

RAPHAEL LOBATO COLLET JANNY TEIXEIRA

DO PRINCÍPIO DA MORALIDADE ADMINISTRATIVA COMO FUNDAMENTO
ÉTICO-JURÍDICO PARA REGULAÇÃO DE INTELIGÊNCIA ARTIFICIAL NA
ADMINISTRAÇÃO PÚBLICA

NITERÓI

2023

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Ficha catalográfica automática - SDC/BFD
Gerada com informações fornecidas pelo autor

T266p Teixeira, Raphael Lobato Collet Janny
DO PRINCÍPIO DA MORALIDADE ADMINISTRATIVA COMO FUNDAMENTO
ÉTICO-JURÍDICO PARA REGULAÇÃO DE INTELIGÊNCIA ARTIFICIAL NA
ADMINISTRAÇÃO PÚBLICA / Raphael Lobato Collet Janny Teixeira. -
2023.
131 f.

Orientador: André Saddy.
Dissertação (mestrado)-Universidade Federal Fluminense,
Faculdade de Direito, Niterói, 2023.

1. Moralidade Administrativa. 2. Ética. 3. Inteligência
Artificial. 4. Administração Pública. 5. Produção
intelectual. I. Saddy, André, orientador. II. Universidade
Federal Fluminense. Faculdade de Direito. III. Título.

CDD - XXX

Bibliotecário responsável: Debora do Nascimento - CRB7/6368

RAPHAEL LOBATO COLLET JANNY TEIXEIRA

DO PRINCÍPIO DA MORALIDADE ADMINISTRATIVA COMO FUNDAMENTO
ÉTICO-JURÍDICO PARA REGULAÇÃO DE INTELIGÊNCIA ARTIFICIAL NA
ADMINISTRAÇÃO PÚBLICA

Dissertação de mestrado apresentada em cumprimento às exigências para a obtenção do título de mestre no Programa de Pós-Graduação *Stricto Sensu* em Direito Constitucional da Universidade Federal Fluminense. Linha de Pesquisa: Instituições políticas, Administração Pública e Jurisdição Constitucional.

Orientador: Prof. Dr. Eur. André Saddy

NITERÓI

2023

RAPHAEL LOBATO COLLET JANNY TEIXEIRA

DO PRINCÍPIO DA MORALIDADE ADMINISTRATIVA COMO FUNDAMENTO
ÉTICO-JURÍDICO PARA REGULAÇÃO DE INTELIGÊNCIA ARTIFICIAL NA
ADMINISTRAÇÃO PÚBLICA

Dissertação de mestrado apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Direito Constitucional da Universidade Federal Fluminense como requisito parcial para a obtenção do título de Mestre em Direito Constitucional.

Data de aprovação: 31 de maio de 2023.

Prof. Dr. André Saddy (Orientador)

Universidade Federal Fluminense

Prof. Dr. Paulo Roberto dos Santos Corval

Universidade Federal Fluminense

Prof. Dr. Daniel Nunes Pêcego

Universidade Estadual do Rio de Janeiro

Prof. Dr. Prof. Dr. Eduardo Jose Guedes Magrani

Berkman Klein Center for Internet & Society da

Universidade de Harvard

AGRADECIMENTOS

De toda a dissertação, talvez a parte mais difícil seja a dos agradecimentos! Difícil porque nunca será suficiente agradecer por todas as valiosas contribuições de tantos amigos e colegas ! Além disso, sempre se incorre no risco de esquecer alguém que, talvez ocultamente, tenha contribuído, seja auxiliando no trâmite interno de documentos, seja com suas orações oferecidas pelo sucesso deste projeto.

De qualquer forma, não podemos deixar de nominar algumas pessoas que contribuíram substancialmente para esta jornada acadêmica!

Agradeço, primeiramente, ao meu orientador, prof. André Saddy, que desde o primeiro contato mostrou-se solícito e interessado em participar deste trabalho. Genuíno acadêmico, professor e formador de professores, não só para docência, mas também para pesquisa e publicação científicas. Levarei suas lições (e correções, rs) pelo resto da minha vida acadêmica e profissional!

Agradeço àqueles que, desde a inscrição no processo seletivo do PPGDC-UFF, apoiaram e incentivaram de forma valiosa, especialmente ao prof. Walter Rodrigues. Também agradeço especialmente aos professores Daniel Pêcego, Eduardo Magrani e Paulo Corval pelas valiosas contribuições feitas nas bancas de qualificação e defesa da dissertação.

Agradeço a todos os meus queridos irmãos, Lia, Rachel, Ruth, Gabriel, Daniel e Esther, que acompanharam, torceram e comemoraram cada etapa vencida deste itinerário acadêmico.

