do aprendizado de máquina ocorreu justamente com o *Deep Learning-DL*, e a maioria das atuais aplicações daquilo que se chama de "Inteligência Artificial" nada mais é do que DL. Durante muito tempo, entre os anos 1970 até os anos 1990, prevaleciam nas aplicações de IA os chamados Sistemas Especialistas, que eram mecanismos inteligentes baseados em regras[195]. O programador avaliava o problema do mundo real e tentava modelá-lo por meio de regras, que depois seriam aplicadas por um motor de inferência a novos dados de entrada. A deficiência dessa abordagem é que, por vezes, muito difícil e laborioso

[194] NORVIG, Peter Peter; NORVIG, Peter. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Elsevier, 2013. Tradução de Regina Célia Simille, p. 932.

[195] SEJNOWSKI, Terrence J. **A revolução do aprendizado profundo**. Rio de Janeiro: Alta Books, 2019. Traduzido por Carolina Gaio, p. 35-37.

criar as regras específicas para cada situação, e mais ainda para as exceções que se intersectam com a regra em alguns pontos. Um programa de reconhecimento da imagem de um gato, por exemplo, dependeria de escrever em código minuciosamente o que "é" um gato e centenas, talvez milhares de regras sobre o que não é um gato, mas sim uma onça, um puma, ou outro felino. Ora, não é assim que funciona a mente humana, a mais avançada forma de inteligência que conhecemos. Simplesmente sabemos muitas coisas que não podemos verbalizar em termos estritos. Aquilo que chamamos de "senso comum", por exemplo, é constituído de um vasto conhecimento sobre leis físicas e sociais que não podem ser codificadas, tanto mais porque não sabemos exatamente quais são elas, embora as apliquemos no dia a dia.

Nos anos 1980, Douglas Lenat, por meio de um projeto chamado CYC[196], tentou codificar o "senso comum" de um ser humano. O programa chegou a acumular mais de 1 milhão de regras, sem, no entanto, conseguir abranger algo que um humano comum aprende ainda na infância.

De fato, há muitas coisas que um ser humano sabe, mas não consegue expressar em palavras, e muito menos de forma quantitativa. Santo Agostinho escreveu que sabia o que era o tempo, mas bastava alguém pedir-lhe para dizer o que era, que não sabia mais[197]. Esse célebre pensamento ilustra a maneira como funciona a inteligência humana. Muita coisa é aprendida simplesmente por exemplos e repetições de padrões, sem a necessidade de que a mente analise todos os aspectos e relações do objeto conhecido.

O *Deep Learning* parte de uma abordagem diferente, mais próxima do funcionamento do cérebro humano. As dificuldades enfrentadas pela abordagem baseada em regras e heurísticas, estimulou os pesquisadores em IA a buscar saídas que fossem mais factíveis. Segundo Sejnowski[198], quatro coisas indicavam que era ruim trilhar pelo caminho da criação de regras para desenvolver um sistema inteligente, porque: a) o cérebro humano trabalha primeiro com reconhecimento de padrões, as regras surgem depois; b) é preciso uma prática repetitiva para que o cérebro domine atividades mais complexas; c) o cérebro não se orienta por regras no dia a dia, embora possa trabalhar com elas em um nível mais profundo do pensamento; e, finalmente, d) nossos cérebros têm bilhões de neurônios que se intercomunicam, o que sugere que ele trabalha com processamento paralelo dos dados de entrada e não com processamento linear (arquitetura de Von Neumman).

---

[196] https://www.cyc.com/the-cyc-platform. Acesso em 20 dez 2020

[197] "O que é, por conseguinte, o tempo? Se ninguém mo perguntar, eu sei; se o quiser explicar a quem me fizer a pergunta, já não sei". (AGOSTINHO, Santo. **Confissões**. São Paulo: Companhia das Letras, 2017. Tradução de Lorenzo Mammì. Edição do Kindle, p. 237.)

[198] SEJNOWSKI, op. cit., p. 41-42.

Foi essa abordagem que permitiu a maior parte dos progressos efetivos na área de IA aplicada a negócios. Os modelos de *deep learning* muitas vezes atingem 95% de acurácia, em certas tarefas, o que era impensável antes do uso dessa técnica. E esse nível de acurácia, como explica Adriano Mussa, não é raro em *Deep Learning*:

> Esse percentual elevado de acurácia dos algoritmos de *Deep Learning* não é exceção. Ele tem sido observado em uma infinidade de aplicações de setores e contextos diferentes, mostrando sua forte robustez.[199]

Assim, os modelos de *Deep Learning* asseguram aa principais vantagens do uso de IA em negócios ou qualquer aplicação que dependa de julgamentos: rapidez, economia e acurácia.

3.4 Riscos

Como demonstrado no item anterior, a maior parte das decisões automatizadas que são atualmente colocadas em prática resultam de modelos de *Deep Learning* em Redes Neurais Artificiais. Convém, assim, ter presente que os riscos que foram levados em conta pelo próprio legislador estão associados a esse método de tratamento de dados.

O primeiro e mais conhecido risco das decisões automatizadas decorrentes de DL é o da opacidade. Pelo próprio volume de cálculos e pela quantidade de dados necessária para a concepção de um modelo de DL, não é acessível sequer para os desenvolvedores o processo exato por meio do qual o modelo chegou a esta ou aquela predição ou mesmo decisão, e isso naturalmente pode levantar desconfianças e suposições em relação à higidez do modelo, eventualmente exigindo uma coadjuvação humana para que ele possa ser colocado em prática. Terrence Sejnowski[200] exemplifica bem o problema com o caso dos diagnósticos médicos:

> Embora possam dar resposta correta para um problema, atualmente não sabemos como as redes neurais chegam a ela. Por exemplo, suponha que uma paciente chegue a um pronto-socorro com uma dor aguda no peito. Trata-se de um infarto agudo do miocárdio, o que precisa de intervenção imediata, ou simplesmente um caso grave de indigestão? Uma rede treinada para diagnosticar pode ser mais precisa do que o médico responsável pela triagem; mas, sem uma explicação sobre como a rede tomou a decisão, a relutância em confiar nela seria plausível. Os médicos também são treinados para acompanhar o que equivale a algoritmos, séries de testes e pontos de decisão que os orientam em casos de rotina. O problema é que há casos raros, que estão fora do escopo de seus 'algoritmos', enquanto uma rede neural treinada com muito mais casos, mais do que a média dos médicos verá em toda uma vida, pode muito bem dirimir sobre esses casos raros. Mas você confiaria mais no diagnóstico estatisticamente mais sólido de uma rede neural, sem explicação de como foi feito, do que no de um médico com um diagnóstico plausível?

---

[199] MUSSA, op.cit.,p. 91.

[200] Op.cit., p. 134.

A opacidade também pode decorrer de fatores comerciais. O desenvolvedor do modelo pode até saber explicar como se chegou a certa decisão, mas a exposição desse caminho poderia revelar o seu "segredo comercial ou industrial", que na verdade é a sua fonte de ganhos com o modelo.

A LGPD cuida do ponto, na esteira do GDPR, estabelecendo no art. 20, §1º, um direito à explicação em caso de decisão automatizada, como primeira linha contra a opacidade, mas respeitado o segredo comercial e industrial — e é difícil, na prática, conciliar essas duas coisas.

Se a explicação não for dada pelo controlador ao titular dos dados, sob o argumento da existência de segredo comercial ou industrial, então a Autoridade Nacional de Proteção de Dados - ANPD pode ser acionada para fazer uma verificação sobre possíveis vieses discriminatórios (LGPD, art. 20, §3º). No capítulo seguinte avalia-se melhor essa questão, mas de logo chama a atenção a estreiteza da norma, que deixa duas importantes questões em aberto: a) E se a explicação for negada com outro fundamento, que não o segredo comercial ou industrial? (Por exemplo, a alegação de que o próprio controlador não sabe exatamente como o modelo funciona); b) E se não houver discriminação, mas sim outro tipo de violação a direitos individuais?

Os tecnólogos têm tentado criar modelos que sejam capazes de ser autoexplicativos, a chamada Inteligência Artificial Explicável (XAI, na sigla em inglês para *Explanable Artificial Intelligence*). Mas aqui a questão esbarra na autorreferência. É que o próprio cérebro humano é também uma caixa-preta. De fato, a objeção de que um modelo opera com uma caixa-preta pode ser aplicada também ao cérebro humano. Não há até aqui conhecimento objetivo e minucioso sobre os processos decisórios humanos, exceto que se sabe que há muito mais viés e irracionalidade do que se imaginava. Eventuais explicações humanas, muitas vezes, são meramente retóricas. Modelos de XAI podem recair no mesmo impasse. O risco de opacidade, portanto, permaneceria sendo um problema insolúvel.

Outro ponto, ainda sobre a opacidade, é que a concepção de uma decisão automatizada envolve o tratamento de muitos dados e a revelação do seu processo para um titular poderia ensejar a violação da intimidade de outros titulares, cujos dados também foram levados em conta na decisão, para efeito de comparação, e a quem pode não interessar a divulgação do processo decisório.

Um segundo risco criado pelas Redes Neurais Artificiais de *Deep Learning* é que elas dependem de um grande volume de dados para funcionarem bem, o que gera uma corrida por dados pessoais. Com efeito, a acurácia das decisões automatizadas criadas por Redes Neurais

Artificiais depende fundamentalmente de um vasto conjunto de dados (*Big Data*), especialmente na fase de treinamento e validação, e também para evoluir na fase de aplicação. Essa necessidade faz com que aumentem os riscos ligados à privacidade, porque os desenvolvedores buscarão sempre ter acesso ao máximo de dados pessoais para criarem, validarem e aplicarem modelos preditivos e decisórios de alto desempenho. Com o advento da Internet 5G e a implantação da Internet das Coisas, estima-se que a coleta de dados crescerá exponencialmente, já que atividades triviais do dia a dia e até da intimidade doméstica, como abrir uma geladeira, fechar uma porta, ou ligar uma lâmpada, poderão ser incorporadas à internet e gerarão dados pessoais suscetíveis de serem usados em modelos preditivos. Presumivelmente, isso multiplicará muitas vezes os riscos à privacidade.

Um terceiro risco, ligado ao anterior, é que, além de precisarem de um grande volume de dados, as Redes Neurais expressam apenas o conhecimento que se pode extrair desses dados, não mais que isso. Logo, se há no conjunto de dados de treinamento um viés, proposital ou não, esse viés se replicará indefinidamente nas decisões.

O caso mais conhecido sobre isso ocorreu em Los Angeles (EUA)[201], e dizia respeito ao reconhecimento facial em locais públicos. Descobriu-se que um modelo de reconhecimento facial da polícia cometia mais erros em relação a negros do que em relação a brancos, porque, enquanto os "procurados", na fase de aplicação do modelo, eram na maior parte pessoas negras, na fase de treinamento o modelo fora apresentado a um número maior de faces brancas, tornando-se naturalmente melhor em reconhecer estas do que outras.

Tal situação pode se repetir em muitas outras áreas. A escolha dos dados de treinamento não é um ato neutro; muito menos o é a rotulação dos dados de saída, feita pelos programadores. Aqui a escolha envolve aspectos ideológicos, muitas vezes inconscientemente. Como explica Terrence Sejnowski:

> Todas as redes neurais que classificam entradas são tendenciosas. Em primeiro lugar, a escolha das categorias de classificação incorpora um viés que reflete o preconceito humano na forma como esmiuçamos o mundo. Por exemplo, seria útil treinar uma rede para detectar ervas daninhas em gramados. Mas como identificá-la? A erva daninha de um homem pode ser a flor silvestre de outro. A classificação é um problema muito mais amplo, que reflete vieses culturais. Essas ambiguidades precisam integrar os conjuntos de dados usados para treinar a rede.[202]

---

[201] GARVIE, Clare; FRANKLE, Jonathan. Facial-Recognition Software Might Have a Racial Bias Problem: Depending on how algorithms are trained, they could be significantly more accurate when identifying white faces than African American ones.. *In*: THE ATLANTIC. **The Atlantic**, 6 abr. 2016. Disponível em: https://www.theatlantic.com/technology/archive/2016/04/the-underlying-bias-of-facial-recognition-systems/476991/. Acesso em: 22 dez. 2020.

[202] SEJNOWSKI, op.cit., p. 135.

Pior ainda, com a aplicação do modelo em massa, produzem-se *loopings* que reforçam o viés original. Cathy O´Neal[203] exemplifica esse fenômeno com os modelos de otimização do policiamento ostensivo. Como esses modelos usam dados relativos a pequenas infrações, tais como perturbação da ordem, posse de pequena quantidade de droga e vadiagem, os policiais acabam sendo enviados para patrulhar regiões pobres, onde normalmente acontecem essas infrações. Com o aumento do patrulhamento, aumentam também as prisões por essas pequenas contravenções, o que induz a realimentação e o reforço por *feedback* ao modelo para aumentar o patrulhamento nesses locais.

Ainda no campo dos vieses e seu ciclo vicioso de reforço, observa Caathy O'Neal também que, embora a cor da pele a condição social não sejam incluídas no modelo como parâmetros para a inferência, o fato é que os dados escolhidos (sobre pequenos delitos) para esse tipo de policiamento acabam funcionando como *proxies* para a raça e a pobreza, já que apenas negros e hispânicos da periferia são presos, segundo a autora, por esse tipo de crime nos Estados Unidos. Um indivíduo branco que pratique ações semelhantes num campus universitário dificilmente deparará com uma patrulha policial.

Há um quarto risco, não menos grave, no uso de mecanismos inteligentes para formulação de decisões automatizadas. É que a grandeza que é escolhida para ser otimizada pode subdimensionar outras questões relevantes. Assim, se o modelo visa ao lucro — e a maioria visa a isso, naturalmente — a função de lucro deve ser otimizada pelo modelo, no que não há nada de ilegal ou imoral. Acontece que essa otimização, quando feita em termos matemáticos, é implacável. O modelo não se deterá diante de nenhuma circunstância, a não ser que programado para isso, para aumentar os lucros. Como mecanismo de Inteligência Artificial Fraca ou Estreita, o modelo não é capaz de contextualizar as decisões para além dos dados que lhes foram apresentados, de modo que se o lucro é o que deve ser maximizado, ele fará isso *per fas et per nefas*.

Eventualmente, essa "objetividade" inexorável pode produzir danos imensos, sobretudo quando aplicada em grande escala. E aqui se chega a um risco transversal de todos os modelos matemáticos para produzir decisões automatizadas: a escala. É a escala que gera os maiores danos.

Como explica Cathy O´Neal[204], é a escala transforma o que seria um pequeno incômodo em algo com a força de um tsunami. Ao estabelecer um ciclo de decisão em um número imenso

[203] O'NEIL, op.cit.
[204] O'NEIL, op.cit., p. 48.

de casos idênticos, o modelo em larga escala acaba influenciando o ambiente de duas formas: a) ele reforça em massa um padrão, inferido de situações anteriores (que podem ser injustas); b) ele induz o comportamento futuro das pessoas, que tentarão se ajustar ao modelo.

A escala das decisões automatizadas gera um problema adicional. A regulamentação ou qualquer ação legal que vise a solucionar um problema gerado por algoritmos pode não conseguir atingir o seu objetivo, justamente por não ser escalável. Ou seja, enquanto decisões automatizadas são tomadas *on-line* e em massa, as soluções legislativas tendem a depender de uma análise artesanal, caso a caso. A brutal diferença de velocidade e de volume pode levar a norma à completa ineficácia prática. Assim, as regulamentações precisarão contar com mecanismos de implementação escaláveis. Nesse sentido, observam Kearns & Aaron[205]:

> Regulations and laws certainly have a crucial role to play—as we have emphasized throughout, the specification of what we want algorithms to do and not do for us should remain firmly in the human and societal arenas. But purely legal and regulatory approaches have a major problem: they don't scale. Any system that ultimately relies solely or primarily on human attention and oversight cannot possibly keep up with the volume and velocity of algorithmic decision-making. The result is that approaches that rely only on human oversight either entail largely giving up on algorithmic decision-making or will necessarily be outmatched by the scale of the problem and hence be insufficient. So while laws and regulations are important, we have argued in this book that the solution to the problems introduced by algorithmic decision-making should itself be in large part algorithmic.[206]

Em resumo, os principais riscos dos mecanismos de decisão automatizada são: a) opacidade; b) necessidade de grande volume de dados, com riscos à privacidade; c) viés; d) subdimensionamento de grandezas diversas daquela buscada pelo controlador dos dados; e) escala.

---

[205] KEARNS; ROTH, op.cit., p.192.

[206] Regulamentos e leis certamente têm um papel crucial a desempenhar - como enfatizamos ao longo do texto, a especificação do que desejamos que os algoritmos façam e não façam por nós deve permanecer firmemente nas arenas humana e social. Mas as abordagens puramente legais e regulatórias têm um grande problema: elas não escalam. Qualquer sistema que, em última análise, dependa única ou principalmente da atenção e supervisão humanas, não pode acompanhar o volume e a velocidade da tomada de decisão algorítmica. O resultado é que as abordagens que dependem apenas da supervisão humana implicam em desistir amplamente da tomada de decisão algorítmica ou serão necessariamente superadas pela escala do problema e, portanto, insuficientes. Portanto, embora as leis e os regulamentos sejam importantes, argumentamos neste livro que a solução para os problemas introduzidos pela tomada de decisão algorítmica deve ser em grande parte algorítmica (tradução nossa).

4 DECISÕES AUTOMATIZADAS: OS DIREITOS DO TITULAR

4.1 Introdução

Como visto no capítulo anterior, a formação das decisões automatizadas dá-se através do tratamento de dados pessoais por mecanismos inteligentes, notadamente por modelos que usam aprendizado de máquina (*machine learning*) para otimizar funções previamente designadas pelos desenvolvedores nas fases de treinamento e validação desses modelos (com o uso de dados de treinamento e de teste), em diferentes áreas da economia privada e dos governos.

A pesquisa sobre os direitos do titular em casos de decisões automatizadas precisa, então, considerar todo o processo de concepção da decisão automatizada, e não apenas o momento em que ela é manifestada, visto que podem ocorrer violações de direitos durante todo o *iter* decisório.

Os riscos abstratos que normalmente acompanham a estrutura do processo decisório automatizado (opacidade, violação da privacidade, viés discriminatório, subdimensionamento de grandezas diversas daquela buscada pelo desenvolvedor e escala) podem transformar-se em ameaças concretas ou mesmo danos a interesses específicos dos titulares em hipóteses particulares, de modo que o legislador da LGPD, inspirado no GDPR, prenunciou alguns direitos ao titular dos dados, nos casos de decisões automatizadas (art. 20 da LGPD: direito à revisão e direito à explicação). Porém, esses direitos referem-se apenas ao resultado do processo decisório, e não esgotam o rol das prerrogativas dos titulares previstas na própria LGPD, na legislação correlata e na Constituição Federal, quanto ao processo de formação da decisão, que é uma forma de tratamento de dados[207].

Segundo as Diretrizes do Conselho Europeu de Proteção de Dados[208], as soluções legais na matéria das decisões automatizadas e perfilização visam essencialmente a criar: a) transparência e justiça nas decisões; b) obrigações de prestação de contas pelo controlador; c) direito dos titulares de não serem perfilizados contra a sua vontade, a não ser em situações

[207] A Universidade de Toronto, em articulação com outras instituições, produziu um interessante relatório sobre o uso de decisões automatizadas nas atividades de polícia de imigração e refugiados do Canadá, no qual ficou demonstrado que os riscos aos direitos humanos ocorrem tanto nos resultados do processo decisório, como na formação da decisão. Cf.: 20180926-IHRP-Automated-Systems-Report-Web.pdf (utoronto.ca)

[208] JUSTICE AND CONSUMERS (Europea Union). Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. In: **JUSTICE AND CONSUMERS** (Europea Union). European Commission. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 11 jan. 2021

legalmente previstas; e) necessidade de avaliação de impacto das tecnologias de automatização de decisões.

Dito isso, convém avaliar todo o texto da LGPD, para verificar quais são os direitos atualmente consagrados aos titulares, sob o ângulo das decisões automatizadas.

4.1.1 Direitos dos titulares na LGPD

Para fins de sistematização, a melhor maneira de investigar os direitos dos titulares é analisar o processo decisório automatizado, desde o seu início (com a criação e treinamento do modelo) até o fim do tratamento de dados.

É preciso ter presente, em adição, que, para além dos direitos dos titulares, o processo decisório automatizado também pode ser constrangido por amplas medidas de ordem institucional e de infraestrutura tomadas por governos e empresas sem a participação direta dos titulares de dados pessoais, que desempenham relevante papel no ajuste e na evolução das tecnologias. Essas medidas, todavia, não estão no foco desta pesquisa, que está centrada nos direitos individuais e coletivos dos titulares.

Tendo em conta a dinâmica necessária para a concepção e aplicação de uma decisão automatizada, pode-se classificar os direitos dos titulares da seguinte maneira:

a) Direitos relativos à formação da decisão — como tais entendidos aqueles relacionados à criação do modelo de decisão e à sua operação em geral;
b) Direitos relativos aos resultados da decisão automatizada — como tais entendidos aqueles relacionados a uma decisão em concreto já tomada.

Os direitos relativos à formação da decisão tanto poderão ser individuais como coletivos, como se verá adiante. Já os direitos relativos aos resultados concretos de uma decisão automatizada normalmente serão individuais, embora não se possa excluir a hipótese de tutela coletiva em casos de direitos individuais homogêneos.

Há também alguns direitos que são garantias, isto é, são necessários para o exercício dos demais direitos. Nessa categoria estão: a) o direito de petição; e b) o direito ao devido processo legal.

Assim, temos:

I) Direitos Relativos à Formação da Decisão Automatizada:

1) Direito ao *design* adequado do sistema (art. 49, LGPD), inclusive com preferência para padrões técnicos que possibilitem ao próprio titular a fiscalização (art. 51, LGPD);

2) Direito à confirmação da existência do tratamento (art. 18,I c/c art. 19, LGPD);

3) Direito de opor-se ao tratamento (art. 18, VIII e §2º, LGPD)

4) Direito de acesso aos dados pessoais e aos compartilhamentos (art. 9º c/c art. 18, II e VII, LGPD);

5) Direito de correção (art. 18, III, LGPD);

6) Direito à anonimização (art. 18, IV, LGPD);

7) Direito de consentir ou não no tratamento (arts. 7°, 11 e 14, LGPD);

8) Direito de revogar o consentimento (art. 15, III c/c art. 18, XI, LGPD);

9) Direito à eliminação dos dados (art. 18, VI, LGPD);

II) Direitos Relativos aos Resultados da Decisão Automatizada:

1) Direito de Revisão das Decisões Automatizadas (art. 20, LGPD e art. 5º, VI da LCP);

2) Direito à Explicação (art. 20, §1º, LGPD)[209]

III) Direitos instrumentais:

1) Direito de petição (CF, art. 5º, XXXIV, "a"; LGPD, art. 18,§1º c/c art. 55-J, V) ;

2) Direito ao devido processo legal (CF, art. 5º, LIV; LGPD, art. 4º, §1º)

Vale ressaltar que, como explorado em capítulo anterior, num sentido bastante amplo todos esses direitos podem ser reconduzidos ao *direito fundamental à proteção de dados*, que abrange a ideia de autodeterminação informativa, liberdade e privacidade. Todavia, para efeito de atingir o objeto desta pesquisa, importa ter em consideração os direitos numa perspectiva mais operacional e concreta, em relação à produção de decisões automatizadas.

4.1.2 Outros direitos

Os direitos acima enumerados não são reciprocamente excludentes, antes são cumulativos, e inclusive podem articular-se com muitas outras prerrogativas previstas em leis setoriais de proteção ao consumidor (CDC), à navegação na Internet (Marco Civil da Internet), ao bom pagador (Lei do Cadastro Positivo), ao titular de informações pessoais em bancos de dados (Lei do Habeas Data), e com o Código Civil e leis administrativas (Lei de Acesso à

[209] Na doutrina, há quem defenda que o direito à explicação pode ser exercido mesmo antes da tomada da decisão automatizada: "Vale destacar que o legislador não explicitou o momento no qual devem ser fornecidas as informações, deixou a critério do titular dos dados. Este pode solicitar a qualquer tempo." (SOUZA, Carlos Afonso; PERRONE, Christian; MAGRANI, Eduardo. O direito à explicação entre a experiência europeia e sua positivação na LGPD. *In:* MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle, p.275.)

Informação, por exemplo). A sua enumeração tem por propósito apresentar e sistematizar os direitos do titular decorrentes da LGPD, sem qualquer sentido de exclusão de outros direitos que nasçam dos diferentes diplomas legislativos relacionados, e da própria Constituição Federal, como aliás expressamente menciona a própria LGPD no art. 64, verbis: "Os direitos e princípios expressos nesta Lei não excluem outros previstos no ordenamento jurídico pátrio relacionados à matéria ou nos tratados internacionais em que a República Federativa do Brasil seja parte."

Bruno Bioni, invocando a chamada Teoria do Diálogo das Fontes, afirma que a LGPD coloca-se, ela mesma, como "uma fonte normativa materialmente geral que deve conversar com as demais para governar o uso dos dados pessoais."[210] Dentro dessa visão, o mesmo autor afirma que o relacionamento da LGPD com o restante da legislação pertinente à proteção de dados pessoais deve observar o seguinte[211]:

a) Coerência sistemática: a normativa de proteção de dados deve ser vista como fazendo parte de uma unidade maior, sistêmica; assim, cada lei, atual ou futura, que venha a dispor direta ou indiretamente sobre dados pessoais, pode ser considerada um elemento de um conjunto, devendo qualquer resultado interpretativo a respeito delas guardar compatibilidade interna e externa com todas as disciplinas concorrentes desse mesmo conjunto;
b) Complementaridade ou subsidiariedade: os direitos do titular consagrados na LGPD (que são investigados nesta pesquisa com foco nas decisões automatizadas) não excluem outros que decorram do ordenamento jurídico em geral (notadamente as já referidas leis: CDC, Marco Civil da Internet e Lei do Cadastro Positivo) e que façam parte do Sistema da Proteção de Dados Pessoais.
c) Coordenação ou adaptação sistêmica: as leis de proteção de dados pessoais, inclusive a LGPD, interagem reciprocamente, de sorte que os conceitos e princípios precisam ser lidos de modo conectado, cabendo sempre entender que leis gerais influenciam disposições especiais e vice-versa.

[210] BIONI, Bruno R. **Proteção de dados pessoais**: a função e o limite do consentimento. 2.ed. Rio de Janeiro: Forense, 2020, p. 260.
[211] Ibidem.

4.2 Direitos Relativos à Formação da Decisão Automatizada

4.2.1 Direito ao design adequado do sistema (art. 49), inclusive com preferência para padrões técnicos que possibilitem ao próprio titular a fiscalização (art. 51)

Decisões automatizadas são resultados de tratamento de dados por aparelhos artificiais. Esses aparatos tentam de algum modo reproduzir a forma como a mente humana funciona. Acontece que o mecanismo de operação do cérebro humano não é totalmente conhecido e, além disso, deve a sua eficiência ao um *design* extremamente complexo — que, segundo a teoria mais aceita, é fruto de uma seleção evolutiva mais ou menos caótica —, até aqui reproduzido em frações mínimas por máquinas, que apenas emulam pequenas parcelas dos aspectos tangíveis do pensamento humano.

Steven Pinker, comparando a mente humana aos processos de aprendizado de máquina, observa:

> Nossos órgãos físicos devem seu design complexo às informações contidas no genoma humano, e o mesmo, a meu ver, aplica-se aos nossos órgãos mentais. Não aprendemos a ter um pâncreas, e também não aprendemos a ter um sistema visual, aquisição da linguagem, bom senso ou sentimentos de amor, amizade e justiça. Nenhuma descoberta isolada comprova essa afirmação (assim como nenhuma descoberta isolada comprova que o pâncreas tem uma estrutura inata), mas muitas linhas de evidência convergem nessa direção. A que mais me impressiona é o Desafio do Robô. Cada um dos grandes problemas de engenharia resolvidos pela mente é insolúvel na ausência de hipóteses incorporadas sobre as leis que se aplicam na respectiva arena de integração com o mundo. Todos os programas criados por pesquisadores da inteligência artificial foram especificamente projetados para uma área específica, como linguagem, visão, movimento ou um dos muitos tipos diferentes de bom senso. Nas pesquisas sobre inteligência artificial, o orgulhoso criador de um programa às vezes o apregoa como uma mera amostra de um sistema de uso geral a ser elaborado futuramente, mas todo mundo da área rotineiramente descarta bazófias desse tipo. Predigo que ninguém jamais construirá um robô semelhante a um ser humano — e me refiro a um robô *realmente* semelhante a um ser humano — a menos que o equipe com sistemas computacionais feitos sob medida para resolver diferentes problemas.[212]

O design das máquinas inteligentes é concebido com hipóteses incorporadas que nem sempre correspondem à realidade. Ao contrário do cérebro humano, que evoluiu em contato com o ambiente e segundo a seleção natural, os computadores precisam de uma seleção artificial para se tornarem coerentes com a realidade exterior e para suprirem as suas inerentes defasagens de harmonização ao contexto. As opções do desenvolvedor para conceber um modelo cristalizam uma visão de mundo que pode se revelar errônea e até mesmo nociva à sociedade. Daí a necessidade, muitas vezes, de eliminar ou ajustar certas conjecturas do modelo.

---

[212] PINKER, Steven. **Como a mente funciona**. 3. ed. São Paulo: Companhia das Letras, 2015. Trad. Laura Teixeira Mota, p. 42-43.

Um exemplo interessante de mudança de design para proteger direitos ocorreu quanto o serviço mensageiro *WhatsApp* limitou o número de destinatários em casos de mensagens "virais", sob o argumento de que assim estaria evitando o espalhamento de *Fake News*[213] e, por consequência, garantindo a liberdade de pensamento sem oferecer perigo ao direito à informação. Segundo estimativas do próprio *WhatsApp*, essa simples mudança teria reduzido, por exemplo, em 70% as mensagens com desinformação sobre o coronavírus.[214]

Embora o *WhatsApp* não seja um algoritmo criador de decisões automatizadas, a situação ilustra bem duas coisas aplicáveis a qualquer mecanismo eletrônico estruturado para o uso interpessoal: 1º) mudanças simples no design de um mecanismo inteligente pode ter reflexos profundos no seu uso social e nas consequências disso para os direitos individuais; 2º) as hipóteses incorporadas nos modelos normalmente primam pela eficiência na função específica para que serve o mecanismo — no caso, entregar mensagens mais rapidamente para um número maior de pessoas, com fidelidade —, minimizando outras externalidades relacionadas, que, no entanto, podem se mostrar importantes nas aplicações reais.