Agradeço a todos os meus colegas de mestrado, especialmente na pessoa de Isabella Macedo, sempre disponível para esclarecer dúvidas e discutir questões acadêmicas contemporâneas. Da mesma forma, agradeço ao colega Jan Mertens pelas ricas discussões na temática de Direito Constitucional. Não poderia deixar de agradecer a todos os colegas do GDAC, grupo de pesquisa em direito administrativo contemporâneo; do Centro Cultural Boa Viagem e da Petrobras, especialmente a Flávio Romano, pelos debates técnicos relacionados à Inteligência Artificial.

Agradeço ao pessoal administrativo do PPGDC-UFF, especialmente à Luciana, secretaria, e ao prof. Pedro Curvello, coordenador, sempre ágeis e solícitos no atendimento das formalidades relativas ao programa de mestrado.

Por fim agradeço à ajuda dos céus, proporcionada especialmente pelo servo de Deus Marcelo Henrique Câmara, a quem espero em breve participar da cerimônia de sua beatificação.

A todos, o meu mais profundo OBRIGADO!

Aos meus queridos pais, Rubens e Margarida, que me deixaram a mais nobre herança que um filho pode desejar: a fé, os irmãos e a arte de viver.

*"Caminante, son tus huellas
el camino y nada más;
Caminante, no hay camino,
se hace camino al andar.
Al andar se hace camino
y al volver la vista atrás
se ve la senda que nunca
se ha de volver a pisar.
Caminante, no hay camino
sino estelas en la mar."*

(poema "Cantares" de Antonio Machado)

RESUMO

A presente dissertação discorre acerca da aplicação do princípio da moralidade administrativa como fundamento jurídico para regulação e regulamentação da adoção de Inteligência Artificial (IA) pela Administração Pública. Foram estudados os aspectos técnicos relacionados à IA, principais campos de aplicação, diagnóstico e perspectivas de sua utilização na Administração. A seguir, apresentou-se um histórico dos problemas éticos que exurgiram em razão da ideologia do progresso técnico, desembocando, hodiernamente, nos riscos relacionados à adoção da IA pelo Estado, com destaque para a ditadura algorítmica em prejuízo da privacidade e autodeterminação informacional; algoritmos opacos de IA (*black box*); substituição da subjetividade do gestor pela “subjetividade” do robô; vieses algorítmicos (*bias*); e impedimento ou dificuldade de acesso a dados públicos. A pesquisa empreendida conclui, como solução, recorrer ao princípio constitucional da moralidade administrativa como fundamento e fonte de força jurídica para a autorregulação pública setorial pelos diferentes órgãos e entidades da Administração Pública, que deverão regulamentar e orientar o uso ético e confiável de sistemas inteligentes. Ademais, propõe diretrizes para auxiliar a Administração Pública na tarefa de gênese normativa, quais sejam: i) Excelência mediante o uso de IA; ii) Centralidade do ser humano; iii) Segurança e Precaução no uso de sistemas inteligentes; iv) Controle e Supervisão por seres humanos; v) Explicabilidade e Revisabilidade das decisões automatizadas; vi) Correção de vieses algorítmicos; vii) Transparência e Disponibilidade de dados públicos; viii) Privacidade e Segurança da Informação; e ix) Responsabilidade pelo adequado funcionamento. No que tange à metodologia adotada, empreendeu-se uma pesquisa bibliográfica, especialmente da produção científica nacional mais recente, bem como da legislação e da jurisprudência brasileira.

Palavras-chave: Moralidade Administrativa. Ética. Inteligência Artificial. Administração Pública

ABSTRACT

This research discusses the application of the principle of administrative morality as a legal basis for regulating ethical issues related to the adoption of Artificial Intelligence (AI) by Public Administration. Technical aspects related to AI, main fields of application, diagnosis, and perspectives for its use in Administration were studied. Next, a history of the ethical problems that arose due to the ideology of technical progress was presented, resulting, in our times, in the risks related to the adoption of AI by the State, with emphasis on the algorithmic dictatorship to the detriment of privacy and informational self-determination; opacity in AI algorithms (black box); replacement of the subjectivity of the manager by the “subjectivity” of the robot; algorithmic biases; and impediment or difficulty of access to public data. The research undertaken concludes, as a solution, to resort to the constitutional principle of administrative morality as the foundation and source of legal force for sectoral public self-regulation by the different organs and entities of the Public Administration, which must regulate and guide the ethical and reliable use of intelligent systems. Furthermore, it proposes guidelines to assist the Public Administration in the task of normative genesis, which are: i) Excellence through the use of AI; ii) Centrality of the human being; iii) Safety and Precaution; iv) Control and Supervision by human beings; v) Explainability and Revisability of automated decisions; vi) Correction of algorithmic biases; vii) Transparency and Availability of public data; viii) Privacy and Information Security; and ix) Responsibility for proper functioning. Regarding the methodology adopted, a bibliographic research was carried out, especially of the most recent national scientific production, as well as Brazilian legislation and jurisprudence.