Como já referido no capítulo referente ao direito à proteção de dados, o Design Sensível ao Valor (VSD – *Value Sensitive Design*) implica alguns compromissos para ajustes em processos e técnicas, de modo que estas se harmonizem com valores humanos, para além de sua mera funcionalidade. Firdeman & Hendry, afirmam que esses compromissos incluem: a) a proposição-chave de que a relação entre tecnologia e valores humanos é fundamentalmente interacional; b) a análise das partes interessadas, direta ou indiretamente; c) as distinções entre os valores do designer, valores explicitamente apoiados pelo projeto e valores das partes interessadas; d) os níveis de análise individual, do grupo e da sociedade; e) as investigações conceituais integrativas e iterativas, técnicas e empíricas; f) a coevolução da tecnologia e da estrutura social; e g) um compromisso com o progresso, mas não com a perfeição[215].

A abordagem do VSD é dinâmica e está, ela mesma, em evolução, para incorporar e processar as variáveis específicas de cada campo de avaliação das técnicas e de suas consequências sociais e para os valores humanos. O uso de critérios de avaliação ampliados

---

[213] SINGH , Manish. WhatsApp's new limit cuts virality of 'highly forwarded' messages by 70%. *In*: TECHCRUNCH. **TECHCRUNCH** [*S. l.*], 27 abr. 2020. Disponível em: https://techcrunch.com/2020/04/27/whatsapps-new-limit-cuts-virality-of-highly-forwarded-messages-by-70/. Acesso em: 21 jan. 2021.

[214] TOGOH , Isabel. WhatsApp Viral Message Forwarding Drops 70% After New Limits To Stop Coronavirus Misinformation. *In*: FORBES. **Forbes**. [*S. l.*], 27 abr. 2020. Disponível em: https://www.forbes.com/sites/isabeltogoh/2020/04/27/whatsapp-viral-message-forwarding-drops-70-after-new-limits-to-stop-coronavirus-misinformation/?sh=26e60cbc490d. Acesso em: 20 jan. 2021.

[215] FRIEDMAN, Batya; HENDRY, David G. **Value Sensitive Design**: shaping technology with moral imagination. Cambridge (ma): The Mit Press, 2019, p. 4-5.

serve não apenas para julgar as consequências ou os danos de uma tecnologia já em operação, mas sobretudo para influenciar o *design* de novas tecnologias, durante todo o processo de sua concepção. A LGPD adota essa visão ampliada especificamente em relação à proteção da privacidade, no art. 46, §2º, ao afirmar que as medidas de segurança, técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou qualquer forma de tratamento inadequado ou ilícito, devem ser adotadas desde a concepção do produto ou serviço até a sua execução.

O Design Sensível ao Valor deve empenhar-se também na crítica dos próprios valores humanos, em sua articulação recíproca, bem como no seu contato com elementos da filosofia, da antropologia, da psicologia, sociologia, dos estudos da interação homem-máquina. Juridicamente, a riqueza dessa abordagem permite uma apreciação mais lúcida de riscos e danos de novas tecnologias, facilitando a tomada de decisão[216], tanto pelos interessados como pelas autoridades regulatórias, e serve como poderoso vetor interpretativo que deve ser levado em conta assim no momento da formulação como da aplicação das normas jurídicas pertinentes a esse campo do interesse.

As chamadas Tecnologias de Facilitação da Privacidade (*Privacy Enhancing Technologies* – PETs) podem ser consideradas um bom exemplo de Design Sensível ao Valor. Nessas tecnologias, como explica Bruno Bioni, está embutida "a ideia de que a proteção de dados pessoais deve orientar a concepção de um produto ou serviço"[217]. Denomina-se esse tipo de solução como *privacy by design*-PbD.

Há inúmeros exemplos de medidas de *privacy by design*[218]. Visando à assegurar a inviolabilidade da privacidade, a criptografia é uma ferramenta de PbD. Na mesma linha de proteção da intimidade, a anonimização dos dados pessoais também é uma técnica de *privacy by design*. Por ela, evita-se a associação dos dados ao seu titular. A navegação anônima é outro mecanismo importante de proteção da privacidade, com grande semelhança com o processo de anonimização.

Ann Cavoukian, uma das pioneiras no estudo da *privacy by design*, afirma que a necessidade desse tipo de solução decorre da observação de que respostas meramente regulatórias não têm a eficácia necessária para proteger a privacidade no ambiente da internet. Segundo ela, a PbD busca uma funcionalidade total das tecnologias, em que elas assegurem não

[216] FRIEDMAN; HENDRY, op.cit.
[217] BIONI, op. cit., p. 167.
[218] BIONI, op. cit., p. 168.

apenas a privacidade do usuário, mas igualmente a sustentabilidade das organizações que tratam os dados. A *privacy bv design* estende-se sobre três aplicações principais: 1) sistemas de Tecnologia da Informação; 2) modelos de negócio responsáveis; e 3) projetos físicos e de infraestrutura de rede. A articulação desses elementos, respeitando tanto a dignidade dos usuários como a sustentabilidade das empresas, é que pode assegurar a implementação de uma atmosfera sadia na rede, para além do consentimento e da regulação.[219]

O art. 49 da LGPD determina que: "Os sistemas utilizados para o tratamento de dados pessoais devem ser estruturados de forma a atender aos requisitos de segurança, aos padrões de boas práticas e de governança e aos princípios gerais previstos nesta Lei e às demais normas regulamentares." Nesse ponto, a lei consagra um direito ao *design* adequado. O art. 51 da LGPD, por sua vez, afirma: "A autoridade nacional estimulará a adoção de padrões técnicos que facilitem o controle pelos titulares dos seus dados pessoais".

Pensando em termos de tratamento de dados por computadores eletrônicos — únicos mecanismos que de fato produzem decisões automatizadas —, pode-se inferir que o *design* decorre tanto da infraestrutura física dos elementos que arquivam e permitem o trânsito dos dados pessoais (em um sentido amplo, todo o *hardware* e a arquitetura das redes), como dos sistemas, isto é, dos modelos que processam os dados pessoais (ou seja, dos *softwares*).

O direito ao design adequado está ligado a todo esse conjunto de insumos necessários para a produção da decisão automatizada, mas apenas surge se e quando o titular ou algum legitimado para a defesa de direitos coletivos demonstra a ocorrência de um dano concreto ou de um risco plausível à proteção de dados pessoais. Em outras palavras, o direito ao design adequando não implica que os titulares possam subordinar os modelos de negócios às suas próprias concepções de proteção de dados, mas sim que os controladores precisam levar em conta, em sua livre criação de riquezas por mecanismos inteligentes, os danos e ameaças concretas à proteção de dados dos titulares.

Dificilmente se pode cogitar de uma situação em que o design ofenda especificamente apenas um ou poucos titulares de dados. Pela própria disposição e ordem das tecnologias que geram decisões automatizadas, normalmente pensadas em grande escala, a violação de direitos pelo design tende a ser coletiva. Daí porque o direito ao design adequado frequentemente se mostrará como um direito coletivo (difuso, ou coletivo em sentido estrito). Mas não se pode excluir a hipótese de ele surgir como um direito individual homogêneo.

---

[219] PRIVACY BY DESIGN (Canadá). **The 7 Foundational Principles**. 2011. Disponível em: https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf. Acesso em: 23 out. 2020.

O direito ao design adequado tanto pode ser invocado pelos mecanismos extrajudiciais próprios, tais como demandas perante o próprio controlador, como perante a ANPD, como pode ser objeto de judicialização pelos meios legalmente previstos — neste caso, mais naturalmente por ações coletivas. É fundamental, em qualquer caso, que algum dano ou ameaça à proteção de dados seja ligada ao *design*, em termos objetivos e bem fundamentados.

4.2.2 Direito à confirmação da existência do tratamento (LGPD, art. 18,I c/c art. 19)

Este direito represente uma espécie de interpelação. O titular pode suspeitar, por algum motivo, da existência de tratamento de seus dados pessoais, para fins de produção de decisão automatizada. Em tal caso, tem o direito de requisitar — e o uso desse verbo não deixa dúvidas de que se trata de um direito potestativo — do controlador que afirme ou negue a existência do tratamento.

Na LGPD, a abrangência da definição de tratamento é muito ampla, por isso que qualquer uma das atividades englobadas nessa definição deve ser tida como tratamento e, portanto, confirmada ao titular, em caso de pedido nesse sentido. Com efeito, diz o art. 5º, X da LGPD que tratamento consiste em:

> (...) toda operação realizada com dados pessoais, como as que se referem a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração.

No caso específico de decisão automatizada, a sua produção envolve necessariamente o acesso e o processamento de dados, mas não ficam excluídas as demais operações.

O direito à confirmação da existência de tratamento será exercido mediante requerimento expresso do titular ou de representante legalmente constituído, dirigido ao agente de tratamento (LGDP, art. 18, §3º). Tal requerimento deve ser atendido pelo controlador sem custos para o titular, nos prazos e nos termos previstos em regulamento (LGPD, art. 18, §5º).

O controlador pode responder ao requerimento de forma imediata e simplificada, mediante um formulário eletrônico — inclusive pode ser uma resposta automatizada, à falta de proibição legal, desde que tal resposta avalie concretamente o requerimento, vedado o uso de resposta-padrão negativa para todos os requerimentos, porquanto isso implicaria a supressão prática do direito à confirmação.

É franqueada ao controlador também a possibilidade de responder ao requerimento por meio de declaração clara e completa, que indique a origem dos dados, a inexistência de registro, os critérios utilizados e a finalidade do tratamento, observados os segredos comercial e

industrial, fornecida no prazo de até 15 (quinze) dias, contado da data do requerimento do titular (LGPD, art. 18, II).

A resposta deve ser oferecida em formato que favoreça o acesso aos dados pelo titular (LGPD, art. 19, §§1º e 2º) — aqui uma manifestação setorial do *design* adequado. Em caso de tratamento baseado no consentimento prévio do titular ou em contrato, o titular poderá solicitar cópia eletrônica integral de seus dados pessoais, observados os segredos comercial e industrial, nos termos de regulamentação da autoridade nacional, em formato que permita a sua utilização subsequente, inclusive em outras operações de tratamento (LGPD, art. 19, §3º).

A ANPD pode alterar os prazos de reposta para setores específicos, que trabalhem com o tratamento de dados (LGPD, art. 18, §4º).

A LGPD não esclarece qual a solução para o caso de omissão do controlador em responder ao pedido de esclarecimento. Pode-se cogitar de medidas indutivas e sancionatórias, tais como advertências, multas simples, multas diárias, e interdições temporárias ou definitivas do tratamento de dados, a cargo da ANPD (LGPD, art. 52), sem prejuízo de outras sanções previstas no Código de Defesa do Consumidor e outras leis setoriais pertinentes. Em todo caso, não havendo a confirmação do tratamento, pode-se cogitar também do uso do *habeas data* para suprir a omissão (art.7º, I, Lei 9.507/97).

No caso das decisões automatizadas, o direito à confirmação pode-se manifestar também como o direito de saber se algum processo automatizado de perfilização ou de decisão foi adotado relação ao titular. É importante para o titular ter essa informação porque será a partir dela que ele poderá exercer outros direitos, tais como o direito de pedir a revisão da decisão automatizada, ou o direito de pedir explicações sobre a formação da decisão (LGPD, art. 20).

4.2.3 Direito de opor-se ao tratamento (art. 18, VIII e §2º, LGPD)

O tratamento de dados pessoais, inclusive para a formação de decisões automatizadas, apenas pode ocorrer nas hipóteses em que a lei assim o autorize. Os instrumentos de criação de decisões automatizadas, em especial os mecanismos de aprendizado de máquina, dependem fundamentalmente do processamento de dados para o seu funcionamento. Logo, é indispensável, para que a formação da decisão automatizada seja juridicamente hígida, que o tratamento subjacente tenha base legal.

A LGPD estabelece uma grande divisão entre as formas de legitimação do tratamento de dados: 1º) o tratamento, para uma finalidade específica, baseado no consentimento livre, informado e inequívoco do titular; e 2º) o tratamento independente do consentimento do titular,

também para finalidades específicas, nos casos legalmente previstos (LGPD, art. 7º, II a X; art. 11, II; art. 14, §3º; art. 26, §1º, I, III, IV e V).

Evidentemente, só se pode cogitar do direito de opor-se ao tratamento na segunda situação. Se o tratamento foi baseado em consentimento e o titular não deseja que ele continue, então o que pode ocorrer é o direito de revogar o consentimento (LGPD, art. 9º, §2º), não o direito de opor-se ao tratamento.

O direito de opor-se ao tratamento, no caso de decisões automatizadas, traduz-se no direito de pedir que cesse a produção desse tipo de decisão, em relação ao titular, quando ocorrer dano ou ameaça a algum interesse seu, juridicamente protegido.

Por estarem baseadas em inferências estatísticas e outras operações matemáticas, as decisões automatizadas implicam perigos relacionados à discriminação algorítmica, excesso de generalização, criação de estereótipos, desprezo ao contexto, minimização de particularidades concretas. Desse modo, elas representam alto risco para os direitos fundamentais do titular e, mesmo que possam ser legalmente produzidas sem o consentimento deste em situações legalmente autorizadas, devem ser sempre ostensivas e sujeitas ao escrutínio do próprio titular. O direito de oposição ao tratamento, portanto, funciona como um meio de assegurar ao titular algum poder de fiscalização sobre os tratamentos automatizados que supostamente não dependem do seu consentimento. É uma forma de assegurar ao titular o poder de impedir o tratamento clandestino dos seus dados pessoais em modelos de perfilização e de decisão automatizada.

Pode suceder de haver divergência, entre o controlador e o titular, a respeito da possibilidade ou não do tratamento dos dados sem o consentimento do titular. Nesses casos, o direito de opor-se normalmente só será eficazmente exercido ou perante a ANPD, ou por via judicial. Mas, no geral, o controlador deverá respeitar o direito de oposição do titular, se ele tiver base legal.

Um caso particularmente delicado é aquele em que o controlador alegue "legítimo interesse" para efetuar o tratamento (LGPD, art. 10). Nessa hipótese, as operações devem ser rigorosamente registradas (LGPD, art. 37), e, além do mais, a ANPD pode exigir Relatório de Impacto à Proteção de Dados Pessoais – RIPDP, sem prejuízo do segredo industrial ou comercial (LGPD, art. 10, §3º). Acresce que a oposição ao tratamento, nesta situação, naturalmente induzirá o debate sobre a existência ou não de legítimo interesse do controlador que justifique a dispensa do consentimento do titular — o que não será uma questão simples, dado que o legislador utilizou-se aqui de conceitos jurídicos indeterminados. A tendência, em

casos assim, deverá ser de que a concretização do direito de oposição se resolva no âmbito judicial.

4.2.4 Direito de acesso aos dados pessoais e aos compartilhamentos (art. 6º, IV e VI c/c art. 9º e art. 18, II e VII, LGPD)

A opção por utilizar dados pessoais para gerar decisões automatizadas decorre do modelo de negócios de cada empresa, sendo progressivamente mais comum na economia digital e nas *startups*, como mecanismo de redução de custos e de ampliação de ganhos em nichos do mercado.

Os governos também têm lançado mão de técnicas de aprendizado de máquina para criar modelos para apreciação de pedidos de benefícios sociais, no campo da vigilância e da fiscalização de atividades, no campo tributário, e em muitas áreas em que atividades repetitivas e de processamento maciço de dados podem ser assumidas por máquinas.

Assim, o compartilhamento de dados entre empresas, ou entre os governos e empresas, é uma medida de grande relevância porque potencializa os modelos de aprendizado de máquina. Ao oferecer dados abundantes e diversificados aos modelos, os compartilhamentos permitem que as máquinas engendrem soluções novas e mais eficientes para problemas logísticos, mercadológicos, de *marketing*, de segurança pública, de atendimento à saúde, etc.

Acontece que o compartilhamento aumenta os riscos à privacidade e à autodeterminação informativa dos titulares, porquanto os dados pessoais, quando reunidos a outros, associados a outras finalidades, oferecem maior grau de conhecimento sobre o titular, suas características e seus hábitos. De fato, a privacidade é inversamente proporcional à quantidade de dados em poder de terceiros, e diretamente proporcional à dispersão desses dados por bancos de dados diferentes. Quando se aumenta a quantidade de dados num só banco de dados, a privacidade do titular decai duplamente: por um lado, o controlador passa a ter mais dados sobre o titular; por outro, a maior diversidade dos dados oferece um campo de visão mais amplo sobre o titular.

É fundamental, por isso, que o titular tenha pleno acesso aos seus próprios dados, que estejam em poder do controlador, de modo a poder conhecer e controlar a produção de conhecimento sobre si mesmo. Trata-se de um pressuposto para a autodeterminação informativa. A gestão dos dados pessoais por controladores particulares ou mesmo pelos governos não pode se dar de modo secreto para o titular e, ademais, somente pode ser feita nas hipóteses e nos limites legalmente previstos.

O direito ao acesso aos dados já estava previsto, ainda numa perspectiva das tecnologias analógicas, na Constituição Federal de 1988. A primeira hipótese de cabimento do *habeas data*

(CF, art. 5º, LXXII, "a") nada mais é que uma consagração do direito de acesso aos dados, embora neste caso pela via judicial:

> LXXII - conceder-se-á *habeas data:* a) para assegurar o conhecimento de informações relativas à pessoa do impetrante, constantes de registros ou bancos de dados de entidades governamentais ou de caráter público; (...)

Sendo mais preciso, o certo é dizer-se que a Constituição garantiu o acesso às informações sobre o impetrante, e não apenas aos dados. Como já explicado em passagens anteriores, a informação decorre de dados analisados e processados. Ou seja, o *habeas data* confere ao impetrante o direito de saber não apenas que dados estão em poder do administrador do banco de dados, mas também que tipo de informação tem sido gerada pelo controlador a partir desses dados, o que inclui logicamente o direito de saber em que finalidades terão sido aplicados os dados e com quem eles terão sido compartilhados. Nessa mesma linha, o que faz o art. 9º da LGPD é dar ao titular o direito de acesso às informações produzidas sobre si, e não apenas aos seus dados, *verbis*:

> Art. 9º O titular tem direito ao acesso facilitado às informações sobre o tratamento de seus dados, que deverão ser disponibilizadas de forma clara, adequada e ostensiva acerca de, entre outras características previstas em regulamentação para o atendimento do princípio do livre acesso:
> I - finalidade específica do tratamento;
> II - forma e duração do tratamento, observados os segredos comercial e industrial;
> III - identificação do controlador;
> IV -informações de contato do controlador;
> V - informações acerca do uso compartilhado de dados pelo controlador e a finalidade;
> VI - responsabilidades dos agentes que realizarão o tratamento; e
> VII - direitos do titular, com menção explícita aos direitos contidos no art. 18 desta Lei.

No contexto das decisões automatizadas, o direito de acesso às informações é essencial e, em parte, se confunde com o direito à explicação, que será visto mais adiante. Sendo certo que os modelos inteligentes produzem conhecimento sobre o titular e, a partir desse conhecimento, tomam decisões que geram algum impacto significativo sobre os seus direitos, é indispensável que o titular possa saber que conclusões têm sido extraídas sobre si dos seus dados, inclusive para que possa eventualmente contestar essas conclusões (direito de correção), bem como pedir explicações (direito à explicação), ou mesmo a revisão da decisão automatizada (direito de revisão). A operacionalidade de vários direitos, portanto, articula-se com o direito de acesso às informações, sendo ele pressuposto lógico para o exercício de outras prerrogativas do titular. Não é por outra razão que a LGPD eventualmente trata o direito de

acesso como princípio do livre acesso ou princípio da transparência[220], de modo a permitir ao intérprete margem de avalição em cada caso desse vetor hermenêutico, o que é particularmente importante para acomodar avanços tecnológicos não totalmente previstos pelo legislador e a preservação do segredo comercial ou industrial que está na base do modelo de aprendizado de máquina.

O direito de acesso às informações está igualmente previsto na Lei do Cadastro Positivo (Lei 12.414/2011, art. 5º, II), especificamente no que diz respeito aos cadastros de crédito, mas com instrumentos que podem analogicamente ser invocados para outros tipos de bancos de dados e aplicações:

> Art. 5º São direitos do cadastrado:
> (...)
> II - acessar gratuitamente, independentemente de justificativa, as informações sobre ele existentes no banco de dados, inclusive seu histórico e sua nota ou pontuação de crédito, cabendo ao gestor manter sistemas seguros, por telefone ou por meio eletrônico, de consulta às informações pelo cadastrado; (Redação dada pela Lei Complementar nº 166, de 2019) (Vigência)
> (...)
> IV - conhecer os principais elementos e critérios considerados para a análise de risco, resguardado o segredo empresarial;
> V - ser informado previamente sobre a identidade do gestor e sobre o armazenamento e o objetivo do tratamento dos dados pessoais; (Redação dada pela Lei Complementar nº 166, de 2019) (Vigência)
> (...)
> VII - ter os seus dados pessoais utilizados somente de acordo com a finalidade para a qual eles foram coletados.

O direito de acesso deve ser exercido e atendido da mesma forma que se exerce e se atende o direito à confirmação sobre tratamento automatizado (LGPD, art. 19). A ANPD poderá dispor sobre a padronização de mecanismos técnicos para facilitar o acesso aos dados pelo titular (LGPD, art. 40).

Importa enfatizar que o direito de acesso pertence ao titular, podendo ser exercido por ele ou por procurador contra quem detém as informações, não se confundindo com o direito de acesso à informação a que alude o art. 5º, XXXIII da Constituição Federal, regulamentado pela

---

[220] LGPD, art. 6º, IV e VI: "Art. 6º As atividades de tratamento de dados pessoais deverão observar a boa-fé e os seguintes princípios:

(...) IV - livre acesso: garantia, aos titulares, de consulta facilitada e gratuita sobre a forma e a duração do tratamento, bem como sobre a integralidade de seus dados pessoais;(...)
VI - transparência: garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial;(...)"

Lei de Acesso à Informação (Lei 12.527/2011). Este último é um direito difuso, que tem por objetivo assegurar a publicidade e transparência dos atos do Poder Público. O primeiro é direito individual que visa a assegurar a autodeterminação informativa.

Diferentemente das informações relacionadas ao Poder Público, aquelas que decorrem de dados pessoais são, para terceiros não autorizados, presumivelmente sigilosas, excetuando-se os dados tornados manifestamente públicos pelo próprio titular (LGPD, art. 7º, §4º). Os agentes de tratamento devem adotar medidas de segurança, técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou qualquer forma de tratamento inadequado ou ilícito (LGPD, art. 46). Donde se também se conclui que o direito de acesso a que se alude aqui pertence apenas ao titular ou a alguém por ele autorizado.

4.2.5 Direito de correção (art. 18, III, LGPD)

Pela própria lógica de funcionamento dos modelos, pode-se dizer que o resultado do processo de produção da decisão automatizada está diretamente ligado aos *dados de entrada* e indiretamente aos dados de treinamento. Os dados de saída são a própria decisão automatizada, de modo que a correção deles se resolverá pelo chamado direito à revisão, adiante analisado.

O direito de correção, em se tratando de decisão automatizada, fica melhor colocado em relação apenas aos dados de entrada. Eventual pedido de correção dos dados de treinamento e validação do modelo estará mais ligado ao direito ao design adequado, visto como tais dados normalmente pertencem a grande número de titulares e a sua correção acaba sendo um problema de interesse difuso.

O uso mais evidente do direito de correção está ligado à retificação de informações errôneas a respeito do titular. Isso é tanto mais importante, no ambiente digital, porque muitos dados do titular são coletados informalmente e em massa pelos modelos, considerando, por exemplo, o IP de uma máquina ligado ao titular (que também pode ser usada por outras pessoas, eventualmente), ou certo histórico de navegação, muitas vezes não havendo certeza sobre a autenticidade das informações, visto como as decisões automatizadas são concebidas com amplo uso de estatísticas e probabilidades, o direito de correção pode retificar desvios e erros pró

A aplicação mais evidente do direito de correção está ligada à retificação de informações errôneas a respeito do titular. Isso é tanto mais importante, no ambiente digital, porque muitos dados do titular podem ser coletados sem estruturação prévia e em massa, por agentes humanos

ou sensores. Esse modo de coleta de dados está sujeito a algumas anomalias, como explicam Leandro Castro e Daniel Ferrari:

> Quase todas as bases de dados reais apresentam algum tipo de anomalia, que pode ser causada por fatores como atividades maliciosas (por exemplo, furtos, fraudes, intrusões, etc.), erros humanos (por exemplo, erros de digitação ou de leitura), mudanças ambientais (por exemplo, no clima, no comportamento de usuários, nas regras ou nas leis do sistema, etc.), falhas em componentes (por exemplo, peças, motores, sensores, atuadores, etc. entre outros.[221]

Outro ponto que pode dar ensejo à aplicação do direito à correção decorre da circunstância de que os dados, na internet, são coletados muitas vezes no pressuposto de que certo IP (*Internet Protocol*) está associado a certa pessoa. Ora, tanto por motivos de fraude, informalidade do cadastro, como de compartilhamento de aparelhos, pode suceder de um aparelho ser utilizado por pessoa diversa daquela em nome de quem está registrado. E isso pode levar à errônea associação de certo conjunto de dados a determinada pessoa, que na verdade não produziu aqueles dados. Embora haja algumas tentativas de autenticar movimentações da internet, como senhas e cadastros biométricos, ainda se mostra possível, à falta de um heurema com fé pública presumida, que suceda de o alegado titular repudiar os dados como seus. E isso pode ter efeitos importantes no processo de formação de uma decisão automatizada que leve em conta características pessoais do titular.

Também pode-se pedir a correção de dados para atualizá-los. Muitas vezes, o modelo pode usar dados desatualizados de algum titular para produzir decisões automatizadas a respeito dele, o que pode gerar danos, por meio de inferências defasadas em relação ao contexto do momento da decisão. Imagine-se um mecanismo automático de avaliação de risco de crédito que leve em conta um débito prescrito, ou uma dívida já paga, ou uma obrigação judicialmente anulada. Em todos esses casos, o direito de correção pode ser utilizado pelo titular para suprimir ou retificar os dados em poder do controlador.

Mas, bem entendido, é preciso que os dados a serem corrigidos estejam de fato equivocados. Se os dados estão corretos e o titular, sob o argumento de pedir a correção, requer que eles sejam alterados para criar uma situação falsa, isso não está abrangido no direito de correção. Também não está no direito à correção o chamado "direito ao esquecimento". Este tem outro fundamento e visa a fim diverso do direito à correção, sendo normalmente realizado pela via judicial, a partir do Marco Civil da Internet, como ressalta Ricardo Villas Bôas Cueva:

---

[221] CASTRO, Leandro Nunes de; FERRARI, Daniel Gomes. **Introdução à mineração de dados**: conceitos básicos, algoritmos e aplicações. São Paulo: Saraiva, 2016, p. 270.

> O novo diploma [Marco Civil da Internet] introduziu a reserva de jurisdição, ou seja, a responsabilidade do provedor pela retirada do conteúdo infringente tem início a partir da notificação judicial, que deve determinar claramente o conteúdo a ser removido, mediante a indicação específica da URL, sob pena de nulidade.[222]

Eventualmente, o próprio titular pode ser o autor do erro nos dados, em casos de lapsos no preenchimento de formulários ou resposta a questionários. Observada a boa-fé, também nesses casos é possível invocar o direito à correção dos dados, com base no princípio da qualidade dos dados (LGPD, art. 6º, V), que garante aos titulares, a exatidão, clareza, relevância e atualização dos dados, de acordo com a necessidade e para o cumprimento da finalidade de seu tratamento.

4.2.6 Direito à anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou tratados em desconformidade (art. 18, IV, LGPD)

Enquanto os dados podem ser ligados a uma pessoa natural, eles são dados pessoais. Neste caso, os processos de proteção passam pela criação de barreiras físicas ou controles de acesso a eles (encriptação, por exemplo). Se os dados, porém, deixam de poder ser associados ao seu titular, segundo técnicas computacionais razoavelmente conhecidas, então eles perdem o caráter de dado pessoal e, assim, a anonimização é uma forma de proteção pelo empobrecimento informacional dos dados[223]. A eliminação dos dados, por sua vez, idealmente é uma proteção absoluta, visto como os dados desaparecem em tal caso e, por consequência, não podem mais ser usados para nenhuma finalidade. Logo, não podem ser violados.

Os dados anonimizados estão, a princípio, fora da disciplina da LGPD, porque deixam de ser dados pessoais, já que não podem mais revelar nenhuma informação sobre os seus titulares. O mesmo se passa naturalmente com os dados eliminados.

Os modelos de decisão automatizada, para aumentarem a sua acurácia, normalmente dependem de uma grande massa de dados pessoais. Nesse contexto, o controlador pode tratar dados pessoais muito além daqueles estritamente necessários para a rodagem do modelo. Esses dados em excesso, pela escala em que são tratados, acabam gerando pequenos incrementos de produtividade que representam diferenças de ganho para empresas. Por isso há estímulo econômico e concorrencial para o uso de dados excessivos; é justamente neles que pode surgir um *insight* comercial promissor e exclusivo.

---

[222] CUEVA, Ricardo Villas Bôas. Proteção de Dados Pessoais e Direito ao Esquecimento. In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle, p.637.

[223] Cf.: NELSON, Gregory S.. Practical Implications of Sharing Data: A Primer on Data Privacy, Anonymization, and De-Identification. **As Global Forumproceedings**, p. 2-23, abr. 2015. Disponível em: https://www.pharmasug.org/proceedings/2016/IB/PharmaSUG-2016-IB06.pdf. Acesso em: 20 jan. 2021.

Como já referido em passagens anteriores, o aumento do número de dados pessoais em poder do controlador implica o aumento do conhecimento dele sobre os titulares e, consequentemente, a ampliação do poder de predição probabilística sobre comportamentos futuros desses mesmos titulares. Como diz Cathy O´Neil, os modelos, especialmente aqueles que ela chama de "armas de destruição matemática" (*Weapons of Math Destruction*), são "caixas-pretas que absorvem dados e cospem conclusões"[224].

Por meio do direito de acesso aos dados o titular pode saber quais dos seus dados pessoais estão sendo utilizados para a produção da decisão automatizada. Se houver algum que pareça excessivo, desnecessário ou tratado fora das hipóteses legalmente previstas, o titular pode pedir a anonimização ou eliminação desses dados.

Também se pode pedir a anonimização, bloqueio ou eliminação de dados tratados em desconformidade com a lei (por exemplo, sem o consentimento do titular, nos casos em que tal consentimento é exigível).

Na hipótese de decisões automatizadas, se o titular tomo conhecimento do tratamento apenas depois do resultado do tratamento, isto é, depois da decisão automatizada ter sido produzida, os direitos previstos no art. 20 da LGPD (direito à explicação e direito à revisão) podem ocupar o espaço que é reservado aos direitos vistos até aqui. Em tese, esses dois direitos exercem funções de conhecimento, correção e anonimização. Tal circunstância não torna inúteis os direitos de confirmação do tratamento, acesso, correção, anonimização e eliminação de dados impertinentes, pois tais direitos mostram-se indispensáveis para atacar tratamentos ilegítimos que ainda não produziram a decisão automatizada, embora se dirijam a isso.

O direito à anonimização, bloqueio ou eliminação de dados excessivos visa a assegurar a minimização do uso de dados. A minimização do uso de dados é, na verdade, para além de um direito do titular, um princípio informativo da LGPD (art. 6º, III: "necessidade: limitação do tratamento ao mínimo necessário para a realização de suas finalidades, com abrangência dos dados pertinentes, proporcionais e não excessivos em relação às finalidades do tratamento de dados"). Tal princípio tem inspiração no GDPR (art. 5º, 1, "c": "Os dados pessoais devem ser: adequados, pertinentes e limitados ao que é necessário relativamente às finalidades para as quais são tratados ('minimização dos dados')".

[224] O´NEIL, op. cit., p. 208, tradução nossa. Texto original: "(...) a black box that takes in data and spits out conclusions"

Um exemplo, lembrado por Lydia F. de La Torre[225], de um conjunto de dados excessivo pode se dar quando uma agência de recrutamento coloca trabalhadores em uma variedade de empregos. Envia aos candidatos um questionário geral, que inclui perguntas específicas sobre condições de saúde que só são relevantes para ocupações manuais específicas. Seria irrelevante e excessivo obter tais informações de um indivíduo que se candidatou a um cargo de escritório.

Dados desatualizados também podem gerar problemas para o titular. A princípio, pode-se pensar que, em tema de decisão automatizada, o próprio controlador tem o máximo de interesse em atualizar os dados pessoais com que trabalha, com o objetivo de aumentar a acurácia e a confiabilidade do modelo. Isso pode ser verdadeiro em muitos casos nos quais o ganho econômico depende da atualização dos dados, mas se, mesmo desatualizados, os dados geram riqueza para o controlador, não há estímulo econômico para a atualização e, assim, pode ocorrer de o modelo trabalhar de modo rentável com dados desatualizados. Cathy O'Neil cita o exemplo de um modelo de seleção de trabalhadores de baixa renda para vagas de emprego. Segundo ela, esse tipo de modelo trabalha com verdadeiros "rebanhos"[226], não havendo uma preocupação maior em atualizar dados, "a menos que algo dê errado", ou seja, que as contratações venham a apresentar algum problema específico.

4.2.7 Direito de consentir ou não no tratamento (arts. 7°, 11 e 14, LGPD) e Direito de revogar o consentimento (art. 8º, §5º, art. 15, III c/c art. 18, XI, LGPD)

O consentimento do titular para o tratamento, compartilhamento e transferência internacional de dados ainda é um importante fundamento para a legitimação do processo de tratamento de dados. Porém, como demonstrado no capítulo 2 desta pesquisa, há diversas fragilidades que debilitam o consentimento em algumas circunstâncias específicas.

O consentimento, que foi definido como "manifestação livre, informada e inequívoca pela qual o titular concorda com o tratamento de seus dados pessoais para uma finalidade determinada" (LGPD, art. 5º, XII), é normalmente exigível para o tratamento de dados pessoais (LGPD, art. 7º, I e V), com exceção de algumas situações de interesse público (LGPD, art. 7º, II, III, IV, VI), ou de interesse presumível do próprio titular dos dados (LGPD, art. 7º, VII e VIII), ou no "legítimo interesse" do controlador ou de terceiro, em especial para a proteção do crédito (LGPD, art. 7º, IX e X).

---

[225] DE LA TORRE , Lydia F. What is "data minimization" under EU Data Protection Law?. *In*: MEDIUM. **Medium**. [*S. l.*], 22 jan. 2019. Disponível em: https://medium.com/golden-data/what-is-data-minimization-under-eu-data-protection-law-b0e30fbb856e. Acesso em: 27 jan. 2021.

[226] O´NEIL, op. cit., p. 111.

O papel do consentimento no tratamento de dados gravita entre dois polos ideológicos: a) o polo americano, que confere maior papel ao consentimento e, portanto, à autogestão dos dados pessoais pelo titular; e b) o polo europeu, que cria restrições legais coercitivas para o tratamento, a despeito do consentimento[227].

Como mostrado no Capítulo 2, a doutrina atual levanta uma série de questões a respeito da efetiva proteção conferida pelo consentimento, realçando a incapacidade prática de o cidadão comum, no ecossistema da internet, gerir seus próprios dados em muitas situações. Bruno Bioni, por isso, fala de uma "travessia do protagonismo do consentimento". Segundo esse autor, *verbis*:

> Em que pese ter sempre havido dúvidas em torno da racionalidade e do poder de barganha dos titulares de dados pessoais para que eles empreendessem um controle efetivo sobre seus dados pessoais, o consentimento permaneceu sendo o elemento nuclear da estratégia regulatória da privacidade informacional. (...) Tem-se, assim, um quadro regulatório encapsulado por uma compreensão reducionista do conteúdo a que se deve referir autodeterminação informacional que, passadas mais de duas décadas, não mais se ajusta ao contexto subjacente dos dados pessoais como ativo econômico em constante circulação. (...) Nessa conjuntura, faz-se necessário reavaliar tal estratégia regulatória e a própria compreensão do conteúdo do que é autodeterminação informacional. Devem-se canalizar esforços para identificar a problemática em torno de uma estratégia regulatória e dogmática anacrônica pensada nos anos 1980, que enfrente uma demanda social dos anos 2000 (...)[228]

Em todo caso, o certo é que o consentimento é uma garantia mínima do indivíduo contra o uso abusivo dos seus dados, nas situações que a regulação não apanha o fato diretamente. Na verdade, o consentimento como elemento de legitimação das ações de terceiros sobre a vida, os bens e a liberdade de outrem está no cerne da vida jurídica das sociedades democráticas, a tal ponto que as teorias políticas clássicas sobre a própria formação das sociedades modernas apegam-se à alguma faceta do contratualismo, sempre no pressuposto de que os indivíduos são seres racionais e auto-interessados, estando, por conseguinte, na melhor posição para avaliarem a conveniência das trocas econômicas, políticas e sociais nas quais se envolvem.[229] Embora esse pressuposto não seja verdadeiro em muitas circunstâncias, o consentimento tem ainda um papel de destaque na legitimação do tratamento de dados pessoais.

Visando a tornar o consentimento mais consistente, a lei exige que o agente de tratamento use de boa-fé e se atenha à finalidade específica considerada no momento do

---

[227] Cf.: SOLOVE, Daniel. Autogestión de la privacidad y el dilema del consentimiento. **Revista Chilena de Derecho y Tecnología**, n. 3, p. 11-47, 23 jan. 2014. Universidad de Chile. http://dx.doi.org/10.5354/0719-2584.2013.30308. Disponível em: https://rchdt.uchile.cl/index.php/RCHDT/article/view/30308. Acesso em: 31 jan. 2021.

[228] BIONI, op. cit., p. 130-131.

[229] CHIAPPIN, J. R. N.; LEISTER, Carolina. O contratualismo como método: política, direito e neocontratualismo. **Revista de Sociologia e Política**, v. 18, n. 35, p. 09-26, fev. 2010. FapUNIFESP (SciELO). Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-44782010000100002&lng=pt&tlng=pt. Acesso em: 27 jan. 2021.

consentimento (LGPD, arts. 6º, I), à adequação entre o tratamento de dados e o contexto (LGPD, art. 6º, II), à eficiência do tratamento, com o uso mínimo possível dos dados (LGPD, art. 6º, III), à qualidade dos dados considerados (LGPD, art. 6º, V). E, mesmo em dados pessoais tornados públicos, para os quais se dispensa a manifestação de consentimento para o tratamento, a finalidade deve ser considerada (LGPD, art. 7º, §§3º, 4º e 7º). Autorizações genéricas de tratamento, por sua vez, são tidas como nulas (LGPD, art. 7, §4º).

Se o caso é de tratamento que dispensa a nução do titular (LGPD, art. 7º, II a X e §4º; art. 11, II; art. 14, §3º), não há falar em direito ao consentimento. Todavia, em todos os demais casos, o consentimento informado para finalidade específica é exigível.

No caso das decisões automatizadas, o consentimento mostra-se particularmente relevante. Como já mencionado em passagens anteriores, as decisões automatizadas partem de dados pessoais, que são processados em articulação com dados de outros titulares, para gerar predições e escolhas automatizadas, baseadas em estereótipos de condutas ou características pretéritas. Assim, o uso desses mecanismos implica julgamentos *a priori* sobre os titulares, e, por isso, devem ser claramente informados e previamente autorizados, a não ser que haja autorização legal explícita para o tratamento de dados sem o consentimento do titular.

Embora a LGPD (art. 20), ao contrário do GDPR (art. 22), não tenha previsto expressamente o direito de não ficar sujeito a decisão automatizada, é certo que esse direito existe entre nós, naqueles casos em que o tratamento dos dados depende de consentimento. É que a decisão automatizada nasce de um tratamento de dados, de sorte que, se este precisa ser consentido, aquela também fica sujeita ao consentimento.

É importante ressaltar que, mesmos nos casos em que a lei dispensa o consentimento, isso não exonera os agentes de tratamento das demais obrigações previstas na LGPD, especialmente da observância dos princípios gerais e da garantia dos direitos do titular (LGPD, art. 7º, §6º).

Entre os princípios, merece destaque o da finalidade. É na reorientação da finalidade que um dado inocentemente oferecido pode gerar conhecimento contra os interesses do titular. Como observam Laura Mendes e Gabriel Soares, o consentimento precisa ser avaliado a posteriori, em articulação com as "legítimas expectativas" presumíveis no momento da aceitação do tratamento de dados, *verbis*:

> Essas "legítimas expectativas" passam a ser avaliadas a partir de elementos como: (i) o contexto em que a suposta violação ocorreu (qual era o ambiente social que estruturava o fluxo de informações analisado?); (ii) os atores envolvidos (quem eram os emissores, receptores e sujeitos do fluxo de informação?); (iii) os atributos da informação analisada (com que tipo de informação se estava lidando? Informações

> médicas, bancárias, preferências pessoais?); (iv) os princípios de transmissão aplicáveis (quais eram os constrangimentos aplicáveis ao fluxo de informações analisado, ele estava condicionado à confidencialidade, reciprocidade, necessidade?).[230]

Sob a perspectiva do controlador, uma questão que surge das "legítimas expectativas" geradas pela concordância do titular com o tratamento dos seus dados liga-se ao problema da revogação do consentimento. A LGPD (art. 8º, §5º) consagra expressamente a revogabilidade do consentimento, sem indicar qualquer condição ou encargo associado a esse ato.

Embora tenha ficado expressamente ressalvado, na parte final do §5º do art. 8º da LGPD, "os tratamentos realizados sob amparo do consentimento anteriormente manifestado enquanto não houver requerimento de eliminação", não esclarece a lei qual a solução a dar para casos em que essa revogação gere danos ao controlador. Se é certo que a revogação *ad nutum* do consentimento precisa ser assegurada, tendo em conta a autodeterminação informativa do titular, não menos exato é que o controlador também precisa de garantias mínimas para desenvolver a sua atividade econômica com base nos negócios celebrados. Assim, a doutrina considera que a revogação é livre para o titular, nos casos em que o consentimento foi anteriormente dado, mas se tal revogação for abusiva e vier a gerar danos para o controlador, isso pode ensejar a responsabilização extracontratual do titular. Nesse sentido, observa Danilo Doneda:

> A eventual conduta abusiva de quem revoga o consentimento pode ensejar um dever de reparação, uma vez que essa conduta caracterize dano a quem anteriormente teria recebido a autorização para tratar os dados pessoais dessa pessoa. Essa reparação, pelos motivos já expostos, não tem caráter negocial; ela também não restringe em nenhum modo a possibilidade de revogação nem a vincula a qualquer outro ato — pois esta deve ser sempre uma faculdade de quem consente, cuja restrição implicaria diminuição de seu poder de autodeterminação.[231]

A apuração da abusividade da revogação deve ocorrer pelos meios ordinários, especialmente tendo em conta a teoria geral do abuso de direito, com mais ênfase para a verificação de eventual invocação do *venire contra factum proprium*.[232]

---

[230] MENDES, Laura Schertel; FONSECA, Gabriel Campos Soares da. Proteção de dados para além do consentimento: tendências de materialização. **In**: In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle, p.101

[231] DONEDA, op. cit., p. 305.

[232] Ibidem.

4.3 Direitos Relativos aos Resultados da Decisão Automatizada

Como já referido anteriormente, os direitos mais típicos, em casos de decisões automatizadas, são aqueles que se dirigem contra o resultado das decisões já tomadas, mais especificamente o *direito à explicação* e o *direito à revisão*.

É interessante notar que esses dois direitos apresentam notória semelhança com os fundamentos jurídicos normalmente aceitos para recorrer contra atos judiciais (esclarecimento, reforma ou anulação do julgado recorrido)[233]. Considerando-se que a doutrina aponta como razões políticas principais para a existência de recursos contra atos judiciais o controle dos juízes inferiores pelos superiores e a busca de uma solução justa, então fica claro que a semelhança não é casual.

Com efeito, também quanto às decisões automatizadas os motivos para a explicação e a revisão são dois: a) controle do funcionamento dos mecanismos automatizados; e b) busca de maior justiça nas decisões.

4.3.1 Direito à explicação

O art. 20, §§1º e 2º dispõem da LGPD:

> Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade.
> § 1º O controlador deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, observados os segredos comercial e industrial.
> § 2º Em caso de não oferecimento de informações de que trata o § 1º deste artigo baseado na observância de segredo comercial e industrial, a autoridade nacional poderá realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais.

A partir da leitura dos §§1º e 2º, acima transcritos, verifica-se que há um *direito à explicação* em casos de decisões automatizadas no Brasil. Embora se possa discutir a terminologia (*direito à informação* ou *direito à explicação*), é certo que o direito positivo nacional consagra, em favor do titular, o direito de conhecer os critérios e os procedimentos utilizados para a produção de uma decisão automatizada que afete os seus interesses.

Com explicam Souza, Perrone e Mgrani, "similarmente ao GDPR, na lei brasileira podemos fundar um direito à explicação a partir de três pontos principais: o princípio da

---

[233] MOREIRA, José Carlos Barbosa. **Comentários ao Código de Processo Civil**. 7. ed. Rio de Janeiro: Forense, 1998. Vol. V, p. 207.

transparência, o direito de acesso à informação e como um pressuposto para o exercício dos outros direitos e, particularmente, do direito a requerer revisão de decisões automatizadas."[234]

Doshi-Velez e Kortz afirmam que uma decisão torna-se explicada quando é interpretável em termos humanos e se pode responder, pelo menos, às seguintes perguntas: a) quais os critérios que levaram à decisão?; b) Alterar algum dos critérios mudaria a decisão?; e c) Por que casos semelhantes resultaram em decisões diferentes e vice-versa?[235]

A questão dos critérios utilizados é fundamental porque alguns são lícitos, outros não. Apenas mediante a explicação de quais foram os fatores utilizados pelo modelo de decisão automatizada é que se pode avaliar se o procedimento foi ou não lícito. Assim, por exemplo, se um modelo de cálculo automatizado de salários por produtividade numa empresa vier a levar em conta a circunstância de que certos empregados têm necessidades especiais, para reduzir-lhes a retribuição, então se estará utilizando elementos ilícitos na predição, visto como a Constituição Federal proíbe o uso desse fator para o cálculo do salário (CF, art. 7º, XXXI). Muitas vezes, o controlador poderá invocar o segredo comercial ou industrial para não revelar os fatores do cálculo. Neste caso, como mencionado no §2º do art. 20, caberá à Autoridade Nacional de Proteção de Dados auditar o modelo, para verificar se ele não está embutindo preconceito nos seus mecanismos de predição e decisão.

A discriminação algorítmica pode ocorrer não apenas pelo uso de fatores de decisão vedados, mas também por erro estatístico, generalização injusta ou pela limitação de direitos, conforme apontam Laura Mendes, Marcela Mattiuzzo e Mônica Fujimoto[236] . A *discriminação por erro estatístico* é aquela que decorre de algum equívoco "que seja genuinamente estatístico, abrangendo desde dados incorretamente coletados, até problemas no código do algoritmo, de modo que ele falhe em contabilizar parte dos dados disponíveis, contabilize-os de forma incorreta etc."[237] A *discriminação por generalização injusta* é a que "embora o modelo funcione bem e seja estatisticamente correto, leva a uma situação na qual algumas pessoas são equivocadamente classificadas em certos grupos"[238]. Finalmente, a *discriminação limitadora*

---

[234] SOUZA; PERRONE; MAGRANI, op.cit, p. 273.

[235] DOSHI-VELEZ, Finale *et al*. Accountability of AI Under the Law: the role of explanation. **Cornell University**, nov. 2017. Disponível em: https://arxiv.org/abs/1711.01134v1. Acesso em: 27 jan. 2021.

[236] MENDES, Laura Schertel; MATTIUZZO, Marcela; FUJIMOTO, Mônica Tiemy. Discriminação algorítmica à luz da Lei Geral de Proteção de Dados. In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle.

[237] Ibidem.

[238] Ibidem.

de direitos é aquela em que há uma relação muito estreita entre o fator empregado no algoritmo e a decisão a ser tomada, a ponto de inviabilizar o exercício de certo direito[239].

Em todos os casos de discriminação, o direito à explicação funciona como potente elemento de dissuasão e de reparação de injustiças. Afinal, é a partir das explicações sobre o modo de funcionamento do modelo que se pode desvendar eventuais discriminações, as quais muitas vezes entram no modelo de maneira não intencional.

A explicação gera um custo adicional para o uso do modelo, de tal modo que, em certos contextos de mínima lesividade ou de pouca escala do modelo, a exigência de explicação pode mostrar-se excessivamente onerosa ou até mesmo economicamente proibitiva do uso do modelo[240].

Em dadas circunstâncias, o direito à explicação pode confundir-se com o direito ao design adequado. Quando a explicação que se busca não é apenas para um caso específico, mas sim para o funcionamento geral do modelo, é mais apropriado falar-se deciframento do próprio design, e não da explicação de uma decisão[241]. E aqui surge uma questão importante, que consiste em saber se a explicação tem de ser dada em termos acessíveis para um leigo, ou se ela precisa ser técnica, abordando todos os aspectos da tomada de decisão automatizada. Segundo o que decorre do princípio da transparência (LGPD, art. 6º, VI; art. 14, §6º; art. 55-J, XIX), exceto em se tratando da relação do controlador com a ANPD, a explicação tem de se dar em termos compreensíveis para um leigo, abstraindo-se as minúcias computacionais e estatísticas do modelo. De fato, a lei (LGPD, art. 6º, VI) fala em transparência como a "garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial."

No Brasil, o direito à explicação precede a LGPD. A Lei 12.414/2011 – Lei do Cadastro Positivo (LCP), em seu art. 5º, IV, afirma que o cadastro (equivalente ao titular, da LGPD) tem direito a conhecer os principais elementos e critérios considerados para a análise de risco. A diferença é que na LCP apenas se cuida do *credit scoring* e, ademais, o direito à explicação nela referido não se dirige expressamente apenas a processos automatizados, embora na prática eles sejam os mais amplamente utilizados.

---

[239] Ibidem.

[240] Cf.: DOSHI-VELEZ, Finale *et al*. Accountability of AI Under the Law: the role of explanation. **Cornell University**, nov. 2017. Disponível em: https://arxiv.org/abs/1711.01134v1. Acesso em: 27 jan. 2021.

[241] Cf.: MULHOLLAND, Caitilin; FRAJHOF, Isabella Z. Inteligência Artificial e a Lei Geral de Proteção de Dados Pessoais: breves anotações sobre o direito à explicação perante a tomada de decisões por meio de machine learning. In: MULHOLLAND, Caitlin; FRAZÃO, Ana (org.). **Inteligência Artificial E Direito**: ética regulação e responsabilidade. São Paulo: Thomson Reuters Brasil, 2019, p. 281.

Discutiu-se na jurisprudência se a formação do *credit scoring* dependeria do consentimento do cadastrado, tendo o STJ firmado na Súmula 550 a sua posição no sentido de que o consentimento é desnecessário, haja vista a autorização expressa da lei, mas se ressaltou a explicação ao interessado como elemento de legitimação do cadastro, *verbis*:

> Súmula 550: A utilização de escore de crédito, método estatístico de avaliação de risco que não constitui banco de dados, dispensa o consentimento do consumidor, que terá o direito de solicitar esclarecimentos sobre as informações pessoais valoradas e as fontes dos dados considerados no respectivo cálculo.

4.3.2 Direito de Revisão das Decisões Automatizadas (art. 20, LGPD e art. 5º, VI da LCP)

Embora a LGPD tenha a sua fonte de inspiração no *General Data Protection Regulation-GDPR* da União Europeia, cujo art. 22 [242] expressamente prevê, em dadas condições, o direito à intervenção humana sobre decisões automatizadas, não temos dispositivo semelhante no país. Isso porque o § 3º do art. 20 da Lei nº 13.709, de 14 de agosto de 2018, alterado pelo art. 2º do Projeto de Lei de Conversão nº 7, de 2019 (MP nº 869/2018), que trouxe essa previsão[243], foi vetado pelo Presidente da República.

Nas razões do veto, o Presidente da República fez constar o seguinte:

> A propositura legislativa, ao dispor que toda e qualquer decisão baseada unicamente no tratamento automatizado seja suscetível de revisão humana, contraria o interesse público, tendo em vista que tal exigência inviabilizará os modelos atuais de planos de negócios de muitas empresas, notadamente das *startups*, bem como impacta na análise de risco de crédito e de novos modelos de negócios de instituições financeiras, gerando efeito negativo na oferta de crédito aos consumidores, tanto no que diz respeito à qualidade das garantias, ao volume de crédito contratado e à composição de preços, com reflexos, ainda, nos índices de inflação e na condução da política monetária[244].

Extrai-se das razões do veto que as decisões automatizadas presumivelmente favorecem modelos de negócio contemporâneos, que estão assentados sobretudo na inovação disruptiva, mediante a manipulação de amplas bases de dados para a criação de novas comodidades com valor econômico. Assim, a previsão legal *tout court* de revisão humana sobre qualquer decisão automatizada poderia, de fato, criar um obstáculo exagerado aos negócios. Já a ausência do

---

[242] THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION. **Regulation nº 2016/679**. On the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). [*S. l.*], 27 abr. 2016. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=PT. Acesso em: 11 jul. 2020.

[243] "Art. 20 (...)§ 3º A revisão de que trata o *caput* deste artigo deverá ser realizada por pessoa natural, conforme previsto em regulamentação da autoridade nacional, que levará em consideração a natureza e o porte da entidade ou o volume de operações de tratamento de dados."

[244] BRASIL. Presidência da República. Mensagem nº 288, de 8 de julho de 2019. **Diário Oficial da União,** Brasília, n. 130, p. 8, 9 jul. 2019. Disponível em: http://www.in.gov.br/en/web/dou/-/despachos-do-presidente-da-republica-190107781. Acesso em: 11 jul. 2020.

texto expresso, por outro lado, não parece implicar a vedação da revisão humana, mas apenas permite que a revisão possa se dar por modo automatizado também.

A leitura do dispositivo sugere que a revisão, entre nós, é de apenas uma instância, ou seja, não há previsão de recurso contra a própria revisão — sem embargo, naturalmente, do acesso à via judicial. Nesse sentido, afirmam Souza, Perrone e Magrani:

> Deve-se frisar que a regulação brasileira explicitou que a garantia será de uma revisão. Isso parece aproximar o país dos sistemas implementados no Reino Unido e na Irlanda, que, como vimos, definiram o direito de contestar, respectivamente, como direito de requerer uma reconsideração e de apelar.[245]

Outro ponto importante, decorrente da lei, e já salientado anteriormente, é que o direito à revisão apenas pode ser exercido após a decisão ter sido tomada, não havendo nenhuma obrigação para o controlador de realizar revisões *ex ante*, sem prejuízo, é claro, de eventuais auditorias da ANPD. Nesse sentido, observa Marco Almada:

> Outra consequência da escolha do legislador brasileiro por um direito à revisão é que o dispositivo da LGPD estabelece que o direito do titular de dados é uma resposta ao fato que gera a lesão ou ameaça a interesse juridicamente tutelado. Os controladores de sistemas de automação ficam, pois, desobrigados de intervenções *ex ante* ao longo do ciclo de desenvolvimento, que permitiriam o envolvimento mais direto dos titulares de dados e de suas perspectivas (...)[246]

Estão sujeitos a rever as decisões automatizadas todos os controladores privados, abrangidas as empresas públicas e sociedades de economia mista que exerçam atividade econômica (LGPD, art. 24, *caput*). As pessoas de direito público e as empresas públicas e sociedades de economia mista que executem políticas públicas, bem como os serviços notarias e de registro (LGPD, art. 23, §4º) ficam sujeitos ao regime dos arts. 23 a 32 da LGPD. Porém, como acentua Marco Almada, essa distinção de regimes não elimina o direito à revisão:

> Tal tratamento diferenciado, contudo, não exclui as decisões automatizadas que ocorrem no contexto das pessoas jurídicas de direito público do alcance do direito de revisão, uma vez que a LGPD, no artigo 23, § 3º, prevê que "[o]s prazos e procedimentos para exercício dos direitos do titular perante o Poder Público observarão o disposto em legislação específica", o que inclui não só eventuais leis especiais futuras, mas também aquelas já existentes, em especial a Lei do Habeas Data (Lei 9.507/1997), a Lei Geral do Processo Administrativo (Lei 9.784/1999) e a Lei de Acesso à Informação (Lei 12.527/2011). A disciplina do exercício dos direitos do titular, incluindo aí o direito à revisão de decisões, deve ocorrer em conformidade com esta legislação, não só no que tange às regras postas, mas também aos princípios consagrados nestes textos legais. Logo, a discussão a respeito do direito à revisão será

[245] SOUZA; PERRONE; MAGRANI, op.cit, p. 277.
[246] ALMADA, Marco. Revisão humana de decisões automatizadas. **Pósdebate 2019-USP**, São Paulo, p. 1-22, nov. 2019. p. 7.

> também relevante para as iniciativas de uso de inteligência artificial e automação na administração pública.[247]

O papel da ANPD na revisão da decisão automatizada é relevante, porque essa entidade pode ser chamada a auditar algum modelo de decisão automatizada cujo controlador alegue segredo comercial ou industrial para não dar explicações, seja sobre a própria decisão ou sobre a revisão dela. Logo, o direito à revisão, apesar de ser apenas de uma instância extrajudicial, a princípio, não se esgota totalmente entre o controlador e o titular, sempre que haja a invocação de segredo industrial ou comercial (LGPD, art. 20, §3º).

Os agentes de tratamento também podem criar mecanismos de boas práticas em revisão de dados, tanto individual como coletivamente (LDPG, art. 50, *caput*)

Em adição, a ANPD pode, por meio de sua competência regulatória, exercer influência sobre os procedimentos de operacionalização do direito à revisão (LGPD, art. 55-J, XIII).

4.4 Direitos Instrumentais

4.4.1 Direito de petição (CF, art. 5º, XXXIV, "a"; LGPD, art. 18,§1º c/c art. 55-J, V)

O direito de petição tem por objetivo invocar a atenção dos poderes públicos para uma questão[248]. Ele cabe a qualquer pessoa que se sinta prejudicada por ação ou omissão que possa ser corrigida por atuação de órgão público.

A sua base é constitucional e ele funciona como importante garantia de outros direitos fundamentais (CF, art. 5º, XXXIV, "a"). A Constituição assegura a sua amplitude, independentemente de qualquer condição ou encargo, e a sua gratuidade.

O direito de petição pode aparecer como uma queixa ou reclamação, ou como um pedido de informação, ou, finalmente, como a manifestação do direito de opinião[249].

No caso específico da proteção de dados, o direito de petição tem relação com a atividade reguladora e de fiscalização da ANPD. O art. 18, §1º da LGPD assegura ao titular dos dados pessoais o direito de peticionar em relação aos seus dados contra o controlador perante a autoridade nacional. Essa petição pode ter por objetivo o exercício de qualquer direito ou o esclarecimento de qualquer situação relacionada à proteção de dados.

Assim, o direito de petição pode visar ao acesso a dados pessoais, à confirmação de tratamento, à correção de dados incompletos, à anonimização de dados, ao bloqueio de dados,

---

[247] Ibidem, p. 9.
[248] SILVA, José Afonso da. **Curso de Direito Constitucional Positivo**. 35. ed. São Paulo: Malheiros Editora, 2012., p. 443.
[249] Ibidem.

enfim a concretizar qualquer direito do titular, por isso mesmo é considerado aqui um direito instrumental, isto é, uma garantia para outros direitos.

Em se tratando de petição contra ação ou omissão de controlador, por questão mesmo de demonstração de interesse, é indispensável que o titular demonstre à ANPD a existência de alguma resistência do agente de tratamento em atender à sua demanda independentemente de intervenção da autoridade pública (LGPD, art. 55-J, V). Nesse sentido, os controladores e operadores, no âmbito de suas competências, pelo tratamento de dados pessoais, individualmente ou por meio de associações, poderão formular regras de boas práticas e de governança que estabeleçam as condições de organização, o regime de funcionamento, os procedimentos, incluindo reclamações e petições de titulares (LGPD, art. 50, *caput*).

Ao direito de petição corresponde a obrigação da ANPD de apreciar a reclamação, queixa ou pedido de informação apresentado pelo titular. Além disso, é também obrigação da ANPD criar regulamentação própria para o exercício do direito de petição, inclusive prevendo prazo para resposta da demanda (LGPD, art. 55-J, V).

4.4.2 Direito ao devido processo legal (CF, art. 5º, LIV; LGPD, art. 4º, §1º)

O devido processo legal tem uma longa história no direito constitucional, especialmente no direito americano, e pode ser encarado em muitas dimensões.

Em um sentido muito amplo, o devido processo legal envolve tudo que diga respeito à vida, à liberdade ou à propriedade[250]. Em outra acepção, também muito ampla, o devido processo legal liga-se às ideias de razoabilidade e proporcionalidade. Esta última acepção tem sido empregada pelo Supremo Tribunal Federal em algumas ocasiões, inclusive para controlar a atividade legislativa, como, por exemplo, quando o STF considerou inconstitucional o uso de sanções indiretas para cobrar tributo[251], por serem irrazoáveis; ou quanto o STF considerou inconstitucional, também por irrazoabilidade, lei estadual que concedia gratificação de férias a servidor aposentado[252] ; ou, finalmente, quando o STF considerou inconstitucional Medida Provisória que mandava excluir, para fins de promoção por antiguidade, um ano do tempo de serviço de servidores do Poder Executivo federal[253].