Keywords: Administrative Morality. Ethic. Artificial Intelligence. Public Administration

LISTA DE ABREVIATURAS E SIGLAS

ANPD	Autoridade Nacional de Proteção de Dados
CEJ/CJF	Centro de Estudos Judiciários do Conselho da Justiça
CGU	Controladoria-Geral da União
CRFB	Constituição da República Federativa do Brasil
EBIA	Estratégia Brasileira para Inteligência Artificial
EIA	Estudo de Impacto Ambiental
GDPR	Regulamento do Direito Europeu sobre Privacidade e Proteção de Dados Pessoais
GSI	Gabinete de Segurança Institucional da Presidência da República
IA	Inteligência Artificial
IoT	Internet das Coisas
LAI	Lei de Acesso à Informação
LGPD	Lei Geral de Proteção de Dados,
ML	Aprendizado de Máquina (<i>Machine Learning</i>)
NLP	Processamento de Linguagem Natural
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
PL	Projeto de Lei
PNSI	Política Nacional de Segurança da Informação
RIMA	Relatório de Impacto Ambiental
RIPD	Relatório de Impacto de Proteção de Dados
RNA	Rede Neural Artificial
RPA	Automação de Processos Robóticos
SAAS	Software como Serviço
STF	Supremos Tribunal Federal
TCU	Tribunal de Contas da União

SUMÁRIO

INTRODUÇÃO	4
CAPÍTULO I - ASPECTOS TÉCNICOS RELACIONADOS À IA E DIAGNÓSTICO DE IMPLANTAÇÃO NA ADMINISTRAÇÃO PÚBLICA	12
1.1 Definição de IA	12
1.2 Modalidades: IA Forte e IA Fraca	14
1.3 Tipos de Aprendizagem de Máquina	17
1.4 Principais domínios de aplicação	18
1.5 Diagnóstico do TCU sobre situação atual e perspectivas da adoção da IA.....	19
CAPÍTULO II - DA PROBLEMÁTICA ÉTICA RELACIONADA À TECNOLOGIA E À IA.....	24
2.1 A ideologia do progresso técnico: da redenção à destruição da humanidade.....	24
2.2 Problemas éticos relacionados à adoção da IA pela Administração Pública.....	32
2.2.1 Ditadura Algorítmica, em prejuízo da privacidade e autodeterminação informacional..	34
2.2.2 Algoritmos opacos de IA	35
2.2.3 Substituição da subjetividade do gestor pela “subjetividade” do robô.....	36
2.2.4 Vieses algorítmicos (<i>bias</i>)	37
2.2.5 Falta de publicidade e impedimento ou dificuldade de acesso a dados públicos	39
2.3 Solução: Ética como principal instrumento norteador da regulação de IA.....	40
CAPÍTULO III – DIREITO, ÉTICA E O PRINCÍPIO DA MORALIDADE ADMINISTRATIVA	43
3.1 A relação entre moral e direito.....	43
3.2 Análise administrativo-constitucional do princípio da moralidade administrativa... 	47
3.2.1 Previsão constitucional do princípio da moralidade administrativa	47
3.2.2 Da construção de normas a partir do exercício de interpretação constitucional por órgãos e entes da Administração Pública.....	53
3.3 Como regular e regulamentar a IA na Administração Pública	54
3.3.1 Modelos regulatórios de IA.....	54
3.3.2 Hetero e autorregulação pública dos aspectos ético-jurídicos da IA.....	60
3.3.3 Diretrizes para a regulamentação da IA: subjetividades incidentes	62
3.3.3.1 Liberdade conformadora ou configuradora para regulamentação da adoção de IA	62
3.3.3.2 Margem de livre apreciação do princípio da moralidade administrativa	65
3.3.4 Análise do PL de IA apresentado pela comissão de juristas	70

CAPÍTULO IV - PROPOSTA DE DIRETRIZES PARA CONSTRUÇÃO DE NORMAS ÉTICO-JURÍDICAS RELATIVAS AO DESIGN E USO DE IA.....	75
4.1 Referências de parâmetros éticos para IA	75
4.1.1 As Três Leis da Robótica de Isaac Asimov	75
4.1.2 Os Princípios sobre IA da OCDE e a Declaração do G20 sobre Princípios para IA Centrada nos Humanos	76
4.1.3 Diretrizes de Ética para IA Confiável da União Europeia.....	78
4.1.4 Referências nacionais de princípios reitores da IA	81
4.2 Proposta de diretrizes para regulação da IA na Administração Pública.....	83
4.2.1 Excelência mediante o uso de IA	84
4.2.2 Centralidade do ser humano	86
4.2.3 Segurança e Precaução no uso de sistemas inteligentes	89
4.2.4 Controle e Supervisão por seres humanos	92
4.2.5 Explicabilidade e Revisibilidade das decisões automatizadas	94
4.2.6 Correção de vieses algorítmicos	95
4.2.7 Transparência e Disponibilidade de dados públicos	97
4.2.8 Privacidade e Segurança da Informação.....	98
4.2.9 Responsabilidade pelo adequado funcionamento.....	101
4.3 Considerações finais	102
CONCLUSÕES.....	104
REFERÊNCIAS	112