---

[250] NERY JUNIOR, Nelson. **Princípios do processo na Constituição Federal**: processo civil, penal e administrativo. 11. ed. São Paulo: Editora Revista dos Tribunais, 2013, p. 92.

[251] BRASIL. Segunda Turma do Supremo Tribunal Federal. Recurso Extraordinário Com Agravo nº 915424. Relator: Ministro Celso de Melo. Brasília, DF, 20 de outubro de 2015. **Diário de Justiça Eletrônico**. Brasília, 30 nov. 2015.

[252] BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº 1158. Relator: Ministro Celso de Mello. Brasília, DF, 19 de dezembro de 1994. **Diário de Justiça Eletrônico**. Brasília, 26 maio 1995.

[253] BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº 1975. Brasília, DF, 25 de maio de 1999. **Diário de Justiça Eletrônico**. Brasília, 14 dez. 2001.

Em sentido mais restrito, o devido processo legal tem dimensão mais procedimental, como possibilidade efetiva de acesso à justiça[254], com possibilidade de acionar a proteção judicial, respeitando-se o direito de defesa e estabelecendo-se autêntico contraditório entre os litigantes.

Na LGPD (art. 4º, §1º), o devido processo legal é expressamente assegurado para o caso de tratamento de dados para fins de segurança pública, defesa nacional, segurança do Estado e atividades de investigação e repressão de infrações penais. Como essas situações estão declaradamente fora do escopo da LGPD, então o legislador se preocupou em atribuir aos titulares de dados pessoais, nessas situações, o direito ao devido processo legal e À observância dos princípios gerais da própria LGPD.

Na prática, considerando-se a amplitude que se pode dar ao devido processo legal e aos princípios da LGPD, as garantias oferecidas aos titulares de dados nas situações acima mencionadas não são menores do que aquelas outorgadas a eles em situações abrangidas pela LGPD.

Outra vertente doutrinária tem associado o devido processo legal, particularmente no âmbito do uso de modelos para a produção de decisões automatizadas, à ideia de um devido processo tecnológico. Levanta-se, em favor dessa tese, a questão segundo a qual as violações de direitos promovidas pelos modelos de decisão automatizada são tão complexas e desconhecidas ainda que se mostra necessário recorrer a uma garantia flexível[255] o suficiente para abrangê-las sem a necessidade de intervenção legislativa.

Essa é uma visão bastante promissora, já que oferece ao intérprete a abertura necessária para compreender e resolver graves problemas em relação aos modelos decisórios, sejam eles de *design*, de aplicações inapropriadas, ou usos sociais abusivos.

---

[254] Ibidem, p. 100.

[255] CRAWFORD, Kate; SCHULTZ, Jason. Big Data and Due Process: toward a framework to redress predictive privacy harms. **Boston College Law Review**, Boston, v. 55, n. 1, p. 93-128, 29 jan. 2014. Disponível em: https://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4. Acesso em: 17 jul. 2020.

CONCLUSÃO

A presente dissertação iniciou-se a partir do projeto de investigação em torno de um possível catálogo de direitos dos titulares de dados pessoais, em face de decisões automatizadas, à luz da Lei Geral de Proteção de Dados.

Os resultados encontrados, após pesquisa bibliográfica pertinente, são consistentes com a hipótese inicialmente levantada. O advento das tecnologias digitais reconfigurou o direito à privacidade, que já vinha sofrendo um processo de ressignificação desde o final do século XIX, de tal maneira que presentemente é mais adequado falar-se de um direito à proteção de dados.

O direito à proteção de dados, no ecossistema da internet e dos computadores digitais, apresenta várias dimensões, que ora o aproximam de um direito do consumidor (na relação do usuário com provedores de conteúdo), ora do direito à privacidade (em decorrência sobretudo da massiva coleta de dados pessoais nas rotinas econômicas da economia 4.0), ora do direito fundamental à própria imagem e à autodeterminação informativa.

Neste último sentido, o direito à proteção de dados encontra relevantes desafios em decorrência de usos, por agentes de tratamento, de modelos de decisão automatizada para a predição de características e comportamentos humanos com base em inferências estatísticas e probabilidades, extraídas de grandes bancos de dados. A pesquisa aponta que esse tipo de aplicação das novas tecnologias tem grande utilidade e produz comodidades de notável importância para as sociedades atuais, porém foram encontradas também situações, não necessariamente intencionais, em que os modelos preditivos produzem discriminação e iniquidade. Ademais, observou-se que essas situações, na maior parte das vezes, ocorrem sem que o titular dos dados tenha consciência da complexidade em que está envolvido, de modo que o seu consentimento não é, por si mesmo, um fundamento seguro para a construção de um sistema robusto de proteção de dados.

Verificou-se que existe a necessidade doutrinária de definir o que é e de como se forma uma decisão automatizada, à luz da LGPD. Intentando atingir esse objetivo, esboçou-se a seguinte definição:

> Decisão automatizada é todo julgamento feito exclusivamente por máquina, com base em predição decorrente de tratamento automatizado de dados pessoais de entrada, segundo um modelo ou algoritmo condicionado por dados de treinamento, que afete imediatamente interesse juridicamente tutelado de pessoa natural, excetuados aqueles que tenham fins particulares e não econômicos, jornalísticos ou científicos.

Em seguida, observou-se que os direitos dos titulares, em casos de decisão automatizada, podem ser divididos em três grandes grupos: a) direitos relativos à formação da decisão; b) direitos relativos aos resultados da decisão automatizada; e c) direitos instrumentais. Foram identificados, na LGPD e legislação correlata, os seguintes direitos relacionados à formação da decisão automatizada: 1) direito ao *design* adequado do sistema (art. 49, LGPD), inclusive com preferência para padrões técnicos que possibilitem ao próprio titular a fiscalização (art. 51, LGPD); 2) direito à confirmação da existência do tratamento (art. 18,I c/c art. 19, LGPD); 3) direito de opor-se ao tratamento (art. 18, VIII e §2º, LGPD) 4) direito de acesso aos dados pessoais e aos compartilhamentos (art. 9º c/c art. 18, II e VII, LGPD); 5) direito de correção (art. 18, III, LGPD); 6) Direito à anonimização (art. 18, IV, LGPD); 7) direito de consentir ou não no tratamento (arts. 7°, 11 e 14, LGPD); 8) direito de revogar o consentimento (art. 15, III c/c art. 18, XI, LGPD); 9) direito à eliminação dos dados (art. 18, VI, LGPD).

Quanto aos resultados da decisão automatizada, foram catalogados os seguintes direitos: a) direito de revisão das decisões automatizadas (art. 20, LGPD e art. 5º, VI da LCP); e b) direito à explicação (art. 20, §1º, LGPD).

Enfim, quanto aos direitos instrumentais, reconheceu-se a presença de dois: a) direito de petição (CF, art. 5º, XXXIV, "a"; LGPD, art. 18,§1º c/c art. 55-J, V); e b) direito ao devido processo legal (CF, art. 5º, LIV; LGPD, art. 4º, §1º).

Verificou-se também que o inter-relacionamento do direito à proteção de dados com o direito do consumidor e com outras leis setoriais pode ensejar o reconhecimento, em dadas circunstâncias, de outros direitos do titular de dados.

A pesquisa reconhece, por fim, que o catálogo de direitos acima não é exaustivo e, além disso, que há espaço para aprofundamentos, especialmente no que concerne ao devido processo legal como ferramenta multiuso flexível para abranger uma série de problemas imprevistos pelo legislador.

REFERÊNCIAS

ABREU, Jacqueline de Souza. Tratado de Proteção deTratamento de Dados Pessoais para Segurança Pública: contornos do regime jurídico pós-LGPD. In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle.

ACCIOLI, Wilson. **Instituições de Direito Constitucional**. 3. ed. Rio de Janeiro: Forense, 1984.

AGOSTINHO, Santo. **Confissões**. São Paulo: Companhia das Letras, 2017. Tradução de Lorenzo Mammì. Edição do Kindle

AGRAWAL, A; GANS, J.; GOLDFARB. **Máquinas preditivas***:* a simples economia da Inteligência Artificial. Rio de Janeiro: Alta Books, 2018. Tradução de Wendy Campos.

ALMADA, Marco. Revisão humana de decisões automatizadas. **Pósdebate 2019-USP**, São Paulo, p. 1-22, nov. 2019

ALBERS, Marion. A complexidade da proteção de dados. **Direitos Fundamentais & Justiça**, Belo Horizonte, a. 10, n. 35, jul./dez. 2016, p. 19-45, jul./dez.

AMBROSE, Meg Leta; AMBROSE, Ben M.. When robots lie a comparison of auto-defamation law. **2014 Ieee International Workshop On Advanced Robotics And Its Social Impacts**, [S.L.], p. 56-61, set. 2014. IEEE. http://dx.doi.org/10.1109/arso.2014.7020980.

ANDERSON, Chris. The Long Tail. *In*: CONDÉ NAST. **Wired**, 10 jan. 2004. Disponível em: https://www.wired.com/2004/10/tail/. Acesso em: 7 out. 2020.

ANDREW Ng: Artificial Intelligence is the New Electricity. Stanford: Stanford Graduate School of Business, 2 fev. 2017. 1 vídeo (1h 27 min). Publicado por Stanford Graduate School of Business. Disponível em: https://www.youtube.com/watch?v=21EiKfQYZXc. Acesso em 30 dez. 2020.

ANDRADE, Manuel A. Domingues. **Teoria geral da relação jurídica**. Coimbra: Almedina, 1992. v. 1.

ASSIS, J.M. Machado de. **O jornal e o livro**. São Paulo: Companhia das Letras, 2011.

BAMFORD, James. The Most Wanted Man in the World. *In*: CONDÉ NAST. **Wired**, 2014. Disponível em: https://www.wired.com/2014/08/edward-snowden/. Acesso em: 29 set. 2020.

BARRETO, Angela Maria. Informação e conhecimento na era digital. **Transinformação**, Campinas, v. 17, n. 2, p. 111-122, maio 2005. Disponível em: https://www.redalyc.org/pdf/3843/384334739002.pdf. Acesso em: 06 jun. 2020.

BBC NEWS (Brasil). **Huawei, Trump, Bolsonaro e China**: o que o brasil tem a ganhar e perder se ceder aos eua no 5g?. o que o Brasil tem a ganhar e perder se ceder aos EUA no 5G?.

2020. Disponível em: https://www.bbc.com/portuguese/brasil-54634201. Acesso em: 10 out. 2020.

BREEN, Jason. YouTube or YouLose? Can YouTube Survive a Copyright Infringement Lawsuit. **Bepress Legal Series.Working Paper 1950**, Los Angeles, p. 1-37, 18 jan. 2007. Disponível em: https://law.bepress.com/cgi/viewcontent.cgi?article=9209&context=expresso. Acesso em: 10 dez. 2020.

BINNS, Reuben; GALLO, Valeria. Automated Decision Making: the role of meaningful human reviews. *In*: ICO. **Information Commissioner's Office**. 12 abr. 2019. Disponível em: https://ico.org.uk/about-the-ico/news-and-events/ai-blog-automated-decision-making-the-role-of-meaningful-human-reviews/. Acesso em: 11 jan. 2021.

BIONI, Bruno R. **Proteção de dados pessoais**: a função e o limite do consentimento. 2.ed. Rio de Janeiro: Forense, 2020.

BITTAR, Carlos Alberto. **Os direitos da personalidade**. 8. ed. São Paulo: Saraiva, 2015. Ebook.

BOWLING, Michael; FÜRNKRANZ, Johannes; GRAEPEL, Thore; MUSICK, Ron. Machine learning and games. **Machine Learning**, [s.l.], v. 63, n. 3, p. 211-215, 10 maio 2006. Springer Science and Business Media LLC. http://dx.doi.org/10.1007/s10994-006-8919-x.

BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº 6.387. Relator: Ministra Rosa Weber. Brasília, DF, 24 de abril de 2020. **Diário de Justiça Eletrônico**. Brasília, 07 maio 2020.

_____. Presidência da República. Mensagem nº 288, de 8 de julho de 2019. **Diário Oficial da União,** Brasília, n. 130, p. 8, 9 jul. 2019. Disponível em: http://www.in.gov.br/en/web/dou/-/despachos-do-presidente-da-republica-190107781. Acesso em: 11 jul. 2020.

_____. Segunda Turma do Supremo Tribunal Federal. Recurso Extraordinário Com Agravo nº 915424. Relator: Ministro Celso de Melo. Brasília, DF, 20 de outubro de 2015. **Diário de Justiça Eletrônico**. Brasília, 30 nov. 2015.

_____. Terceira Turma do Superior Tribunal de Justiça. Recurso Especial nº 1193764 SP. Relator: Ministra Nancy Andrighi. Brasília, DF, 10 de dezembro de 2010. **Diário de Justiça Eletrônico**. Brasília, 8 ago. 2011.

_____. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº 1975. Brasília, DF, 25 de maio de 1999. **Diário de Justiça Eletrônico**. Brasília, 14 dez. 2001.

_____. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº 1158. Relator: Ministro Celso de Mello. Brasília, DF, 19 de dezembro de 1994. **Diário de Justiça Eletrônico**. Brasília, 26 maio 1995.

_____. Tribunal Pleno do Supremo Tribunal Federal. Inquérito nº 503-RJ. Relator: Ministro Sepúlveda Pertence. Brasília, DF, 24 de junho de 1992. **Diário de Justiça**. Brasília, 23 mar. 1993.

BRKAN, Maja; BONNET, Grégory. Legal and Technical Feasibility of the GDPR's Quest for Explanation of Algorithmic Decisions: of black boxes, white boxes and fata

morganas. **European Journal Of Risk Regulation**, [s.l.], v. 11, n. 1, p. 18-50, mar. 2020. Disponível em: https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/legal-and-technical-feasibility-of-the-gdprs-quest-for-explanation-of-algorithmic-decisions-of-black-boxes-white-boxes-and-fata-morganas/7324CDE80A300179C170C5BA8CA7E851. Acesso em: 17 jul. 2020.

BURROWS, Charles. The History of Radio Wave Propagation up to the End of World War I. **Proceedings Of The Ire**, [s.l.], v. 50, n. 5, p. 682-684, maio 1962. Institute of Electrical and Electronics Engineers (IEEE). http://dx.doi.org/10.1109/jrproc.1962.288097. Disponível em: https://ieeexplore.ieee.org/document/4066757. Acesso em: 31 maio 2019.

CALO, M. Ryan. Digital Market Manipulation. **George Washington Law Review**, Washington, D.C., v. 82, n. 4, p. 995-1051, ago. 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2309703. Acesso em: 17 jul. 2020.

CANCELIER, Mikhail Vieira de Lorenzi. O direito à privacidade hoje: perspectiva histórica e o cenário brasileiro. **Seqüência**: Estudos Jurídicos e Políticos, Florianópolis, v. 38, n. 76, p. 213-240, 20 set. 2017. Universidade Federal de Santa Catarina (UFSC). Disponível em: https://periodicos.ufsc.br/index.php/sequencia/article/view/2177-7055.2017v38n76p213/34870. Acesso em: 04 set. 2020.

CASTRO, Leandro Nunes de; FERRARI, Daniel Gomes. **Introdução à mineração de dados**: conceitos básicos, algoritmos e aplicações. São Paulo: Saraiva, 2016

CHACE, Calum. **Surviving AI**: the promise and peril of artificial intelligence. Oxford: Three Cs, 2015. Kindle Edition.

CHARTIER, Roger. As práticas da escrita. In: CHARTIER, Roger; ARIÈS, Philippe (org.). **Histórias da Vida Privada**: da renascença ao século das luzes. Da Renascença ao Século das Luzes. São Paulo: Companhia das Letras, 1991. 3 v. Tradução de Hidelgard Feist.

CHIAPPIN, J. R. N.; LEISTER, Carolina. O contratualismo como método: política, direito e neocontratualismo. **Revista de Sociologia e Política**, v. 18, n. 35, p. 09-26, fev. 2010. FapUNIFESP (SciELO). Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-44782010000100002&lng=pt&tlng=pt. Acesso em: 27 jan. 2021.

CLARKE, Roger. Profiling: a hidden challenge to the regulation of data surveillance. **Journal Of Law, Information And Science**, v. 2, n. 4, jan. 1993. Disponível em: http://www.austlii.edu.au/cgi-bin/sinodisp/au/journals/JlLawInfoSci/1993/26.html?query=. Acesso em: 10 dez. 2020.

CITRON, Danielle Keats. Technological Due Process. **Washington University Law Review**, Washington, D.c., v. 85, n. 6, p. 1249-1313, ago. 2008. Disponível em: https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview. Acesso em: 17 jul. 2020.

CRAWFORD, Kate; SCHULTZ, Jason. Big Data and Due Process: toward a framework to redress predictive privacy harms. **Boston College Law Review**, Boston, v. 55, n. 1, p. 93-128, 29 jan. 2014. Disponível em: https://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4. Acesso em: 17 jul. 2020.

CREVELD, Martin Van. **Ascensão e declínio do Estado**. São Paulo: Martins Fontes, 2004. Tradução de Jussara Simões.

CUEVA, Ricardo Villas Bôas. Proteção de Dados Pessoais e Direito ao Esquecimento. In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle.

DE LA TORRE , Lydia F. What is "data minimization" under EU Data Protection Law?. *In*: MEDIUM. **Medium**. [*S. l.*], 22 jan. 2019. Disponível em: https://medium.com/golden-data/what-is-data-minimization-under-eu-data-protection-law-b0e30fbb856e. Acesso em: 27 jan. 2021.

DONEDA, Danilo. **Da privacidade à proteção de dados pessoais**: elementos de formação da lei geral de proteção de dados. 2. ed. São Paulo: Thomson-reuters Brasil, 2019.

DOSHI-VELEZ, Finale *et al*. Accountability of AI Under the Law: the role of explanation. **Cornell University**, nov. 2017. Disponível em: https://arxiv.org/abs/1711.01134v1. Acesso em: 27 jan. 2021.

ELIAS, Norbert. **O processo civilizador**: formação do estado e civilização. Rio de Janeiro: Jorge Zahar Editor, 1993. 2 v. Tradução de Ruy Jungmann.

ERTEL, Wolfgang. Introduction. **Undergraduate Topics In Computer Science**, p. 1-21, 2017. Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-58487-4_1.

_____. **Introduction to Artificial Intelligence**. Cham (Switzerland): Springer, 2017. Tradução de Nathanael T. Black.

ESPANHA. AGENCIA ESPAÑOLA DE PROTECCIÓN DE DADOS. . **Introduction to 5G technologies and their risks in terms of privacy**. 2020. Disponível em: https://www.aepd.es/sites/default/files/2020-06/nota-tecnica-privacidad-5G-en.pdf. Acesso em: 05 out. 2020.

FOREIGN INTELLIGENCE SURVEILLANCE ACT. **Section 702 Overview**, 2008. Disponível em: https://www.dni.gov/files/icotr/Section702-Basics-Infographic.pdf. Acesso em: 29 set. 2020.

FRIEDMAN, Batya; HENDRY, David G. **Value Sensitive Design**: shaping technology with moral imagination. Cambridge (ma): The Mit Press, 2019.

GARVIE, Clare; FRANKLE, Jonathan. Facial-Recognition Software Might Have a Racial Bias Problem: Depending on how algorithms are trained, they could be significantly more accurate when identifying white faces than African American one. *In*: THE ATLANTIC. **The Atlantic**, 6 abr. 2016. Disponível em: https://www.theatlantic.com/technology/archive/2016/04/the-underlying-bias-of-facial-recognition-systems/476991/. Acesso em: 22 dez. 2020.

GIACAGLIA, Giuliano. Data is the New Oil. *In*: HACKER NOON. **Hacker Noon**. [*S. l.*], 9 fev. 2019. Disponível em: https://hackernoon.com/data-is-the-new-oil-1227197762b2. Acesso em: 17 jul. 2020.

GLEICK, James. **A informação**: uma história, uma enxurrada. São Paulo: Companhia das Letras, 2013. Tradução de Augusto Kalil.

GRAND CHAMBER. Judgment of the court in Case C-311/18. *In*: CURIA. **O TRIBUNAL DE JUSTIÇA DA UNIÃO EUROPEIA**. [*S. l.*], 16 jul. 2020. Disponível em: http://curia.europa.eu/juris/document/document.jsf?text=&docid=228677&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=4956180. Acesso em: 29 set. 2020.

HANNÁK, Anikó; WAGNER, Claudia; GARCIA, David; MISLOVE, Alan; STROHMAIER, Markus; WILSON , Christo. Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr. **In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing**, New York, p. 1914–1933, Fev. 2017.

HAUSSEN, Doria Fagundes. **Rádio e Política**: tempos de vargas e perón.. 1992. 324 f. Tese (Doutorado) - Curso de Doutorado em Ciências das Comunicações, Universidade de São Paulo, São Paulo, 1992.

HILDEBRANDT, Mireille. Privacy as Protection of the Incomputable Self: from agnostic to agonistic machine learning. **Theoretical Inquiries In Law**, Tel Aviv, v. 20, n. 1, p. 83-121, jan. 2019. Disponível em: https://www7.tau.ac.il/ojs/index.php/til/article/view/1622/1723. Acesso em: 17 jul. 2020.

HISTORY.COM EDITORS. Morse Code & the Telegraph. *In*: A&E TELEVISION NETWORKS. **HISTORY**. [*S. l.*], 9 nov. 2009. Disponível em: https://www.history.com/topics/inventions/telegraph. Acesso em: 30 mai. 2019.

HITACHI-UTOKYO LABORATORY (H-UTOKYO LAB). **Society 5.0**: a people-centric super-smart society. Tokyo: Springer, 2018. Edição do Kindle

HOROWITZ, Michael C.. Artificial Intelligence, International Competition, and the Balance of Power. **Texas National Security Review**, Austin, v. 1, n. 3, p. 37-57, maio 2018.

ICO. What does the UK GDPR say about automated decision-making and profiling?. *In*: ICO. **Information Commissioner's Office**. Disponível em: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-does-the-uk-gdpr-say-about-automated-decision-making-and-profiling/. Acesso em: 11 jan. 2021.

____. What is automated individual decision-making and profiling?. *In*: ICO. **Information Commissioner's Office**. Disponível em: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-is-automated-individual-decision-making-and-profiling/#:~:text=Automated%20decision%2Dmaking%20is%20the,to%20award%20a%20loan%3B%20and. Acesso em: 27 jan. 2021.

ISAACSON, Walter. **Os inovadores**: uma biografia da revolução digital. São Paulo, Companhia das Letras, 2014. Tradução de Berilo Vargas, Luciano Vieira Machado e Pedro Maria Soares.

KAHNEMAN, Daniel. **Rápido e devagar**: duas formas de pensar. Rio de Janeiro: Objetiva, 2012. Tradução de Cássio Arantes de Leite.

KEARNS, Michael; ROTH, Aaron. **The Ethical Algorithm**: the science of socially aware algorithm design. New York: Oxford University Press, 2019. Edição Kindle.

KELLY, Kevin. **Inevitável**: as 12 forças tecnológicas que mudarão nosso mundo. Rio de Janeio:Alta Books, 2019, Tradução de Cristina Yamagami

KUBAT, Miroslav. **An Introduction to Machine Learning**. 2. ed. Coral Gables: Springer, 2017.

LEE, Tian-Shyug; CHEN, I-Fei. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. **Expert Systems with Applications**, [*s. l.*], v. 28, n. 4, p. 743-752, mai. 2005.

LEE, Kai-Fu. **Inteligência artificial**: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos. Rio de Janeiro: Globo Livros, 2019. Tradução de Marcelo Barbão

LEHR, David; OHM, Paul. Playing with the Data: what legal scholars should learn about machine learning. **Uc Davis Law Review**, Davis, v. 51, n. 2, p. 653-717, dez. 2017. Disponível em: https://lawreview.law.ucdavis.edu/issues/51/2/Symposium/51-2_Lehr_Ohm.pdf. Acesso em: 17 jul. 2020.

LÉVY, Pierre. **As tecnologias da inteligência**: o futuro do pensamento na era da informática. São Paulo: Editora 34, 1993. Tradução de Carlos Irineu da Costa.

LUCAS. *In:* A BÍBLIA. Versão Almeida Revista e Corrigida. Barueri: Sociedade Biblica do Brasil, 2009.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY (ed.). **MIT Technology Review**. Self-driving cars. Topics. Disponível em: https://www.technologyreview.com/topic/smart-cities/self-driving-cars/. Acesso em: 02 jun. 2020.

MENDES, Laura Schertel. **Privacidade, proteção de dados e defesa do consumidor**: linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014.

MENDES, Laura Schertel; FONSECA, Gabriel Campos Soares da. STF reconhece direito fundamental à proteção de dados: comentários sobre o referendo da Medida Cautelar nas ADIs 6387, 6388, 6389, 6390 e 6393. **Revista de Direito do Consumidor**, v. 130/2020, p. 471-478. Jul./Ago, 2020.

METCALFE'S Law. In: JONATHAN LAW. **A Dictionary of Business and Management**. 5. ed. Milford: Oxford University Press, 2009. Disponível em: https://www.oxfordreference.com/view/10.1093/oi/authority.20110810105406240. Acesso em: 10 out. 2020.

MIRANDA, Jorge. **Manual de Direito Constitucional**. 4. ed. Coimbra: Coimbra Editora Ltda, 1990.

MIT TECHNOLOGY REVIEW INSIGHTS. How AI is humanizing health care: Artificial intelligence is helping health-care professionals do their jobs better, giving them the tools to build a smarter, more efficient ecosystem. *In*: MASSACHUSETTS INSTITUTE OF TECHNOLOGY (ed.). **MIT Technology Review**. [*S. l.*], 22 jan. 2020. Disponível em:

https://www.technologyreview.com/2020/01/22/276128/how-ai-is-humanizing-health-care/. Acesso em: 2 jun. 2020.

MOREIRA, José Carlos Barbosa. **Comentários ao Código de Processo Civil**. 7. ed. Rio de Janeiro: Forense, 1998. Vol. V.

NASH, Roderick Frazier. **The Rights of Nature**: a history of environmental ethics. Madison: The University Of Wisconsin Press, 1989.

NERY JUNIOR, Nelson. **Princípios do processo na Constituição Federal**: processo civil, penal e administrativo. 11. ed. São Paulo: Editora Revista dos Tribunais, 2013.

NORVIG, Peter Peter; NORVIG, Peter. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Elsevier, 2013. Tradução de Regina Célia Simille.

ODLYZKO, Andrew; TILLY, Benjamin. **A refutation of Metcalfe's Law and a better estimate for the value of networks and network interconnections**. 2005. Versão preliminar. Disponível em: https://www.researchgate.net/profile/Benjamin_Tilly/publication/228829389_A_refutation_of_Metcalfe's_Law_and_a_better_estimate_for_the_value_of_networks_and_network_interconnections/links/547f49960cf2ccc7f8b91b2b.pdf. Acesso em: 05 out. 2020.

OECD.ORG. OECD: better policies for better lives. Disponível em: https://www.oecd.org/going-digital/ai/principles/. Acesso em: 30 jun. 2020.

O'HARA, Kieron; HALL, Wendy. Four Internets: The Geopolitics of Digital Governance. **CIGI Papers**, n. 206, p. 128, dez, 2018. Disponível em: https://www.cigionline.org/sites/default/files/documents/Paper%20no.206web.pdf. Acesso em 07 out. 2020.

O'NEIL, Cathy. **Weapons of math destruction:**: how big data increases inequality and threatens democracy. New York:Crown Publishers, 2016. Ebook.

OI FUTURO. **Museu das telecomunicações**. Disponível em: http://museudastelecomunicacoes.org.br/historia-das-telecomunicacoes/. Acesso em: 31 maio 2019.

ONG, Walter. **Language as hermeneutic**: a primer on the word and digitization. Ithaca and London: Cornell University Press, 2017.

PERRY, Walter L.; MCINNIS, Brian; PRICE, Carter C.; SMITH, Susan C.; HOLLYWOOD, John S. **Predictive Policing**: The Role of Crime Forecasting in Law Enforcement Operations. [*S. l.*]: RAND Corporation, 2013. E-book.

PRIVACY BY DESIGN (Canadá). **The 7 Foundational Principles**. 2011. Disponível em: https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf. Acesso em: 23 out. 2020.

RAHWAN, Iyad *et al*. Machine behaviour. **Nature**, [*s. l.*], v. 568, p. 477–486, 24 abr. 2019. Disponível em: https://www.nature.com/articles/s41586-019-1138-y. Acesso em: 2 jun. 2020.

RODOTÁ, Stefano. **A vida na sociedade da vigilância**: a privacidade hoje. Rio de Janeiro: Renovar, 2008. Tradução Danilo Doneda e Luciana Cabral Doneda.

ROMANOSKY, Sacha; ACQUISTI, Alessandro. Privacy costs and persona data protection: economic and legal perspectives. **Berkeley Technology Law Journal**, Sacramento, v. 24, n. 3, p. 1063-1102, 2009. Disponível em: https://www.heinz.cmu.edu/~acquisti/papers/RomanoskyAcquisti-INFORMS-2009.pdf. Acesso em: 04 set. 2020.

SARPESHKAR, Rahul. Analog Versus Digital: extrapolating from electronics to neurobiology. **Neural Computation**, [s.l.], v. 10, n. 7, p. 1601-1638, out. 1998. MIT Press - Journals. Disponível em: https://ieeexplore.ieee.org/abstract/document/6790538. Acesso em: 10 jun. 2020.

SATARIANO, Adam. Facebook May Be Ordered to Change Data Practices in Europe. *In*: NYTCO. **The New York Times.** 9 set. 2020. Disponível em: https://www.nytimes.com/2020/09/09/technology/facebook-european-union-data-privacy.html. Acesso em: 24 set. 2020.

SAVIGNY, F. Carl Von. **Sistema del Derecho Romano Actua**. 2. ed. Madrid: Centro Editorial de Góngora, 2004. Tradução de: Espanhola de Jacinto Mesía e Manuel Poley.

SCHILLER, Arnold; WEISKOPF, Tobias. Automated Censorship in the Digital Space. *In*: YOUNG EUROPEAN FEDERALISTS (Europe). **The New Federalist**, 1 maio 2019. Tradução de Nora Teuma. Disponível em: https://www.thenewfederalist.eu/automated-censorship-in-the-digital-space?lang=fr. Acesso em: 9 dez. 2020.

SCHREIBER, Anderson. **Direitos da personalidade**. 2. ed. São Paulo: Editora Atlas, 2013.,

SCHWAB, Klaus. **A quarta revolução industrial**. São Paulo: Edipro, 2016, Tradução de Daniel Moreira Miranda.

SCIENCE MUSEUM. Goodbye To The Hello Girls: Automating The Telephone Exchang. *In*: SCIENCE MUSEUM GROUP. **Science Museum**. [*S. l.*], 22 out. 2018. Disponível em: https://www.sciencemuseum.org.uk/objects-and-stories/goodbye-hello-girls-automating-telephone-exchange. Acesso em: 31 maio 2019.

SEJNOWSKI, Terrence J. **A revolução do aprendizado profundo**. Rio de Janeiro: Alta Books, 2019. Traduzido por Carolina Gaio

SELBST, Andrew D; POWLES, Julia. Meaningful information and the right to explanation. **International Data Privacy Law**, [s.l.], v. 7, n. 4, p. 233-242, 1 nov. 2017. Oxford University Press (OUP). http://dx.doi.org/10.1093/idpl/ipx022. Disponível em: https://academic.oup.com/idpl/article-abstract/7/4/233/4762325. Acesso em: 13 jul. 2020.

SENADO DO CHILE. **Boletín 13828-19**: sobre protección de los neuroderechos y la integridad mental, y el desarrollo de la investigación y las neurotecnologías.2020. Disponível em: https://www.senado.cl/appsenado/templates/tramitacion/index.php?boletin_ini=13828-19. Acesso em: 02 out. 2020.

SEVCENKO, Nicolau (org.). **História da vida privada no Brasil**. São Paulo: Companhia das Letras, 1998. 3 v.

SHANNON, C. E.. A Mathematical Theory of Communication. **Mobile Computing And Communications Review**, S.l, v. 5, n. 1, p. 3-55, jan. 2001. Disponível em: https://culturemath.ens.fr/sites/default/files/p3-shannon.pdf. Acesso em: 21 maio 2019.

SILVA, José Afonso da. **Curso de Direito Constitucional Positivo**. 35. ed. São Paulo: Malheiros Editora, 2012.

SINGH, Manish. WhatsApp's new limit cuts virality of 'highly forwarded' messages by 70%. *In*: TECHCRUNCH. **TECHCRUNCH** [*S. l.*], 27 abr. 2020. Disponível em: https://techcrunch.com/2020/04/27/whatsapps-new-limit-cuts-virality-of-highly-forwarded-messages-by-70/. Acesso em: 21 jan. 2021.

SKANSI, Sandro. Introduction to Deep Learning: from logical calculus to artificial intelligence. **Undergraduate Topics In Computer Science**, [s.l.], v. 1, n. 1, p. 1-196, jan. 2018. Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-73004-2.

SOLOVE, Daniel. Autogestión de la privacidad y el dilema del consentimiento. **Revista Chilena de Derecho y Tecnología**, [S.L.], n. 3, p. 11-47, 23 jan. 2014. Universidad de Chile. http://dx.doi.org/10.5354/0719-2584.2013.30308. Disponível em: https://rchdt.uchile.cl/index.php/RCHDT/article/view/30308/32095. Acesso em: 01 out. 2020.

SOUZA, Carlos Afonso; PERRONE, Christian; MAGRANI, Eduardo. O direito à explicação entre a experiência europeia e sua positivação na LGPD. *In:* MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle.

SOUSA, Jorge Pedro. **Uma história breve do jornalismo no Ocidente**. Porto: Edições Universidade Fernando Pessoa, 2008.

TABACH, Danielle; LINHARES, Ludmila Anaquim. Transferência Internacional de dados. In: FEIGELSON, Bruno; SIQUEIRA, Antonio Henrique Albani. **Comentários À Lei Geral De Proteção De Dados**. São Paulo: Thomson Reuters Brasil Revista dos Tribunais, 2019. Cap. 5. p. 147-158.

TÁCIO, Caio. **Temas de Direito Público**: estudos e pareceres. Rio de Janeiro: Renovar, 1997.

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION. **Regulation nº 2016/679**. On the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). [*S. l.*], 27 abr. 2016. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=PT. Acesso em: 11 jul. 2020.

THE U.S. NATIONAL ARCHIVES AND RECORDS ADMINISTRATION. **Executive Order 12333:** United States intelligence activities, 1981. Disponível em: https://www.archives.gov/federal-register/codification/executive-order/12333.html. Acesso em: 29 set. 2020.

THOMSON, Amy; BODONI, Stephanie. Google CEO Thinks AI Will Be More Profound Change Than Fire. *In*: **Bloomberg**. [*S. l.*], 22 jan. 2020. Disponível em:

https://www.bloomberg.com/news/articles/2020-01-22/google-ceo-thinks-ai-is-more-profound-than-fire. Acesso em: 2 jun. 2020.

TOGOH, Isabel. WhatsApp Viral Message Forwarding Drops 70% After New Limits To Stop Coronavirus Misinformation. *In*: FORBES. **Forbes**. [*S. l.*], 27 abr. 2020. Disponível em: https://www.forbes.com/sites/isabeltogoh/2020/04/27/whatsapp-viral-message-forwarding-drops-70-after-new-limits-to-stop-coronavirus-misinformation/?sh=26e60cbc490d. Acesso em: 20 jan. 2021.

TURING, A. M.. On Computable Numbers, with an Application to the Entscheidungsproblem. **Proceedings Of The London Mathematical Society**, v. 2-42, n. 1, p. 230-265, jan. 1937. Disponível em: https://www.cs.virginia.edu/~robins/Turing_Paper_1936.pdf. Acesso em: 30 maio 2019.

VARGAS, Milton. **Para uma Filosofia da Tecnologia**. São Paulo: Editora Alfa Omega, 1994.

VINCENT, James. Putin says the nation that leads in AI 'will be the ruler of the world'. *In*: VOXMEDIA. **The Verge**. 4 set. 2020. Disponível em: https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world. Acesso em: 23 set. 2020.

WARREN, Samuel D.; BRANDEIS, Louis D.. The Right to Privacy. **Harvard Law Review**, Cambridge, v. 4, n. 5, p. 193-220, 15 dez. 1890. Disponível em: https://www.cs.cornell.edu/~shmat/courses/cs5436/warren-brandeis.pd. Acesso em: 03 set. 2020.

WEBER, Max. A objetividade do conhecimento nas ciências sociais. In: FERNANDES, Florestan (org.). **Weber**: sociologia. São Paulo: Ática, 1999. Coleção Grandes Cientistas Sociais.

ZARSKY, Tal. The Trouble with Algorithmic Decisions. **Science, Technology, & Human Values**, [s.l.], v. 41, n. 1, p. 118-132, 14 out. 2015. SAGE Publications. http://dx.doi.org/10.1177/0162243915605575. Disponível em: https://journals.sagepub.com/doi/abs/10.1177/0162243915605575. Acesso em: 17 jul. 2020.

DECISÕES AUTOMATIZADAS: DEFINIÇÃO, BENEFÍCIOS E RISCOS

AUTOMATED DECISIONS: DEFINITION, BENEFITS AND RISKS

Nazareno César Moreira Reis[1]

Sumário: 1 Introdução; 2 Conceito de decisões automatizadas; 2.1 Uso de dados pessoais; 2.2 Tratamento automatizado; 2.2.1 Tratamentos automatizados excluídos do alcance da LGPD (tratamentos domésticos, jornalísticos, artísticos e acadêmicos); 2.2.2 Tratamento automatizado regulado subsidiariamente pela LGPD (segurança pública, defesa nacional, segurança do Estado e atividades de investigação e repressão de infrações penais); 2.3 Ameaça ou lesão a interesse juridicamente tutelado; 2.4 Definição; 3 Decisões automatizadas e perfilização; 4 Os benefícios das decisões automatizadas; 4.1 Ciclo Virtuoso da Inteligência Artificial; 5 Os riscos das decisões automatizadas; 6 Conclusão

Resumo: O presente artigo tratou sobre a temática das decisões automatizadas produzidas por máquinas eletrônicas à luz do marco legal sobre o tema no Brasil: A Lei Geral de Proteção de Dados. A pesquisa buscou investigar o que são das decisões automatizadas, quais seus benefícios e riscos à luz das disposições da LGPD. A metodologia usada foi a pesquisa bibliográfica de doutrinas jurídicas e a pesquisa documental de legislação. Diante da complexidade e interdisciplinaridade do tema com outras áreas, mostra-se necessária a análise de textos fora do Direito, tais como de Ciência da Computação e de Filosofia da Tecnologia, em determinados momentos da pesquisa. Para concluir o trabalho, foi traçada uma definição para as decisões automatizada, com base na pesquisa feita, bem como foram pontuados os achados acerca de riscos e benefícios dessa forma decisória.
Palavras-chave: Decisões automatizadas, inteligência artificial, tratamento de dados, dados pessoais e Lei Geral de Proteção de Dados.

Abstract: This article studied the theme of automated decisions produced by electronic machines in the light of the legal framework on the subject in Brazil: The General Data Protection Law. The research sought to investigate what are the automated decisions, what are their benefits and risks in light of the provisions of the LGPD. The methodology used was the bibliographic research of legal doctrines and the documentary research of legislation. In view of the complexity and interdisciplinarity of the theme with other areas, it is necessary to analyze texts outside the law, such as Computer Science and Philosophy of Technology, in certain moments of the research. To conclude the work, an automated decision definition was drawn up, based on the research carried out, as well as the findings about the risks and benefits of this decision-making form were scored.
Keywords: Automated decisions, artificial intelligence, data processing, personal data and General Data Protection Law.

1 INTRODUÇÃO

O tema desta pesquisa diz respeito à articulação entre o direito e as decisões automatizadas produzidas por máquinas eletrônicas. Havendo já um marco legal no país sobre o assunto (Lei 13.709, de 14 de agosto de 2018, conhecida como Lei Geral de Proteção de

[1] Juiz federal e mestre em Direito Constitucional pelo IDP/iCEV.

Dados - LGPD), a investigação volta-se para tentar esclarecer em que consiste as decisões automatizadas, quais os seus benefícios e riscos.

A relevância do tema é grande. O modo de vida atual torna-se cada vez mais dependente das tecnologias da comunicação, em especial dos dispositivos computacionais (fixos ou móveis) e da internet. O trabalho, as compras, o lazer e, enfim, as relações humanas em geral passam por um processo de incorporação à atmosfera digital.

A representação de crescente coleção de fatos da vida em termos de códigos e objetos digitais, por seu turno, tem reduzido a distância entre o plano *off-line* o plano *on-line*, gerando influências recíprocas entre eles e abalando modelos mentais outrora consolidados, em especial aqueles ligados à intimidade e à privacidade.

Como todas as operações com códigos digitais são potencialmente gravadas e tendem a convergir para a internet, no ambiente das redes não há nada informal, nada local e nada completamente oculto à esfera pública. Todas as categorias jurídicas que se utilizam, portanto, das ideias de informalidade, territorialidade ou de sigilo precisam ser reconsideradas à luz dessa nova realidade material e social.[2]

A massa sempre crescente de dados, produzidos simultaneamente em vários contextos da vida coletiva, desde o âmbito doméstico até à política e à cultura, passando pela agricultura, indústria, comércio, ensino, pesquisa, etc., não fica sob o controle absoluto de nenhum indivíduo ou governo. Esses dados aglutinam-se em grupos geralmente volumosos (os *big datas*), fundem-se, refundem-se, apartam-se e se propagam inexoravelmente nos meios digitais; e, mesmo quando protegidos por mecanismos de criptografia que imitam no mundo *on-line* os muros, as paredes e os cofres do mundo *off-line*, apresentam suscetibilidades próprias de sua conformação, que precisam ser consideradas pelo direito.[3]

Embora os dados sejam gravados, no modo digital, sob a mesma lógica e segundo um padrão físico homogêneo (como um sinal eletromagnético), os fatos aos quais eles se referem concedem-lhes pesos jurídicos muito variados. Enquanto o meio físico os iguala, os valores humanos subjacentes os hierarquizam, donde surge uma tensão entre a técnica e a política, que acaba se expressando em termos jurídicos.

---

[2] LÉVY, Pierre. **As tecnologias da inteligência**: o futuro do pensamento na era da informática. São Paulo: Editora 34, 1993. Tradução de Carlos Irineu da Costa, p. 115-134.

[3] HILDEBRANDT, Mireille. Privacy as Protection of the Incomputable Self: from agnostic to agonistic machine learning. **Theoretical Inquiries In Law**, Tel Aviv, v. 20, n. 1, p. 83-121, jan. 2019. Disponível em: https://www7.tau.ac.il/ojs/index.php/til/article/view/1622/1723. Acesso em: 17 jul. 2020; CRAWFORD, Kate; SCHULTZ, Jason. Big Data and Due Process: toward a framework to redress predictive privacy harms. **Boston College Law Review**, Boston, v. 55, n. 1, p. 93-128, 29 jan. 2014. Disponível em: https://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4. Acesso em: 17 jul. 2020.

Assim é que, quando os dados estão ligados a uma pessoa natural identificada ou identificável, isto é, quando dizem respeito a fatos ou atos da vida de um ser humano, indicando aspectos específicos dos seus comportamentos, dos seus gostos, das suas preferências, eles são chamados de "dados pessoais" (LGPD, art. 5º, I), e têm uma proteção legal especial. Se, ademais, forem considerados especificamente os dados da pessoa natural que estão ligados à sua origem racial ou étnica, à sua convicção religiosa, à sua opinião política, à sua filiação a sindicato ou a organização de caráter religioso, filosófico ou político, bem como os referentes à sua saúde ou à sua vida sexual, dados genéticos ou biométricos — então se fala em "dados sensíveis" (LGPD, art. 5º, II), cuja proteção legal é ainda mais forte.

Esses dados (os pessoais e, mais ainda, os sensíveis) apresentam valor maior para o direito porque são expressões da personalidade humana, estando por isso no centro do ordenamento jurídico (CF, art. 1º, III). O esquema doutrinário tradicional, que explica a relação do homem com as coisas por meio dos direitos reais, em especial o direito de propriedade, não atende às necessidades ligadas à proteção dos dados pessoais. É que, por meio desses rastros digitais, com o devido "tratamento", pode-se reconstituir fatos, atos e até pensamentos relacionados a alguém, acossando o ser humano na intimidade de sua vida intelectual, afetiva, moral, política, econômica e social. Os dados pessoais não são, portanto, em relação à pessoa a quem se referem, coisas sobre as quais ela exerce algum direito real, mas sim emanações da sua personalidade, daí porque se fala de um "direito à proteção de dados", e não de um direito de propriedade sobre dados[4].

Com efeito, métodos sofisticados de tratamento de dados, chamados genericamente de Inteligência Artificial, permitem a recopilação de dados dispersos para reconstituir ações humanas e para analisar, prever ou mesmo induzir comportamentos futuros. Enfim, permitem formar uma imagem completa do indivíduo, a partir dos vestígios digitais dos seus movimentos nas redes, prognosticando as suas ações, características, interesses e até pensamentos.

Os usos desse poder novo e espantoso têm sido muito diversificados. As máquinas têm sido usadas para avaliar a capacidade de pagamento de pessoas que pedem empréstimos[5], para estimar preços de mercadorias conforme o consumidor que as queira comprar[6], para prever

---

[4] MENDES, Laura Schertel. **Privacidade, proteção de dados e defesa do consumidor**: linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014. p.120-124.

[5] LEE, Tian-Shyug; CHEN, I-Fei. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. **Expert Systems with Applications**, [*s. l.*], v. 28, n. 4, p. 743-752, mai. 2005.

[6] HANNÁK, Anikó et al. Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr. **Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing**, New York, p. 1914–1933, fev. 2017.

locais que devem receber maior atenção das rondas policiais[7], para dirigir carros autônomos[8], para fazer diagnósticos de doenças[9], em reconhecimento facial ou detecção de objetos por imagens para diversos fins, em *sites* de busca, veículos autônomos, classificação de crédito, publicidade comercial, seleção de pessoal para vagas de emprego, avaliação de produtos, etc.[10]

Na base de toda essa revolução está a Inteligência Artificial e os múltiplos usos que ela é capaz de fazer do grande volume de dados disponíveis na internet ou fora dela, sobretudo por meio do chamado Aprendizado de Máquina (*Machine Learning*). O conjunto dos efeitos sociais desses usos ainda é um território desconhecido, pleno de possibilidades, de esperanças e de muitos receios também.

Em semelhante contexto, o direito precisa se ocupar da disciplina dessas máquinas cognoscentes, visto que elas estão gerando fatos com importantes consequências jurídicas na vida das pessoas. Levanta-se, entre outras, uma questão que promete ocupar crescente atenção dos juristas em toda parte, mas nesta pesquisa com enfoque no Brasil: em que consistem as decisões automatizadas com base no tratamento de seus dados pessoais? Esse é o problema de pesquisa que será desenvolvido, especialmente, à luz da legislação interna, especialmente da Lei Geral de Proteção de Dados – LGPD (Lei 13.709, de 14 de agosto de 2018, com as alterações promovidas pela Lei n. 13.853, de 8 de julho de 2019).

O objetivo geral da pesquisa, desse modo, é avaliar o que são as decisões automatizadas, quais os riscos e benefícios que elas trazem, bem como quais as soluções jurídicas oferecidas pela LGPD para acomodar essa importante inovação tecnológica dentro do sistema jurídico.

De maneira específica, a pesquisa volta-se para a anatomia da decisão automatizada, intentando analisar os elementos fundamentais que a compõem, de modo a verificar por quais razões elas estão sendo tão utilizadas agota, quais suas vantagens, e quais os riscos inerentes ao iter decisório automático.

A metodologia consiste basicamente em pesquisa bibliográfica de doutrinas jurídicas e a pesquisa documental de legislação e jurisprudência. Como o tema apresenta aspecto

---

[7] PERRY, Walter L. *et al*. **Predictive Policing**: The Role of Crime Forecasting in Law Enforcement Operations. [*S. l.*]: RAND Corporation, 2013. E-book.

[8] MASSACHUSETTS INSTITUTE OF TECHNOLOGY (ed.). **MIT Technology Review**. Self-driving cars. Topics. Disponível em: https://www.technologyreview.com/topic/smart-cities/self-driving-cars/. Acesso em: 02 jun. 2020.

[9] MIT TECHNOLOGY REVIEW INSIGHTS. How AI is humanizing health care: Artificial intelligence is helping health-care professionals do their jobs better, giving them the tools to build a smarter, more efficient ecosystem. *In*: MASSACHUSETTS INSTITUTE OF TECHNOLOGY (ed.). **MIT Technology Review**. [*S. l.*], 22 jan. 2020. Disponível em: https://www.technologyreview.com/2020/01/22/276128/how-ai-is-humanizing-health-care/. Acesso em: 2 jun. 2020.

[10] RAHWAN, -Iyad et al. Machine behaviour. **Nature**, [s. l.], v. 568, p. 477–486, 24 abr. 2019. Disponível em: https://www.nature.com/articles/s41586-019-1138-y. Acesso em: 2 jun. 2020.

interdisciplinar, mostra-se necessária eventualmente também a análise de textos ligados a outras áreas, tais como de Ciência da Computação e de Filosofia da Tecnologia.

Em conclusão, a pesquisa responde o problema de pesquisa, que consiste justamente em explicar em que consistem as decisões automatizadas com base no tratamento de seus dados pessoais

2 CONCEITO DE DECISÕES AUTOMATIZADAS

A concepção de uma decisão automatizada envolve vários elementos, e somente se tornou possível com o uso de computadores eletrônicos. Na verdade, apenas em sentido metafórico se pode falar em "decisão" aqui, porque a máquina não age de modo consciente com algum propósito, mas apenas efetua cálculos aritméticos, segundo um programa (algoritmo) e conforme os dados que a alimentam. Logo, as máquinas apenas podem emular a parcela calculável da inteligência humana, não o livre-arbítrio, nem sentimentos, nem emoções.[11]

Como explicam Ajay Agrawal, Joshua Gans e Avi Goldfarb[12], quando a máquina toma uma decisão, ela usa dados de entrada (imagens, textos, sons, etc., que têm de ser reduzidos a um formato digital legível pela máquina), para fazer uma predição. A predição está baseada no "conhecimento" que o algoritmo ou modelo adquiriu na fase de treinamento, com os chamados dados de treinamento. Combinando a predição com o julgamento (escolha da solução, segundo o interesse do programador/desenvolvedor, expresso no algoritmo ou modelo), a máquina de decisão automática indica uma ação a ser efetivada (por humano ou outra máquina) e essa ação leva a um resultado (eventualmente com uma recompensa associada pelo programador). O resultado fornece ao modelo um *feedback* (positivo ou negativo), que assim realimenta todo o processo para decisões futuras.

A diferença entre algoritmo e modelo é fundamental para entender posteriores desdobramentos jurídicos relacionados às decisões automatizadas. Michael Kearns e Aaron Roth[13] explicam que a distinção entre algoritmo e modelo está em que o segundo é o resultado da aplicação do primeiro sobre uma vasta coleção de dados. Enquanto o algoritmo é o conjunto de regras que, aplicadas a um conjunto finito de dados, pode solucionar problemas semelhantes

---

[11] É certo que existem debates filosóficos, filmes e livros sobre a possível ascensão das máquinas inteligentes ao nível da autoconsciência, quando então ocorreria a singularidade, isto é, um crescimento teoricamente infinito da inteligência das máquinas sem a intervenção humana. Porém, tais discussões estão fora do propósito desta pesquisa. Para um bom panorama do tema, Cf.: CHACE, Calum. **Surviving AI**: the promise and peril of artificial intelligence. Oxford: Three Cs, 2015. Kindle Edition.

[12] AGRAWAL, Ajay; GANS, Joshua; GOLDFARB, Avi. **Máquinas Preditivas**: a simples economia da inteligência artificial. Rio de Janeiro: Editora Alta Books, 2018. Tradução de Wendy Campos, p. 74.

[13] KEARNS, Michael; ROTH, Aaron. **The Ethical Algorithm**: the science of socially aware algorithm design. New York: Oxford University Press, 2019. Edição Kindle, p.9.

em tempo finito, o modelo é, por assim dizer, um algoritmo com experiência prática anterior em avaliar dados. O modelo tem, portanto, um *background* que condiciona o seu modo de tratar dados novos, encaixando-os na sua "pré-compreensão". Dizem os referidos autores:

> As we've suggested, many of the algorithms we discuss in this book would more accurately be called models. These models, which make the actual decisions of interest, are the result of powerful machine learning (meta-) algorithms being applied to large, complex datasets. A crude but useful sketch of the pipeline is that the data is fed to an algorithm, which then searches a very large space of models for one that provides a good fit to the data. Think of being given a cloud of 100 points on a piece of paper, each labeled either "positive" or "negative," and being asked to draw a curve that does a good but perhaps imperfect job of separating positives from negatives. The positive and negative points are the data, and you are the algorithm—trying out different curves until you settle on what you think is the best separator. The curve you pick is the model, and it will be used to predict whether future points are positive or negative But now imagine that instead of 100 points, there are 10 million; and instead of the points being on a 2-dimensional sheet of paper, they lie in a 10,000-dimensional space.[14]

A "experiência" do algoritmo com os dados de treinamento é aprimorada por meta-algoritmos que otimizam o trabalho de construção do modelo, mediante a revisão sistemática dos dados de saída, segundo o resultado desejado pelo programador, para melhor agrupá-los e interrelacioná-los. O meta-algoritmo mais conhecido e usado é de *backpropagation*, que resumidamente pode ser descrito como um conjunto de instruções para reanalisar várias vezes os dados de saída e corrigir erros de avaliação porventura verificados, mediante um processamento inverso, melhorando o desempenho do modelo e reequilibrando os pesos dos fatores em jogo para a tomada de decisão[15].

Assim, por exemplo, uma máquina de reconhecimento facial para fins de localização de possíveis foragidos da justiça que estejam circulando em áreas públicas funciona da seguinte maneira: 1º) ela coleta os dados automaticamente (imagens), por meio de câmeras apontadas para os transeuntes em vias públicas (dados de entrada); 2º) o modelo utilizado para analisar

---

[14] Como sugerimos, muitos dos algoritmos que discutimos neste livro seriam chamados de modelos com mais precisão. Esses modelos, que tomam as decisões reais de interesse, são o resultado de poderosos algoritmos de aprendizado de máquina (meta-) aplicados a conjuntos de dados grandes e complexos. Um esboço rudimentar, mas útil, do pipeline é que os dados são alimentados para um algoritmo, que então procura um espaço muito grande de modelos por um que forneça um bom ajuste aos dados. Pense em receber uma nuvem de 100 pontos em um pedaço de papel, cada um rotulado como "positivo" ou "negativo", e ser solicitado a desenhar uma curva que faz um bom, mas talvez imperfeito trabalho de separar os positivos dos negativos. Os pontos positivos e negativos são os dados, e você é o algoritmo - experimentando curvas diferentes até chegar ao que você acha que é o melhor separador. A curva que você escolhe é o modelo e será usado para prever se os pontos futuros são positivos ou negativos. Mas agora imagine que em vez de 100 pontos, há 10 milhões; e em vez de os pontos estarem em uma folha de papel bidimensional, eles ficam em um espaço de 10.000 dimensões (tradução nossa)

[15] Para descrição dos aspectos matemáticos da questão, Cf AGGARWAL, Charu C.. Neural Networks and Deep Learning: a textbook. New York: Springe, 2018, p. 21 e ss. Esse tipo de meta-algoritmo é usado para otimizar mecanismos de *deep learning* que, como se explicará adiante, são os mais utilizados atualmente em aplicações práticas da chamada Inteligência Artificial.

esses dados, comparando-os com as imagens dos foragidos armazenadas em seus arquivos, foi previamente treinado com dados de muitos prisioneiros (dados de treinamento), de modo a fazer a associação tida como "correta" pelo programador; 3º) feito o cruzamento, se for encontrado uma correspondência (*match*), a máquina faz a *predição* de que ali está um foragido, com base no alto nível de probabilidade de a imagem coincidir com a do foragido X, por exemplo; 4º) em seguida, a máquina "julga" e aponta aquele suspeito para o operador; 5º) com base nesse julgamento, adota-se uma ação, que são os atos posteriores (que podem ser humanos ou automatizados também — no caso, a detenção do sujeito) que levarão ao resultado (no caso, prisão correta ou incorreta). Conforme esse resultado tenha sido correto ou incorreto, a depender de uma análise humana posterior, a máquina é informada, por *feedback*, para reforçar ou não aquele julgamento.

As regras de julgamento terão sido dadas pelo programador[16] com base em níveis estatísticos de confiabilidade em ambiente de incerteza, daí a semelhança desse processo automatizado com o funcionamento da mente humana. A grande capacidade de adaptação ao ambiente é o ponto forte da inteligência humana; o cérebro humano é capaz de reconhecer padrões, generalizá-los e de ajustar a decisão tendo em conta pequenas mudanças nesses padrões. Os modelos que trabalham com *machine learning*, em particular os de *deep learning*, buscam reproduzir artificialmente essa capacidade adaptativa do funcionamento orgânico do cérebro e, por isso, estão no centro das mais importantes e avançadas aplicações práticas do que se convencionou chamar da Inteligência Artificial[17].

Observa-se que a decisão automatizada, para além do algoritmo, é fortemente influenciada pelos dados, mais especificamente por três tipos de dados: a) os *dados de treinamento*; b) os *dados de entrada*; e c) os *dados de feedback*. Os dados de treinamento criam o *background* do modelo, numa fase anterior à colocação dele em funcionamento; os dados de entrada, já na fase de aplicação, sinalizam para o modelo o que está no ambiente externo, e os dados de saída são o resultado do processo decisório artificial. Os dados de saída poderão voltar à máquina, como *feedback* positivo ou negativo, para que ela possa se autoajustar ou ser ajustada pelo desenvolvedor.

Somente quando os dados de entrada são *dados pessoais* ou quando o julgamento diz respeito a alguma pessoa natural (caso em que os dados de saída são dados pessoais), é que se

---

[16] Como se verá adiante, existem métodos de aprendizado de máquina em que, embora as regras iniciais sejam dadas pelo programador, o modelo pode autonomamente ponderar os pesos dos dados, a partir de exemplos que lhe são apresentados, alterando a programação inicial.

[17] ERTEL, Wolfgang. **Introduction to Artificial Intelligence**. Cham (Switzerland): Springer, 2017. Tradução de Nathanael T. Black, p.3

pode falar em "decisão automatizada", no direito brasileiro, conforme se extrai do art. 20 da LGPD:

> Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade (...)

Ora, se em toda decisão automatizada o titular dos dados (de entrada ou de saída) tem direito de solicitar a revisão, então sempre haverá um titular em tais casos; logo, sempre estão em jogo dados pessoais nas decisões automatizadas, pois o titular é a "pessoa natural a quem se referem os dados pessoais que são objeto de tratamento" (LGPD, art. 5º, V).

De fato, há muitos processos automatizados na indústria ou na pesquisa científica que, no entanto, não produzem "decisões", no sentido empregado pela legislação brasileira. Em uma pesquisa científica sobre uma bactéria, por exemplo, pode-se usar processos automatizados para avaliar e prever aspectos ou comportamentos dessa forma de vida, sem que se possa falar, no entanto, em "decisão automatizada", na acepção jurídica da expressão. O mesmo pode ocorrer numa fábrica de parafusos que automatize os processos de avaliação da qualidade de seus produtos: isso não gera decisões automatizadas, no sentido empregado pela LGPD.

A LGPD não chega a definir o que seja decisão automatizada, mas a ela se refere para assegurar ao titular de dados pessoais o *direito à revisão* dessa decisão, bem como o *direito à explicação* sobre os processos e os dados utilizados na formulação da decisão, nos seguintes termos:

> Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade. (Redação dada pela Lei nº 13.853, de 2019)
> § 1º O controlador deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, observados os segredos comercial e industrial.
> § 2º Em caso de não oferecimento de informações de que trata o § 1º deste artigo baseado na observância de segredo comercial e industrial, a autoridade nacional poderá realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais.

Adiante analisa-se cada um dos elementos normativos utilizados para a composição de uma definição de decisão automatizada.

2.1 Uso de dados pessoais

O dispositivo legal referido estipula alguns elementos que permitem inferir o conceito de decisão automatizada, para os fins da LGPD. Em primeiro lugar, é preciso que a decisão tenha sido tomada mediante o uso de dados pessoais, visto que a lei fala de direitos do "titular" a respeito dessa decisão. E "titular" tem uma definição precisa na LGPD, a saber: é a "pessoa natural a quem se referem os dados pessoais que são objeto de tratamento" (LGPD, art. 5º, V). Dados pessoais, por sua vez, são aqueles que produzam informações relacionadas a pessoa natural identificada ou identificável (LGPD, art. 5º, I).

Logo, como referido acima, processos de automatização adotados em atividades que não envolvam dados pessoais, não estão abrangidos pela disciplina da LGPD. Um caso particularmente interessante é o da pessoa jurídica. Os dados relativos a pessoas jurídicas não são dados pessoais, de modo que o tratamento automatizado de dados relacionados às pessoas jurídicas não estão no raio de incidência da norma da LGPD que assegura os direitos de revisão e de explicação — embora não fique excluída a hipótese de se buscar tais direitos, sobretudo em casos de assimetria negocial, por aplicação analógica do Código de Defesa do Consumidor, ou de alguma normativa protetiva específica, ou até mesmo por aplicação direta da Constituição, com base na ideia mais geral de proteção de dados como direito fundamental extensível também às pessoas jurídicas.

Outra questão que pode ser levantada aqui é dos dados anonimizados. Eles não são considerados dados pessoais pela LGPD (art. 12), salvo quando o processo de anonimização ao qual foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser revertido. Todavia, a agregação de dados pessoais com posterior anonimização para a criação de modelos preditivos de comportamento humano individual parece estar dentro da disciplina das decisões automatizadas, sobretudo quando venham a afetar algum interesse individual ou coletivo, pois nesses casos os dados de saída serão pessoais.

Assim é que, por exemplo, o autopreenchimento dos *sites* de busca é modelado a partir de um grande número de pesquisas individuais. Mesmo que a anonimização dos dados que deram base para a formulação do modelo retire o caráter pessoal desses dados, é certo que a decisão automatizada de preenchimento pode vir a trazer danos individuais ou coletivos e, por isso, está sujeito à disciplina do art. 20 da LGPD. É o que ocorre, por exemplo, quando o autopreenchimento se refere a alguma pessoa natural específica. Neste caso, o nome da pessoa é um dado pessoal e a decisão automatizada, ao ligar esse nome a um fato, a uma característica,

a uma imagem, enfim a uma informação, produz conhecimento com dados pessoais do interessado.

Há inúmeros exemplos de precedentes, em vários países, sobre a questão do autopreenchimento pelos motores de busca na internet, notadamente o *Google*. Um tribunal em Milão obrigou o *Google* a rever o autopreenchimento de pesquisa que associava automaticamente o nome de uma pessoa, quando pesquisada, à palavra "vigarista"[18]. No Japão, a mesma empresa foi obrigada a excluir um autopreenchimento que associava o nome de um indivíduo a crimes cometidos por um homônimo[19]. Em 2013, na Alemanha, um tribunal federal foi mais longe e obrigou o *Google* a eliminar todos os autopreenchimentos difamatórios, quando provocado pelo respectivo interessado[20].

2.2 Tratamento automatizado

O tratamento de dados por mecanismos eletrônicos (digitais) está no cerne da concepção de decisões automatizadas. É por meio do Aprendizado de Máquina (*Machine Learning*), o tipo de programação mais usado em aplicações práticas, que dados pessoais podem ser transformados em informações e em conhecimento, por dispositivos que funcionam de forma autônoma, mediante associações, agregações e desagregações, arranjos e rearranjos de dados; análises de padrões em vastos conjuntos de dados; inferências estatísticas e estimativas probabilísticas — enfim, técnicas matemáticas convenientes para extrair conhecimentos de dados, mimetizando o funcionamento da inteligência humana, ou, pelo menos, a parte computável da inteligência humana.

Com efeito, a LGPD, para esboçar a ideia de decisão automatizada, estabelece que tal é aquela que tenha sido tomada "unicamente com base em tratamento automatizado" (art.20, LGPD). Assim, a lei parece buscar excluir de seu raio de eficácia tanto as decisões decorrentes diretamente da inteligência humana, como as decisões humanas assistidas por processos automatizados, que não devem ser consideradas decisões automatizadas, segundo a lei brasileira.

Nesta altura, vale lembrar a interessante discussão travada nos Estados Unidos caso Loomis *x* Winsconsin. Em fevereiro de 2013, Eric Loomis foi preso por dirigir um carro

---

[18] A íntegra de decisão pode ser lida em: MONTI, Andrea. Tribunale di Milano: Ord. 24 marzo 2011. *In*: MONTI, Andrea. **ICT LEX**: Diritto, politica, cultura della Rete. [S. l.], 24 mar. 2011. Disponível em: https://www.ictlex.net/?p=1285. Acesso em: 7 jan. 2021.

[19] Cf.: https://www.bbc.com/news/technology-17510651, Acesso em: 7 jan. 2021.

[20] AMBROSE, Meg Leta; AMBROSE, Ben M.. When robots lie a comparison of auto-defamation law. **2014 Ieee International Workshop On Advanced Robotics And Its Social Impacts**, [S.L.], p. 56-61, set. 2014. IEEE. http://dx.doi.org/10.1109/arso.2014.7020980.

roubado e por fugir de uma barreira policial em La Crosse (Wisconsin). Após o regular processamento da acusação, ele foi condenado pelo juiz local a uma pena de 6 anos de prisão. A sentença, ademais, negou a liberação condicional do condenado, sob o argumento, entre outras coisas, de que o COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), um modelo utilizado pelo Judiciário de Wisconsin para calcular o risco de reincidência dos acusados, apontava alto grau de periculosidade em Eric Loomis.

A defesa de Loomis apresentou recurso contra essa condenação, alegando que não se sabia exatamente de que maneira o COMPAS funcionava, e que os seus fabricantes naturalmente não iriam revelar, porque nesse sigilo residiria justamente o valor econômico do produto. Assim, o uso desse tipo de ferramenta, segundo a defesa, violaria o devido processo legal, especialmente o direito de ser sentenciado de forma fundamentada e sem o uso de fatores inverificáveis.

A Suprema Corte de Wisconsin rejeitou a apelação[21], sob o argumento de que o juiz não decidira unicamente com base no tratamento automatizado de dados, mas sim também com base em todo o contexto probatório. A Suprema Corte Americana, para a qual posteriormente foi dirigido um pedido de *writ of certiorari*, rejeitou o julgamento do mérito da questão[22].

Observa-se que a posição do Judiciário americano, nesse caso, tolerando o uso do tratamento automatizado de dados em um tema tão sensível como é a decisão sobre a liberdade de locomoção, apoiou-se no fato de que a deliberação, em última análise, não foi da máquina, mas sim do humano (o juiz) que apreciou o pedido de liberdade condicional, embora ele possa ter levado em conta a predição do modelo, que indicava alto risco de reincidência.

O problema do grau de contribuição humana para a decisão tende a ser geralmente o de mais difícil abordagem, quando se trata de delimitar o alcance da LGPD na questão das decisões automatizadas. O uso do advérbio "unicamente" parece sugerir que qualquer mínima intervenção humana no processo decisório descaracteriza a decisão como sendo automatizada. Isso porque se a decisão tem intervenção humana, qualquer que seja ela, não é possível calcular quanto dessa decisão decorreu de contribuição da máquina, de modo que, pela lei brasileira, tal decisão não é automatizada.

É certo que a decisão automatizada apenas se torna possível mediante ações humanas anteriores, de programadores, investidores, cientistas de dados, engenheiros, matemáticos, etc.. No entanto, chega um ponto em que o modelo pode funcionar autonomamente, produzindo

---

[21] Cf.: https://www.wicourts.gov/sc/opinion/DisplayDocument.pdf?content=pdf&seqNo=171690. Acesso em: 8 jan. 2021.

[22] Loomis v. Wisconsin, 881 N.W.2d 749 (Wis. 2016), cert. negado, 137 S.Ct. 2290 (2017).

deliberações de acordo com o seu modo de funcionamento ordinário, mediante a combinação de dados de entrada segundo um procedimento criado total ou parcialmente por programadores. É neste ponto que a intervenção humana pode descaracterizar a decisão como automatizada.

Se a máquina apenas assiste o humano, fornecendo-lhe elementos para avaliar as melhores alternativas, cabendo a escolha do resultado ao humano, isso não pode ser definido como decisão automatizada, segundo a LGPD. Por outro lado, se o humano apenas ratifica a decisão da máquina, sem possibilidade de criticá-la ou descartá-la, então a decisão é automatizada, apesar de eventualmente ser assinada por um ser humano. Neste último caso, ocorre aquilo que se se chama de *rubber-stamping*[23], ou seja, um mero carimbo do ser humano.

A Autoridade Independente de Dados do Reino Unido[24] e o Conselho Europeu de Proteção de Dados[25] publicaram algumas orientações elucidativas sobre a questão da intervenção humana como causa da descaracterização da decisão como automatizada. Tais orientações podem ser resumidas no seguinte: a) os revisores humanos devem estar envolvidos na verificação da recomendação do sistema e não devem apenas "rotineiramente" aplicar a decisão automatizada (o envolvimento dos revisores deve ser ativo e não apenas simbólico); b) os revisores humanos devem ter uma influência "significativa" (*meaningful*) na decisão automatizada, inclusive com autoridade e competência para ir contra ela; c) os revisores humanos devem "pesar" e "interpretar" a predição da máquina, considerando todos os dados de entrada disponíveis e outros fatores adicionais.

Dois exemplos, citados nas orientações da Autoridade de Proteção de Dados do Reino Unido[26], podem esclarecer a diferença entre decisão automatizada e decisão humana assistida por processos automatizados: 1º) Pense-se numa fábrica que calcula e paga o valor de uma gratificação dos empregados conforme a sua produtividade, apurada por mecanismos automatizados e sem qualquer intervenção humana ou com intervenção humana meramente homologatória — isso é uma decisão automatizada; 2º) agora pense-se numa fábrica que use

---

[23] BINNS, Reuben; GALLO, Valeria. **Automated Decision Making**: the role of meaningful human reviews. In: ICO. Information Commissioner's Office. 12 abr. 2019. Disponível em: https://ico.org.uk/about-the-ico/news-and-events/ai-blog-automated-decision-making-the-role-of-meaningful-human-reviews/. Acesso em: 11 jan. 2021.

[24] ICO. What does the UK GDPR say about automated decision-making and profiling?. *In*: ICO. **Information Commissioner's Office**. Disponível em: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-does-the-uk-gdpr-say-about-automated-decision-making-and-profiling/. Acesso em: 11 jan. 2021.

[25] JUSTICE AND CONSUMERS (Europea Union). Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. *In*: JUSTICE AND CONSUMERS (Europea Union). **European Commission**. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 11 jan. 2021

[26] ICO, op. cit.

mecanismos automatizados para avaliar a pontualidade dos empregados, disparando um aviso a um gerente de recursos humanos sempre que algum empregado, segundo apuração automatizada de dados, chega atrasado mais de tantas vezes — isso não é decisão automatizada, pois a máquina apenas prediz a situação (a falta de pontualidade) e comunica ao ser humano responsável, para que decida e adote a ação adequada.

2.2.1 Tratamentos automatizados excluídos do alcance da LGPD (tratamentos domésticos, jornalísticos, artísticos e acadêmicos)

Conforme o art. 4º, I, II e III da LGPD, para além dos casos de extraterritorialidade, não estão sob a proteção da lei especial brasileira os tratamentos de dados que sejam realizados: a) por pessoa natural para fins exclusivamente particulares e não econômicos; b) para fins exclusivamente jornalísticos ou artísticos; c) para fins acadêmicos, observado o disposto nos arts. 7º a 11 da LGPD.

Significa isso dizer que eventual decisão automatizada tomada com os objetivos acima expostos não está sujeita às restrições da LGPD, notadamente as previstas no art. 20. Tal conclusão decorre logicamente do fato de não ser o tratamento de dados, nesses casos, protegido pela LGPD. Está claro, todavia, que eventual violação a direito individual em tais circunstâncias, especialmente à privacidade ou à imagem do titular de dados pessoais, não deve ficar sem meios de reparação, podendo ser corrigida por instrumentos atípicos mediante a aplicação direta da Constituição, sobretudo por força da cláusula do devido processo legal (CF, art. 5º, LIV).

Afinal, o pressuposto da lei para excluir essas decisões da sua disciplina é de que elas são presumivelmente inofensivas a direitos de terceiros, ou estão albergadas pela liberdade de expressão, ou pela liberdade de investigação científica, de modo que, se for alegado e comprovado dano, ameaça de dano por abuso dessas liberdades, há de existir proteção legal contra a violação, ainda que apenas judicial (CF, art. 5º, XXXV).

2.2.2 Tratamento automatizado regulado subsidiariamente pela LGPD (segurança pública, defesa nacional, segurança do Estado e atividades de investigação e repressão de infrações penais)

Os tratamentos de dados para fins de exclusivos de segurança pública, defesa nacional, segurança do Estado e atividades de investigação e repressão de infrações penais (LGPD, art. 4º, III), embora sujeitos a futura legislação específica (LGPD, art. 4º, §1º), deverão até lá prever medidas proporcionais e estritamente necessárias ao atendimento do interesse público,

observados o devido processo legal (CF, art. 5º, LIV), os princípios gerais de proteção (LGPD, arts. 2º e 6º) e os direitos do titular previstos na própria LGPD (arts. 17 a 22).

Cabe à Agência Nacional de Proteção de Dados – ANPD um papel preponderante de regulamentação e fiscalização, na falta de lei específica, dos tratamentos de dados não inteiramente sujeitos à LGPD, como é a hipótese daqueles relacionados à segurança pública e à defesa nacional. Nesses casos, deverá a ANPD emitir opiniões técnicas ou recomendações, e deverá solicitar aos responsáveis relatórios de impacto à proteção de dados pessoais (LGPD, art. 4º, §3º).

Um ponto de grande interesse na disciplina do tratamento de dados para fins exclusivos de segurança pública, defesa nacional, segurança do Estado e atividades de investigação e repressão de infrações penais (LGPD, art. 4º, III), é que pessoas de direito privado não podem realizar esses tratamentos (LGPD, art. 4º, §2º) — exceto em procedimentos sob tutela de pessoa jurídica de direito público, que serão objeto de informe específico à autoridade nacional —, o que está em conformidade com o disposto nos arts. 142 e 144 da Constituição Federal, que atribuem com exclusividade às Forças Armadas e às Polícias Federal, Rodoviária Federal, Ferroviária Federal, Civis, Militares e Penais, a competência para as atividades de segurança externa e interna do país.

Enquanto não advém a legislação específica disciplinando o tratamento de dados para fins de segurança pública e atividades de investigação, o que se observa, pela remissão ampla feita pelo art. 4º, §1º da LGPD, é que eventuais decisões automatizadas tomadas nesse campo estarão sujeitas juridicamente ao disposto no art. 20 da LGPD, além de também deverem atender aos princípios gerais de proteção e ao devido processo legal.

A questão peculiar no tratamento de dados pessoais para fins de segurança e apuração criminal é que, como se sabe, há uma larga tradição jurídica, tanto legislativa quanto jurisprudencial, construída na era analógica, que abria exceções importantes à privacidade quando se cuidava de medidas investigativas requeridas judicialmente por autoridades policiais ou de segurança em geral, independentemente do consentimento do titular e, em alguns casos, até mesmo de sua ciência.

Assim, a proteção à privacidade se dava pela oposição de obstáculos ao acesso às informações íntimas do cidadão (por exemplo: sigilo bancário, sigilo fiscal, sigilo profissional, inviolabilidade de domicílio [27]); obstáculos esses que, excepcionalmente, poderiam ser

[27] Lei Complementar 105/2001; Lei Complementar 104/2001; Lei 5.172/1966 (Código Tributário Nacional); Decreto-lei 2.848/1940 (Código Penal); Constituição Federal, art. 5º, XI.

afastados, com certas reservas procedimentais, a pedido de autoridades policiais. Entretanto, na era digital, esse tipo de garantia torna-se em certos aspectos anacrônica, porque o indivíduo já não governa seus dados, que estão dispersos e profusos em muitos bancos de dados espalhados pela internet; e esses dados podem ser entrecruzados, por mecanismos de inferência apropriados, permitindo a prospecção indireta de informações sobre o indivíduo sem a necessidade de quebra de sigilos.

Nesse contexto, sem prejuízo dos sigilos tradicionais, é fundamental regulamentar a forma como a autoridade policial pode coletar dados pessoais ou reorientar dados já coletados para outros propósitos; como pode tratar esses dados; como pode correlacioná-los com outros, partindo já do pressuposto de que o acesso aos dados não sigilosos pode, indiretamente, levar ao conhecimento de informações sigilosas.

Jacqueline de Sousa Abreu, a esse propósito, faz as seguintes considerações:

> Se o direito à privacidade servia à proteção de escolhas e espaços individuais para realização de intimidade, o direito à proteção de dados pessoais emerge como uma ampla estrutura de proteção regulatória, em atenção a novas formas de danos e riscos a que cidadãos estão expostos. Está assentado na constatação de que a sociedade da informação expõe o indivíduo a diversos riscos de dano físico, material ou moral que comprometem o exercício de sua autonomia, a níveis individual e coletivo. Tais riscos são decorrentes de práticas e/ou estruturas institucionais que se desviam de noções básicas de justiça: ter uma expectativa legítima de respeito e consideração frustrada em suas relações sociais com empresas e com o Estado (pelo uso inesperado de suas informações, pela falta de segurança razoável dispensada a suas informações, pelo uso discriminatório, para dar alguns exemplos), e não possuir instrumentos de remediação, por exemplo."[28]

Constata-se aqui uma premissa que é constante na proteção de dados no ecossistema digital: as possibilidades de produção de informação a partir de dados são tantas e tão diversificadas que é mais conveniente regulamentar a forma como elas podem ser legitimamente implementadas do que tentar usar critérios materiais proibitivos, que sempre foi a técnica mais usada no mundo analógico.

Pequenos fragmentos de informação sobre o investigado, indícios quase desprezíveis, quando devidamente tratados e colocados em contato com grandes volumes de dados pessoais até mesmo de outras pessoas, colhidos muitas vezes para fins inocentes, podem ter um poder revelador insuspeitado. A importância desses fragmentos e indícios acaba se revelando *a posteriori*, em razão das ferramentas de mineração de dados (*data mining*), e não exatamente

---

[28] ABREU, Jacqueline de Souza. Tratado de Proteção deTratamento de Dados Pessoais para Segurança Pública: contornos do regime jurídico pós-LGPD. In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle, p. 592-593.

da matéria de que tratam. Por isso a técnica de isolar e tutelar mais fortemente certos tipos de dados pode não ser suficiente para a adequada proteção de dados pessoais no campo das investigações criminais e da segurança pública em geral.

2.3 Ameaça ou lesão a interesse juridicamente tutelado

Outro elemento integrante do conceito legal brasileiro de decisão automatizada, constante do art. 20 da LGPD, com forte inspiração na Regulamento Geral sobre a Proteção de Dados (GDPR) da União Europeia, está na necessidade de que a deliberação de máquina ameace ou atinja um interesse juridicamente protegido.

Assim, qualquer demanda, judicial ou extrajudicial, contra o controlador que produza decisões automatizadas está na dependência de que o titular dos dados pessoais alegue e prove a ameaça ou violação, pela decisão automatizada, de algum interesse próprio que tenha a tutela do direito. Depreende-se a contrario sensu que decisões automatizadas inofensivas a direitos individuais ou coletivos não estão sob a tutela da lei — como, de resto, ocorre em qualquer área do direito em relação a atos abnóxios, que recaem no campo da licitude.

Os interesses violáveis por decisões automatizadas são os mais diversos, tais como, por exemplo: liberdade de expressão, numa rede social que use algoritmos para moderar publicações ou filtros de *upload*[29]; imagem, num site de busca que associe automaticamente o nome de uma pessoa natural a uma notícia falsa; direitos autorais, numa rede que publique livremente os conteúdos carregados pelos usuários[30]; patrimônio e imagem, num site de compras que manipule automaticamente preços e ofertas, discriminando pessoas pela localização de sua residência, etc.

Todos esses interesses, quando violados dentro do contexto de processos automáticos de tratamento de dados, podem ser reconduzidos à esfera tutelada pelo direito à proteção de dados e suas manifestações especiais e instrumentais previstas na LGPD.

Embora já existam questões relativamente conhecidas nesse campo das decisões automatizadas, tais como aquelas associadas à discriminação algorítmica, não é possível antecipar todas as possíveis ofensas a interesses protegidos que são suscetíveis de ocorrer por força do tratamento automático de dados pessoais. A maior parte da responsabilidade nesse

---

[29] SCHILLER, Arnold; WEISKOPF, Tobias. Automated Censorship in the Digital Space. *In*: YOUNG EUROPEAN FEDERALISTS (Europe). **The New Federalist**, 1 maio 2019. Tradução de Nora Teuma. Disponível em: https://www.thenewfederalist.eu/automated-censorship-in-the-digital-space?lang=fr. Acesso em: 9 dez. 2020.

[30] BREEN, Jason. YouTube or YouLose? Can YouTube Survive a Copyright Infringement Lawsuit. **Bepress Legal Series.Working Paper 1950**, Los Angeles, p. 1-37, 18 jan. 2007. Disponível em: https://law.bepress.com/cgi/viewcontent.cgi?article=9209&context=expresso. Acesso em: 10 dez. 2020.

campo é atípica e centrada mais nos danos que nas condutas, como acentuado no capítulo anterior.

O art. 44, parágrafo único, da LGPD, bem enfatiza que a responsabilidade por tratamento irregular de dados nasce do dano, quando o agente de tratamento não observa as normas de segurança previstas no art. 46 da LGPD.

Sem dano ou ameaça de dano, não há responsabilidade, porquanto não há o que reparar ou assegurar. Assim, o critério inicial para avaliar a presença de uma situação em que a decisão automatizada pode ser questionada ou mesmo anulada, com base em algum direito do titular, é o dano ou o potencial de dano que ela pode causar a interesse juridicamente protegido. Este é um critério de ordem pragmática que está na essência da própria ideia de direito subjetivo. Como explica Manuel A. Domingues de Andrade,

> De toda maneira, onde há um direito subjectivo, ele foi concedido para que através dele fosse obtido o predomínio de certo interesse; tal como a correspondente obrigação ou sujeição foi imposta para que um outro interesse oposto resultasse subordinado àquele.
> Mas uma coisa é o direito subjectivo em si mesmo e outra coisa é a razão por que, ou o fim em vista do qual, a lei atribui esse direito, ou seja o interesse para cuja prevalência tal direito foi concedido.
> O interesse constitui o substrato do direito subjectivo. É-lhe subjacente; está antes dele. Ou então — se assim se prefere — está para além dele. Em todo caso, está fora dele. Não diz respeito à sua estrutura, mas só à sua função. Não tem que entrar, portanto, na definição do respectivo conceito.[31]

O interesse está, consequentemente, no cerne da função de proteção jurídica. É por meio do interesse que, antes de tudo, se pode avaliar a necessidade e a utilidade de mecanismos jurídicos de tutela contra as decisões automatizadas. Se algum interesse juridicamente protegido for violado ou ameaçado pela decisão automatizada, há, quando menos, o direito de questionar em juízo o ato da máquina, por força da garantia do direito de ação (CF, art. 5º, XXXV). Adicionalmente, pode-se invocar os direitos consagrados na LGPD, e, conforme o caso, no CDC, no CC, na Lei do Cadastro Positivo (Lei 12.414/2011), na Lei de Acesso à Informação (Lei 12.527/2011) e em qualquer outra legislação, inclusive tratados, que, mesmo pensados para relações do mundo analógico, possam ser aplicados por semelhança ao contexto digital, conforme determina o art. 64 da LGPD.

---

[31] ANDRADE, Manuel A. Domingues. **Teoria geral da relação jurídica**. Coimbra: Almedina, 1992. v. 1, p. 8.

2.4 Definição

Baseado nas premissas acima apresentadas, pode-se construir uma definição que expresse objetivamente em que consiste uma decisão automatizada no contexto da Lei Geral de Proteção de Dados – LGPD.

Decisão automatizada é todo julgamento feito exclusivamente por máquina, com base em predição decorrente de tratamento automatizado de dados pessoais de entrada, segundo um modelo ou algoritmo condicionado por dados de treinamento, que afete imediatamente interesse juridicamente tutelado de pessoa natural, excetuados aqueles que tenham fins particulares e não econômicos, jornalísticos ou científicos.

Em adição, há dois tipos de julgamento que podem ser classificados como decisões automatizadas por equiparação: 1º) aqueles que, satisfazendo as condições referidas acima, recebam intervenção humana meramente homologatória (*rubber-stamping*); e 2º) as perfilizações automáticas, conforme se verá adiante.

No núcleo do processo de decisão automatizada estão técnicas estatísticas que permitem a extrapolação de informações, a partir de amostras de dados de uma população. Evidentemente, essas técnicas estão sujeitas a erros e desvios típicos do campo estatístico, embora no geral sejam confiáveis como processo de inferência e predição[32]. Como dados pessoais são indispensáveis para qualquer tipo de decisão automatizada, dentro do contexto da legislação brasileira, a construção de perfis individuais aparece sempre associada a qualquer julgamento feito por máquina e foi equiparada, por lei, à decisão automatizada.

3 DECISÕES AUTOMATIZADAS E PERFILIZAÇÃO

A perfilização está tão intimamente ligada às decisões automatizadas que a LGPD (art. 20) a inclui no próprio conceito destas:

> Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, *incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade.*

Na verdade, porém, a perfilização está mais associada à predição e somente pode ser considerada a decisão automatizada se ela mesma for o objetivo do modelo ou algoritmo. Caso se queira, por exemplo, avaliar a capacidade de pagamento de alguém para efeito de concessão de um empréstimo, a perfilização será parte do tratamento de dados e da predição, mas a decisão

[32] KUBAT, Miroslav. **An Introduction to Machine Learning**. 2. ed. Coral Gables: Springer, 2017, p. 231.

não estará nisso, e sim na concessão ou não do empréstimo. A decisão é sempre uma tomada de posição diante dos dados, e não apenas uma inferência estatística. A predição, que decorre das inferências estatísticas, apontará o provável resultado da operação de empréstimo (digamos, há 80% de chance de o indivíduo pagar o empréstimo dentro do prazo); já a decisão estará em definir o titular dos dados como apto ou não para o empréstimo. Por exemplo, certa instituição financeira pode decidir pelo sim, com 80% de chance de pagamento, mas outra pode exigir um limiar de predição maior (digamos, 90%) para contratar o empréstimo. Portanto, a predição não é ainda a decisão; ela é o prenúncio do que provavelmente ocorrerá, caso a decisão seja tomada em um ou outro sentido, à luz dos dados tratados pelo modelo. A preferência por acolher essa probabilidade como um "sim" ou um "não" é que a decisão.

É difícil pensar a decisão automatizada sem algum grau de perfilização. Visto como os dados pessoais, por definição, sempre estão associados a alguma pessoa natural e devem fazer parte do processo de formação da decisão automatizada, como exposto acima; considerando também que o objetivo prático dessas decisões sempre está de algum modo associado à compreensão de características ou do comportamento pretérito de pessoas naturais, para avaliar as suas características ou seus comportamentos futuros, então algum grau de perfilização quase sempre está na base das decisões automatizadas.

Em certos casos, todavia, pode ocorrer decisão automatizada sem perfilização. A Autoridade Independente do Reino Unido menciona, a esse respeito, o caso de uma correção de prova automatizada[33]. Uma banca examinadora pode usar um sistema automatizado para marcar as folhas de respostas de um exame de múltipla escolha. O sistema é pré-programado com o número de respostas corretas necessárias para alcançar marcas de aprovação e distinção. As pontuações são automaticamente atribuídas aos candidatos com base no número de respostas corretas de cada um e os resultados estão disponíveis *online*. Trata-se de um processo automatizado de tomada de decisão que não envolve criação de perfil. Mas isso apenas ocorre em situações pontuais, que não busquem utilizar o modelo reiteradamente para o futuro, como essa cogitada, e não representa o coração das aplicações de processos automatizados nos processos produtivos.

---

[33] ICO. What is automated individual decision-making and profiling?. *In*: ICO. **Information Commissioner's Office**. Disponível em: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-is-automated-individual-decision-making-and-profiling/#:~:text=Automated%20decision%2Dmaking%20is%20the,to%20award%20a%20loan%3B%20and. Acesso em: 27 jan. 2021

O Regulamento Europeu para a Proteção de Dados (GDPR) define a perfilização (ou "definição de perfil", na tradução portuguesa) como algo diferente da decisão automatizada, embora não seja totalmente fiel a essa distinção em outros pontos. Com efeito, o art. 4º, n. 4 do GDPR associa a perfilização com a análise e a predição, que são anteriores à decisão, *verbis*:

> «Definição de perfis», qualquer forma de tratamento automatizado de dados pessoais que consista em utilizar esses dados pessoais para avaliar certos aspetos pessoais de uma pessoa singular, nomeadamente para analisar ou prever aspetos relacionados com o seu desempenho profissional, a sua situação econômica, saúde, preferências pessoais, interesses, fiabilidade, comportamento, localização ou deslocações.

Já em relação à decisão automatizada, o art. 22, n. 1 do GDPR (que, no ponto, foi praticamente copiado pela LGPD) estipula, *verbis*:

> O titular dos dados tem o direito de não ficar sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar.

A associação da perfilização com as decisões automatizadas decorre da circunstância de que, como visto, somente são consideradas automatizadas decisões que utilizem dados pessoais em seu processo de concepção. Como os dados pessoais, por definição, somente são aqueles referentes a uma pessoa natural, então o modelo capaz de produzir decisão automatizada sempre terá dados referentes a alguma pessoa natural como dados de entrada, daí porque a predição que ele fará resultará no prognóstico sobre alguma característica ou comportamento humano, baseado em características ou comportamentos anteriores da mesma ou de outras pessoas naturais que apresentem certo padrão reconhecido pela máquina. A decisão automatizada será baseada nessa predição, por isso ela de alguma maneira está conectada ao perfil decorrente dos dados de entrada.

Assim, um modelo que crie decisões automatizadas para admitir ou negar a entrada de pessoas numa universidade será previamente alimentado com um vasto conjunto de dados anteriores, sobre a admissão e a rejeição de candidatos (dados pessoais, portanto). Matematicamente, o modelo inferirá padrões desse conjunto de dados pessoais: tanto padrões para os que devem ser admitidos, como para os que devem ser rejeitados. Tão logo sejam inseridos os dados de interesse de um novo candidato (local de residência, notas, renda mensal, idade, enfim o conjunto de dados pessoais do postulante à vaga), o modelo predirá se o caso, à luz dos anteriores, é de admissão ou de rejeição; e a decisão de admitir ou rejeitar será tomada com base no grau da predição. É evidente que, em tal contexto, o novo candidato estará sendo perfilizado pelo modelo, embora não seja a perfilização propriamente o objetivo do tratamento

de dados; ela é, na verdade, uma etapa para a construção da decisão — seguramente uma etapa muito relevante.

O mesmo ocorrerá em um modelo de previsão de fraudes bancárias, ou de cotação de preços de mercadorias com base nos dados do pretenso comprador, ou em um mecanismo policial ou alfandegário que decida automaticamente quem deve ser fiscalizado preferencialmente. Sempre haverá a concepção de tipos genéricos que serão comparados aos dados pessoais dos sujeitos de interesse, para a solução do problema de negócio. Logo, a perfilização é um passo necessário para a tomada de decisões automatizadas, mas não é a própria decisão automatizada.

As decisões automatizadas podem ser realizadas com ou sem definição de perfis; a definição de perfis pode ocorrer sem que dela decorra uma decisão automatizada. Todavia, a definição de perfis e as decisões automatizadas não constituem necessariamente atividades separadas. Um procedimento iniciado como um processo de decisão automatizada pode tornar-se um procedimento de definição de perfis, dependendo da forma como os dados sejam utilizados.[34]

Em outras circunstâncias, a decisão automatizada pode ou não depender de perfilização, segundo o interesse do desenvolvedor na concepção do modelo. Assim, um sistema automatizado de imposição de multas de trânsito, a partir de imagens de câmeras de monitoramento espalhadas nas vias públicas, pode não levar em conta nenhum fator particular do infrator — nesse caso, portanto, desprezando a perfilização. Mas o mesmo modelo pode ser incrementado, para incluir características específicas do infrator (tempo de habilitação, multas anteriores, profissão, etc.), de modo a calibrar o valor da multa. Nesse caso, a perfilização estaria presente na composição da decisão automatizada.[35]

Pela redação da LGPD, no entanto, deve-se admitir que a perfilização, mesmo quando não seja seguida de uma decisão automatizada, mas sim de uma decisão humana assistida por máquina, deve ser considerada em si mesma uma decisão automatizada por equiparação, já que a lei afirma expressamente que estão incluídas entre as decisões automatizadas aquelas

---

[34] Cf.: JUSTICE AND CONSUMERS (Europea Union). Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. In: **JUSTICE AND CONSUMERS** (Europea Union). European Commission. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 11 jan. 2021

[35] O exemplo é dado, com pequenas alterações, na página 7 do Guia de Orientações já citado: Cf.: JUSTICE AND CONSUMERS (Europea Union). Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. In: **JUSTICE AND CONSUMERS** (Europea Union). European Commission. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 11 jan. 2021

"decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade" (artigo 20, LGPD).

Historicamente, a perfilização antecede o uso massivo de processos automatizados de coleta e tratamento de dados. Já nos anos 1980, falava-se do processo de crescente perfilização em várias áreas, especialmente no campo criminal e no âmbito do marketing direcionado[36]. O método de construção de perfis é notoriamente suscetível às técnicas que estão na base dos processos de aprendizado de máquina, daí porque a coleta massiva e o tratamento automatizado de dados pessoais naturalmente implicaram um processo exponencial de perfilização.

De fato, a perfilização, conforme Roger Clarke[37], é uma técnica por meio da qual um conjunto de características de um grupo particular de pessoas é inferido a partir de experiências passadas (das mesmas pessoas ou de pessoas com comportamento assemelhado), de modo tal a formar acervos que podem ser comparados com indivíduos no futuro, para avaliar o quanto estes se ajustam às características típicas do grupo. Bem analisada, a ideia de perfilização, em termos de método para conhecer objetivamente a mecânica dos comportamentos humanos, pode mesmo remontar ao conceito de "tipo ideal", de Max Weber[38], pois na sociologia o estudo de padrões de comportamentos sociais a partir da junção e organização de fragmentos esparsos de condutas individuais e de grupo é uma ferramenta há muito utilizada.

A metodologia matemática, na qual está a essência das técnicas de aprendizado de máquina, usa frequentemente o processo de reunião de objetos por características comuns (conjuntos), para inferir as relações de pertinências ou não de outros objetos. A perfilização por mecanismos automatizados é fundamentalmente um procedimento matemático de coleta, seleção, agrupamento e comparação de dados pessoais.

4 OS BENEFÍCIOS DAS DECISÕES AUTOMATIZADAS

O que leva as empresas e os governos a automatizarem os seus processos decisórios é, sem dúvida, o aumento da capacidade e da velocidade de resposta a demandas repetitivas e a redução de custos que isso proporciona. Por isso mesmo, decisões políticas ou que contenham elementos discricionários ou de estratégia negocial normalmente permanecem sob a

---

[36] CLARKE, Roger. Profiling: a hidden challenge to the regulation of data surveillance. **Journal Of Law, Information And Science**, v. 2, n. 4, jan. 1993. Disponível em: http://www.austlii.edu.au/cgi-bin/sinodisp/au/journals/JlLawInfoSci/1993/26.html?query=. Acesso em: 10 dez. 2020.

[37] Op.cit.

[38] WEBER, Max. A objetividade do conhecimento nas ciências sociais. In: FERNANDES, Florestan (org.). **Weber**: sociologia. São Paulo: Ática, 1999. Coleção Grandes Cientistas Sociais, p. 79-123.

governança estritamente humana, embora possam ser subsidiadas por tratamentos automatizados de dados.

Decisões tomadas em massa, com certo padrão, traduzíveis em termos matemáticos, tais como preços de mercadorias, contratos de empréstimos e análises de risco, são particularmente suscetíveis ao processo de automatização, desde que se tenha um conjunto relevante de dados que permita construir um modelo replicador das decisões anteriores.

O aprendizado de máquina busca imitar a racionalidade humana, a qual, por sua vez, está baseada na observação e organização intelectual do mundo, segundo padrões prévios, para predizer o futuro.

A automatização é um processo fundamentalmente estatístico-matemático: desvendam-se padrões nos dados e, a partir disso, a máquina "aprende" a reconhecê-los e a associá-los às "decisões corretas" respectivas. Cria-se, em suma, uma conexão lógica entre os dados de entrada e a decisão desejável para um futuro presumível. A máquina "aprende" a fazer essa imputação e, a partir de então, pode trabalhar de forma autônoma à vista da entrada de novos dados.

Durante o funcionamento do processo de decisão automatizada, as saídas podem ser otimizadas pelos programadores, mediante um processo ajuste fino do modelo por meio dos dados de *feedback*, ou mesmo por meta-algoritmos, como os de *backpropagation*; assim, o modelo pode criar decisões automatizadas ainda melhores, num processo teoricamente infinito de autoaprendizagem e autocorreção coadjuvado ou não por seres humanos, chamado de programação dinâmica (*dynamic programming*)[39].

No processo de autoaprendizagem são muito relevantes também as exceções, ou seja, aquelas situações que parecem se encaixar em certo padrão, mas na verdade são diferentes. É justamente nesse ponto que o modelo pode produzir decisões enviesadas ou iníquas, por generalizar demais ou de menos o padrão que lhe foi ensinado. Em tese, quanto mais dados são apresentados ao modelo, mais chance de ele encontrar exceções que precisam de um tratamento diferente. Em contraste, o modelo pode ser pobre em dados de treinamento, não atinando para padrões que seriam perceptíveis num conjunto maior de dados. Os dados de treinamento são determinantes para a acurácia de qualquer modelo de aprendizado de máquina atual.

Os modelos podem assumir grande número de processos decisórios em empresas, governos e organizações em geral, liberando recursos humanos e materiais para a assunção das exceções, normalmente ligadas a processos não quantificáveis. Assim, os benefícios da

[39] KUBAT, Miroslav, op.cit.,p. 338.

automatização para as empresas e organizações em geral são, antes de tudo, econômicos. No caso dos governos, a automatização pode trazer maior eficiência em serviços e políticas públicas, além de ser também fator de aperfeiçoamento econômico e administrativo.

Para os consumidores e usuários de serviços públicos, as vantagens dos processos automatizados residem na criação de comodidades cada vez mais personalizadas e, consequentemente, mais adequadas às necessidades específicas de cada indivíduo ou família. Desde a indicação de filmes, livros e produtos em geral, conforme os hábitos de consumo demonstrados em operações anteriores, até o relacionamento com o Fisco ou o deferimento de benefícios sociais ou outras prestações do Poder Público, conforme o perfil do contribuinte ou do grupo familiar, os mecanismos automatizados criam uma sinergia profunda que proporciona altos graus de eficiência em grande escala nos mais diferentes processos produtivos.

Numa visão mais radical e mais otimista, o processo de automatização levará a humanidade a uma Sociedade 5.0, de grande abundância e conforto proporcionado pelas máquinas, mediante a integração total de vários sistemas inteligentes, com a fusão quase completa do mundo *off-line* com o mundo *on-line*.

Embora alguns processos de automatização já tragam benefícios palpáveis para consumidores e usuários de serviços públicos, a ideia de uma sociedade 5.0 é muito mais ampla e profunda, porque imagina toda a vida social imersa no crisol da Inteligência Artificial, sem que haja a necessidade de "acessar" nada, uma vez que a realidade física estará ela mesma envolta e hibridizada com os mecanismos inteligentes, a tal ponto que não será possível perceber qualquer diferença entre estar *on-line* ou *off-line*. Nesse sentido, observou-se:

> In summary, Society 5.0 will feature an iterative cycle in which data are gathered, analyzed, and then converted into meaningful information, which is then applied in the real world; moreover, this cycle operates at a society-wide level.[40]

Seria, assim, um passo à frente da Indústria 4.0, que diz respeito apenas aos processos produtivos da indústria e do comércio, mas não de outros aspectos da vida individual e coletiva. Na Sociedade 5.0 as pessoas individual e coletivamente seriam o centro do processo tecnológico de disseminação da inteligência sobre objetos e sobre todo o ambiente circundante, ou seja, a culminância da perfilização.

---

[40] HITACHI-UTOKYO LABORATORY (H-UTOKYO LAB). **Society 5.0**: a people-centric super-smart society. Tokyo: Springer, 2018. Edição do Kindle, p.24.
Em resumo, o Sociedade 5.0 apresentará um ciclo iterativo no qual os dados são coletados, analisados e, em seguida, convertidos em informações significativas, que são então aplicadas no mundo real; além disso, este ciclo opera em um nível de toda a sociedade (tradução nossa).

4.1 Ciclo Virtuoso da Inteligência Artificial

Pelo visto, as vantagens do processo de automatização resultam da disseminação de "inteligência" sobre objetos inanimados, de tal maneira a "cognificar"[41] o mundo, fazendo com que objetos, tais como eletrodomésticos, automóveis, móveis, e até a infraestrutura das cidades, colaborem ativamente para ganhos de produção das empresas, melhoria de serviços públicos e aumento da qualidade de vida das pessoas.

Andrew Ng, uma das maiores autoridades no tema do aprendizado de máquina, diz, por essa razão, que a Inteligência Artificial é a nova eletricidade[42]. O caráter transversal dessa tecnologia tende a repetir o que ocorreu com a eletricidade, na virada do século XIX para o XX, isto é, tende a exercer influência sobre todas as áreas da vida humana, assim como se deu com a eletricidade. Desde tarefas domésticas, passando pela agricultura, pela indústria, pelo comércio, pelo entretenimento, enfim, tudo será de algum modo afetado pelo processo de espalhamento da inteligência artificial.

Logicamente, a aposta de que a IA será amplamente incorporada em objetos decorre do fato de que há atrativos muito claros na adoção dos mecanismos inteligentes, de modo que se pode presumir razoavelmente que a implementação dessas tecnologias ocorrerá sem a necessidade de qualquer incentivo adicional.

Adriano Mussa fala de um "Ciclo Virtuoso da Inteligência Artificial" para aqueles que implantarem a IA em seus negócios:

> Em linhas gerais, o ciclo funciona da seguinte forma: se a organização desenvolver um produto ou serviço de qualidade satisfatória, ela conseguirá alguns usuários iniciais. Os usuários iniciais, ao utilizarem o produto ou serviço, gerarão dados que serão coletados e armazenados pela organização. Esses dados, se bem tratados por técnicas de Inteligência Artificial, principalmente *Machine Learning*, possibilitarão a melhoria do produto ou serviço. O produto ou serviço aperfeiçoado levará à aquisição de mais usuários. Mais usuários gerarão mais dados; mais dados levarão à melhoria do produto ou serviço e esse ciclo seguirá continuamente.[43]

Ao contrário do que se pode pensar a partir do imaginário criado especialmente pela indústria cinematográfica, a IA não é uma poderosa e maligna ferramenta capaz até de se rebelar contra os seus criadores. A maioria das aplicações de IA hoje são estreitas (*narrow*), isto é, são direcionadas a finalidades bem específicas e limitadas. Não existe ainda, e provavelmente nunca

---

[41] A expressão é de Kevin Kelly. Cf.:KELLY, Kevin. **Inevitável**: as 12 forças tecnológicas que mudarão nosso mundo. Rio de Janeio:Alta Books, 2019, Tradução de Cristina Yamagami, p. 31-65.

[42] ANDREW Ng: Artificial Intelligence is the New Electricity. Stanford: Stanford Graduate School of Business, 2 fev. 2017. 1 vídeo (1h 27 min). Publicado por Stanford Graduate School of Business. Disponível em: https://www.youtube.com/watch?v=21EiKfQYZXc. Acesso em 30 dez. 2020.

[43] MUSSA, Adriano. **Inteligência Artificial - Mitos e Verdades:** as reais oportunidades de criação de valor nos negócios e os impactos no futuro do trabalho. São Paulo: Saint Paul, 2020. Edição Kindle, p.105.

existirá, uma Inteligência Artificial Geral (AGI, na sigla em inglês para *Artificial General Intelligence*), unificada e com aptidão para quaisquer propósitos.

As aplicações de IA atualmente estabelecem uma relação simples do tipo: A → B, em que "A" representa os dados de entrada (*Input*), "→" indica uma relação de implicação condicional, e "B", os dados de saída (*Output*). Os dados de saída resultam, portanto, do tratamento dos dados de entrada pelo modelo. O modelo cria uma conexão estatístico-matemática entre o *Input* e o *Output*, que emula a conexão semântica estabelecida pela inteligência humana, só que numa escala, precisão e velocidade muito maiores e, em compensação, infinitamente mais estreita e descontextualizada também.

Para que o modelo funcione adequadamente, os programadores "ensinam", com dados de treinamento, qual a conexão "correta" a ser estabelecida. Em seguida, o próprio modelo "aprende" o padrão da conexão e passa replicá-la. Quando já na fase de aplicação, os usuários também acabam ajudando o modelo a melhorar, por meio de suas interações, que nada mais são do que rotulações para o modelo. Por exemplo, quando o usuário dá um *like* num produto, ele rotula aquele produto — e todos os que a ele estão ligados — como um *output* desejável para si, caso posteriormente ele faça uma pesquisa de compra. O mesmo ocorre também quando o usuário, por exemplo, marca um *e-mail* como *spam*: o modelo incorpora esse rótulo como negativo, posteriormente qualificando *e-mails* com o mesmo padrão como *spams*.

Observa-se, assim, que à medida que o modelo entra em contato com os usuários e suas rotulações, salvo interferências propositais do programador, ele vai se amoldando às preferências e repulsões que estes manifestam, potencializando os comportamentos tidos como normais. Isso vale para o indivíduo e para o grupo. Há um processo de *perfilização* constante, individual e grupal.

Vê-se também que o modelo carece de muitos dados para ter acurácia e robustez, pois ele só prediz algo com que já tenha tido contato anterior. Por isso Inteligência Artificial e *Big Data* (grandes conjuntos de dados) andam juntos.

Se os dados de entrada e os dados de saída são conhecidos do programador, a máquina será programada para aprender de modo supervisionado (*supervised learning*-*SL*). Neste caso, o programador, na fase de treinamento, alimenta a máquina com os dados de entrada e também com os dados de saída, de modo a estabelecer o vínculo estatístico-matemático.

Aqui, porém, há uma subdivisão importante: a) o SL pode se dar por meio de *Statistical Machine Learning*, isto é, uma forma em que o algoritmo contém previamente fórmulas para calcular probabilidades e com base nelas gerar o *output*; ou b) por meio de *Deep Learning*-DL,

em que o programador não cria totalmente as fórmulas de cálculo, mas apenas esboça um modelo em camadas aparentes de uma Rede Neural Artificial e depois alimenta essa rede com vastos volumes de dados de entrada, associando-os aos dados de saída "corretos" (rotulados), deixando que o próprio algoritmo, por tentativas e erros, encontre os pesos adequados para cada variável de modo tal que essas associações se encaixem de forma correta. Essas tentativas e erros, quando encerradas, geram camadas profundas e ocultas na Rede Neural Artificial, que o próprio modelo cria e que sequer é do conhecimento do próprio programador.

Grosso modo, no *Statistical Machine Learning* o programador ensina a pergunta, a resposta certa e a forma de chegar a ela; no *Deep Learning*, o programador mostra a pergunta e a resposta certa, mas não diz como chegar a ela, cabendo ao modelo criar esse caminho. E o caminho criado pelo modelo pode ser extremamente eficaz — os modelos de DL têm atingido 95% de acurácia de predição —, embora ele estabeleça conexões que, para nós, humanos, não fazem sentido algum, em termos de relação de causa e efeito.

Adriano Mussa, após explicar como funciona um modelo preditivo baseado em *Statistical Machine Learning*, no qual o programador escolhe as variáveis relevantes (área do imóvel, localização, tempo de construção, etc.) e ensina a máquina qual peso dar a cada uma delas, estima como seria o processo de *Deep Learning* na mesma situação:

> Na prática, alimentamos os algoritmos de DL com a camada de *Input* – A e com os dados de resultado, *Output* – B, e são os algoritmos que buscam todas as combinações possíveis de variáveis, testando a criação de inúmeras camadas e neurônios para buscar, matematicamente, a melhor combinação e pesos, que expliquem os preços dos imóveis com a maior acurácia possível, com base em suas características. Em outras palavras, os algoritmos buscam aumentar a acurácia do modelo utilizando as inúmeras combinações de variáveis, criando neurônios e utilizando pesos que otimizem a sua performance, independentemente de elas fazerem ou não sentido para nós, seres humanos.[44]

Os modelos de DL, portanto, buscam extrair diretamente dos dados de entrada a combinação mais eficiente para chegar aos dados de saída, que por sua vez são rotulados conforme o objetivo do programador. Na concepção desse caminho lógico-matemático, o modelo acaba organizando camadas escondidas (*hidden layers*) de "neurônios artificiais" que atribuem pesos às diferentes combinações, preferindo aquelas cuja soma mais se aproxime do resultado de saída desejado. Em outras palavras, as camadas ocultas trabalham otimizando funções matemáticas que sejam capazes de transformar os dados de entrada na resposta informada pelo desenvolvedor do modelo na fase de treinamento.

---

[44] MUSSA, op.cit, p.86

Mesmo o programador original do modelo não saberá completamente como a Rede Neural Artificial chegou àquela combinação, tal a quantidade de cálculos e de arranjos testados pela máquina. Essas camadas intermediárias, assim, formam uma verdadeira "caixa preta" que oculta a maior parte do processo decisório automático. Assim, se por um lado elas tornam o modelo extremamente robusto para obter as respostas desejadas, por outro elas tornam opaco o processo decisório. Nas camadas escondidas dos modelos está a virtude e o vício do DL.

Quanto mais complexo for o problema a ser resolvido pela Rede Neural Artificial, mais camadas ocultas de combinações e pesos podem ser criadas pelo modelo para aumentar a acurácia. Em compensação, mais obscuros se tornam os critérios de cálculo, ou seja, mais densa a caixa-preta.

Nos modelos de *Statistical Machine Learning*, em que o programador escolhe as variáveis que o modelo deve levar em conta, o que ocorre é que o modelo ficará limitado à visão humana de causalidade, que apenas leva em conta os vínculos fortes entre entrada e saída. Se o mesmo problema de negócio for apresentado a um modelo de *Deep Learning*, ele encontrará correlações que não ocorreriam à mente humana, por aparentemente não terem vínculo de causalidade com o resultado.

Um exemplo impressionante, lembrado por Kai Fu Lee[45], é aquele do modelo criado por uma empresa chinesa para decidir automaticamente sobre a concessão de pequenos empréstimos com base em dados do celular do interessado. Uma Rede Neural Artificial, devidamente treinada com milhões de dados históricos de pequenos empréstimos, descobriu que o nível de bateria médio do celular ao longo do dia, a data de nascimento ou a velocidade de digitação do pedido de empréstimo pelo celular do interessado, tinham correlação com a classificação dele como bom ou mau pagador (os bons pagadores geralmente tinham a bateria do celular mais carregada, por exemplo).

Como isso se dá? Após ter acesso a um vasto conjunto de dados de bons e maus pagadores, o modelo de DL, na fase de treinamento, é apresentado a esses dados, tendo o programador previamente informado (rotulado ou etiquetado) os dados de saída (bons ou maus pagadores). A rede neural então, ante os dados de entrada (os mais diversos dados extraídos dos celulares, tais como tempo de uso diário, nível médio de bateria, sites que navega comumente, etc.), não procura "entender" o porquê de aquele ser um bom ou mau pagador — como faria um ser humano, que pensa em termos de causa e efeito — mas sim criar uma função matemática que ligue de maneira ótima os dados de entrada dos bons pagadores aos dados de saída

[45] Op. cit., p.139.

respectivos, rotulados pelo programador. E o mesmo processo é feito com o telefone de muitos maus pagadores. Ao final desse treinamento, o modelo terá encontrado padrões nos bons e nos maus pagadores, levando em conta elementos que, para um ser humano, seriam completamente irrelevantes, tal como a carga média da bateria do celular, referida acima. Por isso que é apenas metafórica a comparação dos processos decisórios automatizados com a inteligência humana. O que a máquina faz é algo muito diferente do pensamento humano, embora chegue a resultados parecidos e eventualmente com maior acurácia. Edsger Dijkstra, a esse propósito, afirmou: "A questão de saber se um computador pode pensar não é mais interessante que a questão de saber se um submarino pode nadar."[46]

Uma descoberta como essa (que a carga média da bateria do celular influencia na probabilidade de que o contrato seja cumprido) pode representar um *insight* comercial que dá ao operador do modelo uma vantagem relevante, em relação aos concorrentes, sobretudo quando se pensa em grande escala. Mas pode também representar, a depender de qual seja o elemento diferenciador revelado pelos dados, uma fonte involuntária de discriminação de pessoas, grupos ou ideias.

A grande revolução do aprendizado de máquina ocorreu justamente com o *Deep Learning-DL*, e a maioria das atuais aplicações daquilo que se chama de "Inteligência Artificial" nada mais é do que DL. Durante muito tempo, entre os anos 1970 até os anos 1990, prevaleciam nas aplicações de IA os chamados Sistemas Especialistas, que eram mecanismos inteligentes baseados em regras[47]. O programador avaliava o problema do mundo real e tentava modelá-lo por meio de regras, que depois seriam aplicadas por um motor de inferência a novos dados de entrada. A deficiência dessa abordagem é que, por vezes, muito difícil e laborioso criar as regras específicas para cada situação, e mais ainda para as exceções que se intersectam com a regra em alguns pontos. Um programa de reconhecimento da imagem de um gato, por exemplo, dependeria de escrever em código minuciosamente o que "é" um gato e centenas, talvez milhares de regras sobre o que não é um gato, mas sim uma onça, um puma, ou outro felino. Ora, não é assim que funciona a mente humana, a mais avançada forma de inteligência que conhecemos. Simplesmente sabemos muitas coisas que não podemos verbalizar em termos estritos. Aquilo que chamamos de "senso comum", por exemplo, é constituído de um vasto

[46] NORVIG, Peter Peter; NORVIG, Peter. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Elsevier, 2013. Tradução de Regina Célia Simille, p. 932.

[47] SEJNOWSKI, Terrence J. **A revolução do aprendizado profundo**. Rio de Janeiro: Alta Books, 2019. Traduzido por Carolina Gaio, p. 35-37.

conhecimento sobre leis físicas e sociais que não podem ser codificadas, tanto mais porque não sabemos exatamente quais são elas, embora as apliquemos no dia a dia.

Nos anos 1980, Douglas Lenat, por meio de um projeto chamado CYC[48], tentou codificar o "senso comum" de um ser humano. O programa chegou a acumular mais de 1 milhão de regras, sem, no entanto, conseguir abranger algo que um humano comum aprende ainda na infância.

De fato, há muitas coisas que um ser humano sabe, mas não consegue expressar em palavras, e muito menos de forma quantitativa. Santo Agostinho escreveu que sabia o que era o tempo, mas bastava alguém pedir-lhe para dizer o que era, que não sabia mais[49]. Esse célebre pensamento ilustra a maneira como funciona a inteligência humana. Muita coisa é aprendida simplesmente por exemplos e repetições de padrões, sem a necessidade de que a mente analise todos os aspectos e relações do objeto conhecido.

O *Deep Learning* parte de uma abordagem diferente, mais próxima do funcionamento do cérebro humano. As dificuldades enfrentadas pela abordagem baseada em regras e heurísticas, estimulou os pesquisadores em IA a buscar saídas que fossem mais factíveis. Segundo Sejnowski[50], quatro coisas indicavam que era ruim trilhar pelo caminho da criação de regras para desenvolver um sistema inteligente, porque: a) o cérebro humano trabalha primeiro com reconhecimento de padrões, as regras surgem depois; b) é preciso uma prática repetitiva para que o cérebro domine atividades mais complexas; c) o cérebro não se orienta por regras no dia a dia, embora possa trabalhar com elas em um nível mais profundo do pensamento; e, finalmente, d) nossos cérebros têm bilhões de neurônios que se intercomunicam, o que sugere que ele trabalha com processamento paralelo dos dados de entrada e não com processamento linear (arquitetura de Von Neumman).

Foi essa abordagem que permitiu a maior parte dos progressos efetivos na área de IA aplicada a negócios. Os modelos de *deep learning* muitas vezes atingem 95% de acurácia, em certas tarefas, o que era impensável antes do uso dessa técnica. E esse nível de acurácia, como explica Adriano Mussa, não é raro em *Deep Learning*:

---

[48] https://www.cyc.com/the-cyc-platform. Acesso em 20 dez 2020

[49] "O que é, por conseguinte, o tempo? Se ninguém mo perguntar, eu sei; se o quiser explicar a quem me fizer a pergunta, já não sei". (AGOSTINHO, Santo. **Confissões**. São Paulo: Companhia das Letras, 2017. Tradução de Lorenzo Mammì. Edição do Kindle, p. 237.)

[50] SEJNOWSKI, op. cit., p. 41-42.

> Esse percentual elevado de acurácia dos algoritmos de *Deep Learning* não é exceção. Ele tem sido observado em uma infinidade de aplicações de setores e contextos diferentes, mostrando sua forte robustez.[51]

Assim, os modelos de *Deep Learning* asseguram aa principais vantagens do uso de IA em negócios ou qualquer aplicação que dependa de julgamentos: rapidez, economia e acurácia.

5 OS RISCOS DAS DECISÕES AUTOMATIZADAS

Como demonstrado no item anterior, a maior parte das decisões automatizadas que são atualmente colocadas em prática resultam de modelos de *Deep Learning* em Redes Neurais Artificiais. Convém, assim, ter presente que os riscos que foram levados em conta pelo próprio legislador estão associados a esse método de tratamento de dados.

O primeiro e mais conhecido risco das decisões automatizadas decorrentes de DL é o da opacidade. Pelo próprio volume de cálculos e pela quantidade de dados necessária para a concepção de um modelo de DL, não é acessível sequer para os desenvolvedores o processo exato por meio do qual o modelo chegou a esta ou aquela predição ou mesmo decisão, e isso naturalmente pode levantar desconfianças e suposições em relação à higidez do modelo, eventualmente exigindo uma coadjuvação humana para que ele possa ser colocado em prática. Terrence Sejnowski[52] exemplifica bem o problema com o caso dos diagnósticos médicos:

> Embora possam dar resposta correta para um problema, atualmente não sabemos como as redes neurais chegam a ela. Por exemplo, suponha que uma paciente chegue a um pronto-socorro com uma dor aguda no peito. Trata-se de um infarto agudo do miocárdio, o que precisa de intervenção imediata, ou simplesmente um caso grave de indigestão? Uma rede treinada para diagnosticar pode ser mais precisa do que o médico responsável pela triagem; mas, sem uma explicação sobre como a rede tomou a decisão, a relutância em confiar nela seria plausível. Os médicos também são treinados para acompanhar o que equivale a algoritmos, séries de testes e pontos de decisão que os orientam em casos de rotina. O problema é que há casos raros, que estão fora do escopo de seus 'algoritmos', enquanto uma rede neural treinada com muito mais casos, mais do que a média dos médicos verá em toda uma vida, pode muito bem dirimir sobre esses casos raros. Mas você confiaria mais no diagnóstico estatisticamente mais sólido de uma rede neural, sem explicação de como foi feito, do que no de um médico com um diagnóstico plausível?

A opacidade também pode decorrer de fatores comerciais. O desenvolvedor do modelo pode até saber explicar como se chegou a certa decisão, mas a exposição desse caminho poderia revelar o seu "segredo comercial ou industrial", que na verdade é a sua fonte de ganhos com o modelo.

---

[51] MUSSA, op.cit.,p. 91.

[52] Op.cit., p. 134.

A LGPD cuida do ponto, na esteira do GDPR, estabelecendo no art. 20, §1º, um direito à explicação em caso de decisão automatizada, como primeira linha contra a opacidade, mas respeitado o segredo comercial e industrial — e é difícil, na prática, conciliar essas duas coisas.

Se a explicação não for dada pelo controlador ao titular dos dados, sob o argumento da existência de segredo comercial ou industrial, então a Autoridade Nacional de Proteção de Dados - ANPD pode ser acionada para fazer uma verificação sobre possíveis vieses discriminatórios (LGPD, art. 20, §3º). No capítulo seguinte avalia-se melhor essa questão, mas de logo chama a atenção a estreiteza da norma, que deixa duas importantes questões em aberto: a) E se a explicação for negada com outro fundamento, que não o segredo comercial ou industrial? (Por exemplo, a alegação de que o próprio controlador não sabe exatamente como o modelo funciona); b) E se não houver discriminação, mas sim outro tipo de violação a direitos individuais?

Os tecnólogos têm tentado criar modelos que sejam capazes de ser autoexplicativos, a chamada Inteligência Artificial Explicável (XAI, na sigla em inglês para *Explanable Artificial Intelligence*). Mas aqui a questão esbarra na autorreferência. É que o próprio cérebro humano é também uma caixa-preta. De fato, a objeção de que um modelo opera com uma caixa-preta pode ser aplicada também ao cérebro humano. Não há até aqui conhecimento objetivo e minucioso sobre os processos decisórios humanos, exceto que se sabe que há muito mais viés e irracionalidade do que se imaginava. Eventuais explicações humanas, muitas vezes, são meramente retóricas. Modelos de XAI podem recair no mesmo impasse. O risco de opacidade, portanto, permaneceria sendo um problema insolúvel.

Outro ponto, ainda sobre a opacidade, é que a concepção de uma decisão automatizada envolve o tratamento de muitos dados e a revelação do seu processo para um titular poderia ensejar a violação da intimidade de outros titulares, cujos dados também foram levados em conta na decisão, para efeito de comparação, e a quem pode não interessar a divulgação do processo decisório.

Um segundo risco criado pelas Redes Neurais Artificiais de *Deep Learning* é que elas dependem de um grande volume de dados para funcionarem bem, o que gera uma corrida por dados pessoais. Com efeito, a acurácia das decisões automatizadas criadas por Redes Neurais Artificiais depende fundamentalmente de um vasto conjunto de dados (*Big Data*), especialmente na fase de treinamento e validação, e também para evoluir na fase de aplicação. Essa necessidade faz com que aumentem os riscos ligados à privacidade, porque os desenvolvedores buscarão sempre ter acesso ao máximo de dados pessoais para criarem,

validarem e aplicarem modelos preditivos e decisórios de alto desempenho. Com o advento da Internet 5G e a implantação da Internet das Coisas, estima-se que a coleta de dados crescerá exponencialmente, já que atividades triviais do dia a dia e até da intimidade doméstica, como abrir uma geladeira, fechar uma porta, ou ligar uma lâmpada, poderão ser incorporadas à internet e gerarão dados pessoais suscetíveis de serem usados em modelos preditivos. Presumivelmente, isso multiplicará muitas vezes os riscos à privacidade.

Um terceiro risco, ligado ao anterior, é que, além de precisarem de um grande volume de dados, as Redes Neurais expressam apenas o conhecimento que se pode extrair desses dados, não mais que isso. Logo, se há no conjunto de dados de treinamento um viés, proposital ou não, esse viés se replicará indefinidamente nas decisões.

O caso mais conhecido sobre isso ocorreu em Los Angeles (EUA)[53], e dizia respeito ao reconhecimento facial em locais públicos. Descobriu-se que um modelo de reconhecimento facial da polícia cometia mais erros em relação a negros do que em relação a brancos, porque, enquanto os "procurados", na fase de aplicação do modelo, eram na maior parte pessoas negras, na fase de treinamento o modelo fora apresentado a um número maior de faces brancas, tornando-se naturalmente melhor em reconhecer estas do que outras.

Tal situação pode se repetir em muitas outras áreas. A escolha dos dados de treinamento não é um ato neutro; muito menos o é a rotulação dos dados de saída, feita pelos programadores. Aqui a escolha envolve aspectos ideológicos, muitas vezes inconscientemente. Como explica Terrence Sejnowski:

> Todas as redes neurais que classificam entradas são tendenciosas. Em primeiro lugar, a escolha das categorias de classificação incorpora um viés que reflete o preconceito humano na forma como esmiuçamos o mundo. Por exemplo, seria útil treinar uma rede para detectar ervas daninhas em gramados. Mas como identificá-la? A erva daninha de um homem pode ser a flor silvestre de outro. A classificação é um problema muito mais amplo, que reflete vieses culturais. Essas ambiguidades precisam integrar os conjuntos de dados usados para treinar a rede.[54]

Pior ainda, com a aplicação do modelo em massa, produzem-se *loopings* que reforçam o viés original. Cathy O´Neal[55] exemplifica esse fenômeno com os modelos de otimização do policiamento ostensivo. Como esses modelos usam dados relativos a pequenas infrações, tais

---

[53] GARVIE, Clare; FRANKLE, Jonathan. Facial-Recognition Software Might Have a Racial Bias Problem: Depending on how algorithms are trained, they could be significantly more accurate when identifying white faces than African American ones.. *In*: THE ATLANTIC. **The Atlantic**, 6 abr. 2016. Disponível em: https://www.theatlantic.com/technology/archive/2016/04/the-underlying-bias-of-facial-recognition-systems/476991/. Acesso em: 22 dez. 2020.

[54] SEJNOWSKI, op.cit., p. 135.

[55] O'NEIL, Cathy. **Weapons of math destruction**: how big data increases inequality and threatens democracy. New York:Crown Publishers, 2016. Ebook.

como perturbação da ordem, posse de pequena quantidade de droga e vadiagem, os policiais acabam sendo enviados para patrulhar regiões pobres, onde normalmente acontecem essas infrações. Com o aumento do patrulhamento, aumentam também as prisões por essas pequenas contravenções, o que induz a realimentação e o reforço por *feedback* ao modelo para aumentar o patrulhamento nesses locais.

Ainda no campo dos vieses e seu ciclo vicioso de reforço, observa Caathy O'Neal também que, embora a cor da pele a condição social não sejam incluídas no modelo como parâmetros para a inferência, o fato é que os dados escolhidos (sobre pequenos delitos) para esse tipo de policiamento acabam funcionando como *proxies* para a raça e a pobreza, já que apenas negros e hispânicos da periferia são presos, segundo a autora, por esse tipo de crime nos Estados Unidos. Um indivíduo branco que pratique ações semelhantes num campus universitário dificilmente deparará com uma patrulha policial.

Há um quarto risco, não menos grave, no uso de mecanismos inteligentes para formulação de decisões automatizadas. É que a grandeza que é escolhida para ser otimizada pode subdimensionar outras questões relevantes. Assim, se o modelo visa ao lucro — e a maioria visa a isso, naturalmente — a função de lucro deve ser otimizada pelo modelo, no que não há nada de ilegal ou imoral. Acontece que essa otimização, quando feita em termos matemáticos, é implacável. O modelo não se deterá diante de nenhuma circunstância, a não ser que programado para isso, para aumentar os lucros. Como mecanismo de Inteligência Artificial Fraca ou Estreita, o modelo não é capaz de contextualizar as decisões para além dos dados que lhes foram apresentados, de modo que se o lucro é o que deve ser maximizado, ele fará isso *per fas et per nefas*.

Eventualmente, essa "objetividade" inexorável pode produzir danos imensos, sobretudo quando aplicada em grande escala. E aqui se chega a um risco transversal de todos os modelos matemáticos para produzir decisões automatizadas: a escala. É a escala que gera os maiores danos.

Como explica Cathy O´Neal[56], é a escala transforma o que seria um pequeno incômodo em algo com a força de um tsunami. Ao estabelecer um ciclo de decisão em um número imenso de casos idênticos, o modelo em larga escala acaba influenciando o ambiente de duas formas: a) ele reforça em massa um padrão, inferido de situações anteriores (que podem ser injustas); b) ele induz o comportamento futuro das pessoas, que tentarão se ajustar ao modelo.

---

[56] O'NEIL, op.cit., p. 48.

A escala das decisões automatizadas gera um problema adicional. A regulamentação ou qualquer ação legal que vise a solucionar um problema gerado por algoritmos pode não conseguir atingir o seu objetivo, justamente por não ser escalável. Ou seja, enquanto decisões automatizadas são tomadas *on-line* e em massa, as soluções legislativas tendem a depender de uma análise artesanal, caso a caso. A brutal diferença de velocidade e de volume pode levar a norma à completa ineficácia prática. Assim, as regulamentações precisarão contar com mecanismos de implementação escaláveis. Nesse sentido, observam Kearns & Roth[57]:

> Regulations and laws certainly have a crucial role to play—as we have emphasized throughout, the specification of what we want algorithms to do and not do for us should remain firmly in the human and societal arenas. But purely legal and regulatory approaches have a major problem: they don't scale. Any system that ultimately relies solely or primarily on human attention and oversight cannot possibly keep up with the volume and velocity of algorithmic decision-making. The result is that approaches that rely only on human oversight either entail largely giving up on algorithmic decision-making or will necessarily be outmatched by the scale of the problem and hence be insufficient. So while laws and regulations are important, we have argued in this book that the solution to the problems introduced by algorithmic decision-making should itself be in large part algorithmic.[58]

Em resumo, os principais riscos dos mecanismos de decisão automatizada são: a) opacidade; b) necessidade de grande volume de dados, com riscos à privacidade; c) viés; d) subdimensionamento de grandezas diversas daquela buscada pelo controlador dos dados; e) escala.

6 CONCLUSÃO

O presente trabalho buscou investigar o que são as decisões automatizadas, quais são os riscos e benefícios que elas trazem, bem como analisar de que maneira o recente marco legal brasileiro sobre o tema, a Lei Geral de Proteção de Dados, trata essa inovação tecnológica. A metodologia adotada para desenvolver o problema de pesquisa gerou algumas conclusões.

Verificou-se que existe a necessidade doutrinária de definir o que é e de como se forma uma decisão automatizada, à luz da LGPD. Intentando atingir esse objetivo, esboçou-se a

---

[57] KEARNS; ROTH, op.cit., p.192.

[58] Regulamentos e leis certamente têm um papel crucial a desempenhar - como enfatizamos ao longo do texto, a especificação do que desejamos que os algoritmos façam e não façam por nós deve permanecer firmemente nas arenas humana e social. Mas as abordagens puramente legais e regulatórias têm um grande problema: elas não escalam. Qualquer sistema que, em última análise, dependa única ou principalmente da atenção e supervisão humanas, não pode acompanhar o volume e a velocidade da tomada de decisão algorítmica. O resultado é que as abordagens que dependem apenas da supervisão humana implicam em desistir amplamente da tomada de decisão algorítmica ou serão necessariamente superadas pela escala do problema e, portanto, insuficientes. Portanto, embora as leis e os regulamentos sejam importantes, argumentamos neste livro que a solução para os problemas introduzidos pela tomada de decisão algorítmica deve ser em grande parte algorítmica (tradução nossa).

seguinte definição: decisão automatizada é todo julgamento feito exclusivamente por máquina, com base em predição decorrente de tratamento automatizado de dados pessoais de entrada, segundo um modelo ou algoritmo condicionado por dados de treinamento, que afete imediatamente interesse juridicamente tutelado de pessoa natural, excetuados aqueles que tenham fins particulares e não econômicos, jornalísticos ou científicos.

A investigação permitiu demonstrar a relação que pode ser estabelecida entre perfilização e decisões automatizadas, evidenciando como a a LGPD trata essa questão. Como visto, as decisões automatizadas podem ser realizadas com ou sem definição de perfis, da mesma forma como a definição de perfis pode ocorrer sem que dela decorra uma decisão automatizada.

Por outro lado, a definição de perfis e as decisões automatizadas não constituem necessariamente atividades separadas, a decisão automatizada pode ou não depender de perfilização, de acordo com o interesse do desenvolvedor na concepção do modelo.

Todavia, a LGPD atenua essas distinções, uma vez que, segundo a sua redação, deve-se admitir que a perfilização, mesmo não sendo seguida de uma decisão automatizada, mas sim de uma decisão humana assistida por máquina, seja considerada como uma decisão automatizada por equiparação, tendo em vista que seu artigo 20 afirma expressamente que estão incluídas entre as decisões automatizadas aquelas "destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade".

Foram analisados também benefícios e riscos dessa forma decisória. Demonstrou-se que um dos principais benefícios que faz com que as empresas e governos busquem por processos decisórios automatizados é o aumento da capacidade e da velocidade de resposta a demandas repetitivas e a redução de custos, quando máquinas tentam imitar a racionalidade humana, tomando decisões em massa, diante de um conjunto relevante de dados que permitem construir um modelo replicador das decisões.

Na outra ponta do processo decisório, restou evidente que também é possível encontrar benefícios, na medida que os consumidores e usuários de serviços públicos encontram respostas cada vez mais personalizadas e, consequentemente, mais adequadas às necessidades específicas de cada indivíduo ou família.

Entre os principais riscos das decisões automatizadas, cinco foram identificados: a) opacidade; b) necessidade de grande volume de dados, com riscos à privacidade; c) viés; d) subdimensionamento de grandezas diversas daquela buscada pelo controlador dos dados; e) escala.

O risco da opacidade decorre de modelos de *Deep Learning* e consiste na ideia de que o processo exato pelo qual levou a uma ou outra decisão carrega tantos dados e exige inúmeros cálculos que não é acessível nem mesmo para seus desenvolvedores, levantando dúvidas e desconfianças em relação à higidez do processo de decisão automatizada. Em outros casos, a opacidade pode decorrer de questões estratégicas, uma vez que, mesmo tendo um caminho conhecido pelos desenvolvedores, divulgar o processo decisório pode não ser interessante por revelar segredos comerciais ou industriais.

A pesquisa levou a concluir também que a LGPD tratou desse risco e garantiu o direito à explicação em casos de decisões automatizadas, tentando equilibrar a defesa contra a opacidade e ao mesmo tempo garantir a preservação ao segredo comercial e industrial.

O segundo risco identificado consiste no fato de que as decisões automatizadas são dependentes de um grande volume de dados para funcionarem corretamente, fato que gera uma corrida desenfreada por dados e isso, consequentemente, põe em risco a privacidade das pessoas.

O terceiro risco encontrado consiste nas decisões automatizadas se basearem apenas no que se pode extrair dos dados e em nada mais. Assim, caso exista no conjunto de dados de treinamento um viés, proposital ou não, esse viés se replicará indefinidamente nas decisões.

O quarto e o quinto riscos encontrados estão, de certa forma, conectados. O quarto risco consiste no uso de mecanismos inteligentes para a formulação de decisões inteligentes, uma vez que a grandeza que é escolhida para ser otimizada pode subdimensionar outras questões relevantes, direcionando a decisão mais para uma posição ou outra de acordo com os interesses envolvidos e quando aplicadas em escalas grandes podem produzir danos enormes. E esse problema da escala foi o quinto risco encontrado, tendo em vista a escala transforma o que seria um simples problema de uma decisão automatizada em um problema gigantesco, reforçando em massa um padrão, induzindo comportamentos futuros e a correção desse grande volume de decisões inadequadas pode ser totalmente ineficaz por precisar de análises humanas, bem mais lentas.

Desse modo, percebe-se que as decisões automatizadas já podem ser consideradas um novo paradigma, trazendo muitos benefícios e riscos, alguns já enfrentados pelo marco legal brasileiro que trata sobre o tema, a Lei Geral de Proteção de Dados, e outros que ainda vão precisar ser enfrentados.

REFERÊNCIAS

ABREU, Jacqueline de Souza. Tratado de Proteção deTratamento de Dados Pessoais para Segurança Pública: contornos do regime jurídico pós-LGPD. In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang; RODRIGUES JUNIOR, Otavio Luiz; BIONI, Bruno (org.). **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021. Edição do Kindle.

AGGARWAL, Charu C.. **Neural Networks and Deep Learning**: a textbook. New York: Springe, 2018.

AGOSTINHO, Santo. **Confissões**. São Paulo: Companhia das Letras, 2017. Tradução de Lorenzo Mammì. Edição do Kindle.

AGRAWAL, A; GANS, J.; GOLDFARB. **Máquinas preditivas***:* a simples economia da Inteligência Artificial. Rio de Janeiro: Alta Books, 2018. Tradução de Wendy Campos.

AMBROSE, Meg Leta; AMBROSE, Ben M.. When robots lie a comparison of auto-defamation law. **2014 Ieee International Workshop On Advanced Robotics And Its Social Impacts**, [S.L.], p. 56-61, set. 2014. IEEE. http://dx.doi.org/10.1109/arso.2014.7020980.

ANDREW Ng: Artificial Intelligence is the New Electricity. Stanford: Stanford Graduate School of Business, 2 fev. 2017. 1 vídeo (1h 27 min). Publicado por Stanford Graduate School of Business. Disponível em: https://www.youtube.com/watch?v=21EiKfQYZXc. Acesso em 30 dez. 2020.

ANDRADE, Manuel A. Domingues. **Teoria geral da relação jurídica**. Coimbra: Almedina, 1992. v. 1.

BREEN, Jason. YouTube or YouLose? Can YouTube Survive a Copyright Infringement Lawsuit. **Bepress Legal Series.Working Paper 1950**, Los Angeles, p. 1-37, 18 jan. 2007. Disponível em: https://law.bepress.com/cgi/viewcontent.cgi?article=9209&context=expresso. Acesso em: 10 dez. 2020.

BINNS, Reuben; GALLO, Valeria. Automated Decision Making: the role of meaningful human reviews. *In*: ICO. **Information Commissioner's Office**. 12 abr. 2019. Disponível em: https://ico.org.uk/about-the-ico/news-and-events/ai-blog-automated-decision-making-the-role-of-meaningful-human-reviews/. Acesso em: 11 jan. 2021.

CHACE, Calum. **Surviving AI**: the promise and peril of artificial intelligence. Oxford: Three Cs, 2015. Kindle Edition.

CLARKE, Roger. Profiling: a hidden challenge to the regulation of data surveillance. **Journal Of Law, Information And Science**, v. 2, n. 4, jan. 1993. Disponível em: http://www.austlii.edu.au/cgi-bin/sinodisp/au/journals/JlLawInfoSci/1993/26.html?query=. Acesso em: 10 dez. 2020.

CITRON, Danielle Keats. Technological Due Process. **Washington University Law Review**, Washington, D.c., v. 85, n. 6, p. 1249-1313, ago. 2008. Disponível em:

https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview. Acesso em: 17 jul. 2020.

CRAWFORD, Kate; SCHULTZ, Jason. Big Data and Due Process: toward a framework to redress predictive privacy harms. **Boston College Law Review**, Boston, v. 55, n. 1, p. 93-128, 29 jan. 2014. Disponível em: https://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4. Acesso em: 17 jul. 2020.

ERTEL, Wolfgang. Introduction. **Undergraduate Topics In Computer Science**, p. 1-21, 2017. Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-58487-4_1.

GARVIE, Clare; FRANKLE, Jonathan. Facial-Recognition Software Might Have a Racial Bias Problem: Depending on how algorithms are trained, they could be significantly more accurate when identifying white faces than African American one. *In*: THE ATLANTIC. **The Atlantic**, 6 abr. 2016. Disponível em: https://www.theatlantic.com/technology/archive/2016/04/the-underlying-bias-of-facial-recognition-systems/476991/. Acesso em: 22 dez. 2020.

HANNÁK, Anikó; WAGNER, Claudia; GARCIA, David; MISLOVE, Alan; STROHMAIER, Markus; WILSON , Christo. Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr. **In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing**, New York, p. 1914–1933, Fev. 2017.

HILDEBRANDT, Mireille. Privacy as Protection of the Incomputable Self: from agnostic to agonistic machine learning. **Theoretical Inquiries In Law**, Tel Aviv, v. 20, n. 1, p. 83-121, jan. 2019. Disponível em: https://www7.tau.ac.il/ojs/index.php/til/article/view/1622/1723. Acesso em: 17 jul. 2020.

HITACHI-UTOKYO LABORATORY (H-UTOKYO LAB). **Society 5.0**: a people-centric super-smart society. Tokyo: Springer, 2018. Edição do Kindle.

ICO. What does the UK GDPR say about automated decision-making and profiling?. *In*: ICO. **Information Commissioner's Office**. Disponível em: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-does-the-uk-gdpr-say-about-automated-decision-making-and-profiling/. Acesso em: 11 jan. 2021.

JUSTICE AND CONSUMERS (Europea Union). Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. *In*: JUSTICE AND CONSUMERS (Europea Union). **European Commission**. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 11 jan. 2021.

KEARNS, Michael; ROTH, Aaron. **The Ethical Algorithm**: the science of socially aware algorithm design. New York: Oxford University Press, 2019. Edição Kindle.

KELLY, Kevin. **Inevitável**: as 12 forças tecnológicas que mudarão nosso mundo. Rio de Janeio:Alta Books, 2019, Tradução de Cristina Yamagami.

KUBAT, Miroslav. **An Introduction to Machine Learning**. 2. ed. Coral Gables: Springer, 2017.

LEE, Tian-Shyug; CHEN, I-Fei. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. **Expert Systems with Applications**, [*s. l.*], v. 28, n. 4, p. 743-752, mai. 2005.

LÉVY, Pierre. **As tecnologias da inteligência**: o futuro do pensamento na era da informática. São Paulo: Editora 34, 1993. Tradução de Carlos Irineu da Costa.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY (ed.). **MIT Technology Review**. Self-driving cars. Topics. Disponível em: https://www.technologyreview.com/topic/smart-cities/self-driving-cars/. Acesso em: 02 jun. 2020.

MENDES, Laura Schertel. **Privacidade, proteção de dados e defesa do consumidor**: linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014.

MIT TECHNOLOGY REVIEW INSIGHTS. How AI is humanizing health care: Artificial intelligence is helping health-care professionals do their jobs better, giving them the tools to build a smarter, more efficient ecosystem. *In*: MASSACHUSETTS INSTITUTE OF TECHNOLOGY (ed.). **MIT Technology Review**. [*S. l.*], 22 jan. 2020. Disponível em: https://www.technologyreview.com/2020/01/22/276128/how-ai-is-humanizing-health-care/. Acesso em: 2 jun. 2020.

MONTI, Andrea. Tribunale di Milano: Ord. 24 marzo 2011. *In*: MONTI, Andrea. **ICT LEX**: Diritto, politica, cultura della Rete. [*S.l.*], 24 mar. 2011. Disponível em: https://www.ictlex.net/?p=1285. Acesso em: 7 jan. 2021.

MUSSA, Adriano. **Inteligência Artificial - Mitos e Verdades**: as reais oportunidades de criação de valor nos negócios e os impactos no futuro do trabalho. São Paulo: Saint Paul, 2020. Edição Kindle.

NORVIG, Peter Peter; NORVIG, Peter. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Elsevier, 2013. Tradução de Regina Célia Simille.

O'NEIL, Cathy. **Weapons of math destruction:** how big data increases inequality and threatens democracy. New York:Crown Publishers, 2016. Ebook.

PERRY, Walter L.; MCINNIS, Brian; PRICE, Carter C.; SMITH, Susan C.; HOLLYWOOD, John S. **Predictive Policing**: The Role of Crime Forecasting in Law Enforcement Operations. [*S. l.*]: RAND Corporation, 2013. E-book.

RAHWAN, Iyad *et al*. Machine behaviour. **Nature**, [*s. l.*], v. 568, p. 477–486, 24 abr. 2019. Disponível em: https://www.nature.com/articles/s41586-019-1138-y. Acesso em: 2 jun. 2020.

SCHILLER, Arnold; WEISKOPF, Tobias. Automated Censorship in the Digital Space. *In*: YOUNG EUROPEAN FEDERALISTS (Europe). **The New Federalist**, 1 maio 2019. Tradução de Nora Teuma. Disponível em: https://www.thenewfederalist.eu/automated-censorship-in-the-digital-space?lang=fr. Acesso em: 9 dez. 2020.

SEJNOWSKI, Terrence J. **A revolução do aprendizado profundo**. Rio de Janeiro: Alta Books, 2019. Traduzido por Carolina Gaio

THOMSON, Amy; BODONI, Stephanie. Google CEO Thinks AI Will Be More Profound Change Than Fire. *In*: **Bloomberg**. [*S. l.*], 22 jan. 2020. Disponível em:

https://www.bloomberg.com/news/articles/2020-01-22/google-ceo-thinks-ai-is-more-profound-than-fire. Acesso em: 2 jun. 2020.

WEBER, Max. A objetividade do conhecimento nas ciências sociais. In: FERNANDES, Florestan (org.). **Weber**: sociologia. São Paulo: Ática, 1999. Coleção Grandes Cientistas Sociais.

civilistica.com revista eletrônica de direito civil

Submissões

Submissões

Fila Arquivos Ajuda

Minhas Submissões Designadas

Nova Submissão

Buscar

763	**Nazareno César Moreira Reis, ...** Decisões automatizadas: definiç...	Submissão

1 de 1 submissões

Platform & workflow by OJS / PKP

Submissões

Decisões automatizadas

Nazareno César Moreira Reis, Ga...

Submissão Avaliação Edição de Texto

Editoração

Arquivos da Submissão

Buscar

1598-1	nazareno123, Artigo. Decisões Automatizadas.doc	28 de outubro de 2021	Texto integral

Baixar Todos os Arquivos

Discussão da pre-avaliação

Adicionar comentários

Nome	De	Última resposta	Respostas	Fechado

Nenhum item

Platform & workflow by OJS / PKP

Nazareno César Moreira Reis

Endereço para acessar este CV: http://lattes.cnpq.br/2569075829621769
ID Lattes: **2569075829621769**
Última atualização do currículo em 25/06/2020

Possui graduação em direito pela Universidade Federal do Piauí (1997) e especialização em Direito Tributário e Finanças Públicas pelo Instituto Brasileiro de Direito Público (2004). Atualmente é Juiz Federal da Justiça Federal - Seção Judiciária do Estado do Piauí e professor do Instituto de Ciências Jurídicas e Sociais Professor Camillo Filho. Tem experiência na área de Direito, com ênfase em Lógica Jurídica, Direito Constitucional, Direito Processual e Novas Tecnologias aplicadas ao Direito **(Texto informado pelo autor)**

Identificação

Nome	Nazareno César Moreira Reis
Nome em citações bibliográficas	REIS, N. C. M.
Lattes iD	http://lattes.cnpq.br/2569075829621769

Endereço

Endereço Profissional	JUSTIÇA FEDERAL- SEÇÃO JUDICIÁRIA DO PIAUÍ, 1a Vara Federal. AVENIDA MIGUEL ROSA, 7315 REDENÇÃO 64018-550 - Teresina, PI - Brasil Telefone: (86) 21071915 Fax: (86) 21072991

Formação acadêmica/titulação

2004 - 2004	Especialização em Direito Tributário e Finanças Públicas. (Carga Horária: 360h). Instituto Brasiliense de Direito Público, IDP, Brasil. Título: O recurso extraordinário na representação de inconstitucionalidade estadual. Orientador: Gilmar Ferreira Mendes.
1994 - 1997	Graduação em direito. Universidade Federal do Piauí, UFPI, Brasil.

Formação Complementar

2002 - 2002	Estágio Profissionalizante para Magistrados. (Carga horária: 95h). Escola Superior de Magistratura do Estado do Piauí, ESMEPI, Brasil.
1998 - 1999	Curso de Preparação à Magistratura Trabalhista. (Carga horária: 360h). Escola Superior da Magistratura Trabalhista da VI Região, ESMATRAVI, Brasil.
1997 - 1997	Programa de Formação de Procurador Autárquico. (Carga horária: 360h). Universidade de Brasília, UnB, Brasil.

Atuação Profissional

Instituto de Ciências Jurídicas e Sociais Professor Camillo Filho, ICF, Brasil.

Vínculo institucional

2015 - Atual	Vínculo: , Enquadramento Funcional: Professor, Carga horária: 4
Outras informações	Disciplina: Direito Processual Civil I

Justiça Federal - Seção Judiciária do Estado do Piauí, JF/PI, Brasil.

Vínculo institucional

2003 - Atual Vínculo: , Enquadramento Funcional: Juiz Federal, Carga horária: 30

Escola Superior de Magistratura do Estado do Piauí, ESMEPI, Brasil.

Vínculo institucional

2008 - 2010 Vínculo: Professor, Enquadramento Funcional: Professor de Direito Processual Civil, Carga horária: 4

Instituto Nacional do Seguro Social, INSS/DF, Brasil.

Vínculo institucional

1997 - 2002 Vínculo: , Enquadramento Funcional: Procurador Autárquico

Tribunal de Justiça do Estado do Piauí, TJPI, Brasil.

Vínculo institucional

2002 - 2003 Vínculo: , Enquadramento Funcional: Juiz de Direito

Áreas de atuação

1. Grande área: Ciências Sociais Aplicadas / Área: Direito / Subárea: Teoria do Direito/Especialidade: Lógica Jurídica.
2. Grande área: Ciências Sociais Aplicadas / Área: Direito / Subárea: hermenêutica jurídica.
3. Grande área: Ciências Sociais Aplicadas / Área: Direito / Subárea: Teoria do Direito.
4. Grande área: Ciências Sociais Aplicadas / Área: Direito / Subárea: Direito Público.

Idiomas

Inglês Compreende Razoavelmente, Fala Razoavelmente, Lê Razoavelmente, Escreve Pouco.

Produções

Produção bibliográfica

Artigos completos publicados em periódicos

Ordenar por

Ordem Cronológica

1. **REIS, N. C. M.**. O Judiciario na sociedade da informacao. Revista Jurídica Consulex, v. XV, p. 29, 2011.

2. **REIS, N. C. M.**. Reflexões sobre a penhora on line. Revista do Tribunal Regional Federal 1. Regiao, v. 20, p. 35-41, 2008.

3. **REIS, N. C. M.**. Ação civil pública. Proteção do patrimônio histórico e arquitetônico de Oeiras. [Sentenças].. Revista de Direito Ambiental, v. 12, p. 347-362, 2007.

4. **REIS, N. C. M.**. A oralidade nos juizados especiais cíveis federais. Revista do Tribunal Regional Federal 1. Região, Brasília/DF, v. 10, n.ano 16, p. 46-52, 2004.

5. **REIS, N. C. M.**. A relativização do ônus da prova e a justiça constitucional: uma breve reflexão sobre a concretização de valores constitucionais em face da inércia legislativa. Revista do Tribunal Regional Federal 1. Região, Brasília/DF, v. 8, n.Ano 15, p. 21-29, 2003.

6. **REIS, N. C. M.**. ? Considerações sobre a definição de ?erros de cálculo? e ?inexatidões materiais? (art. 463 do CPC) nos processos de execução movidos contra entes públicos?. Revista da Procuradoria Geral do INSS, Brasília/DF, v. 6, n.3, p. 91-98, 1999.

Capítulos de livros publicados

1. **REIS, N. C. M.**; **REIS, N. C. M.** . Reserva legal do Código Florestal e o novo proprietário do imóvel: comentários ao recurso especial nº 222.349/PR da primeira turma do Superior Tribunal de Justiça. In: Vladimir Passos de Freitas. (Org.). Julgamentos históricos do direito ambiental. 1ed.Campinas: Millennium, 2010, v. , p. 77-90.

Outras produções bibliográficas

1.

REIS, N. C. M.. Projeto de Código de Processo Civil e o processo eletrônico: um risco de caducidade precoce. Teresina: Jus Navigandi, 2010 (Artigo).

2. **REIS, N. C. M.**. Por que a arbitragem não é jurisdição?. Teresina: Jus Navigandi, 2009 (Artigo).
3. **REIS, N. C. M.**. O procedimento na execução por quantia certa por título extrajudicial contra deveO procedimento na execucao por quantia certa por titulo extrajudicial contra devedor solvente. Teresina: Jus Navigandi, 2007 (Artigo).

Página gerada pelo Sistema Currículo Lattes em 05/11/2021 às 11:33:43

Imprimir currículo