INTRODUÇÃO

A efetiva adoção de Inteligência Artificial (IA) pelo poder público e iniciativa privada é relativamente recente. O exponencial aumento do poder computacional alcançado nas últimas décadas, aliado a constante queda de custos (comprovando o acerto da Lei de Moore)¹ e o desenvolvimento de modelos de redes neurais, especialmente a partir de 2011², tornaram técnica e economicamente viável o uso da IA nos mais diversos campos e atividades, gerando a chamada Quarta Revolução Industrial, ou Indústria 4.0.

A Indústria 4.0 é a realidade na qual tanto automação quanto tecnologia da informação exercem um papel fundamental na indústria, englobando as mais recentes inovações tecnológicas, tais como IA, Internet das Coisas (IoT), Contratos Inteligentes (*Smart Contracts*), *Blockchain*, Software como Serviço (SAAS), 5G etc. A pandemia da COVID-19 acelerou ainda mais o processo de transformação digital a nível mundial, com impacto direto em múltiplos setores da economia, nas relações de trabalho e na prestação de serviços públicos.³

Desta forma, a adoção de tecnologias digitais passou a ser um instrumento essencial de sobrevivência e funcionamento eficiente para qualquer organização, pública ou privada. Por outro lado, a necessidade de *digital literacy*, ou competência digital, torna-se ainda mais evidente, visto que muitos trabalhos rotineiros e de pouco valor agregado tendem a ser substituídos por máquinas inteligentes⁴. Segundo pesquisa realizada pela IBM em 2022, 77% das companhias estão explorando ou utilizando IA⁵.

Em relação ao Poder Público, vislumbram-se inúmeros benefícios decorrentes da IA a partir do estímulo ao seu uso excelente, com ganhos em economicidade e eficiência operacional, otimização na alocação de recursos humanos, melhoria da qualidade de decisão baseadas em dados, criação de novos modelos de prestação de serviços e aprimoramento dos já

¹ Segundo a LEI DE MOORE, o número de transistores dos chips teria um aumento de 100%, pelo mesmo custo, a cada dois anos. Esse padrão continuou a se manter em grande parte da indústria. Cf. CORNELIS, Disco; BAREND, Van der Meulen. *Getting New Technologies Together: Studies in Making Sociotechnical Order*. New York: Walter de Gruyter, 1998, p. 206–207.

² OCDE. Artificial Intelligence in Society. Paris: [OECD Publishing](https://doi.org/10.1787/eedfee77-en), 2019. DOI: <https://doi.org/10.1787/eedfee77-en>. Disponível em: https://www.oecd-ilibrary.org/science-and-technology/artificial-intelligence-in-society_eedfee77-en. Acesso em: 6 mar. 2023.

³ LAMB, Luís C. O Futuro do Trabalho Pós-Pandemia de COVID-19: Reflexões sobre os Impactos da Inteligência Artificial, Ciência e Educação. *Métricas*, São Paulo, 2020. Disponível em: <https://metricas.usp.br/wp-content/uploads/2020/07/MétricasUSP-revisado14Jul2020.pdf>. Acesso em: 27 dez. 2020.

⁴ Ibid.

⁵ GLOBAL AI Adoption Index 2022: New research commissioned by IBM in partnership with Morning Consult. *IBM*, New York, 2022. Disponível em: <https://www.ibm.com/watson/resources/ai-adoption>. Acesso em: 30 out. 2022.

existentes e como instrumento de políticas públicas voltadas a promover o bem-estar social, o desenvolvimento sustentável, o crescimento econômico e cultural e a redução das desigualdades sociais. Segundo estudo realizado pela Deloitte University Press a implementação de automação e IA no governo dos EUA têm potencial para economizar entre US\$ 3,3 bilhões e US\$ 41,1 bilhões⁶. Sistemas inteligentes também pode ser um importante aliado na transparência e controle social do Estado, de forma a aumentar a conformidade das atividades públicas no que tange à moralidade, legalidade, combate à corrupção e malversação de recursos públicos⁷.

No caso do Brasil, a OCDE ressaltou importantes avanços obtidos a partir da introdução de IA na economia digital brasileira, no combate à corrupção, na medicina e na prestação de serviços à população⁸. Por exemplo, já existem diversas aplicações de inteligência computacional para identificação de indícios de fraudes em licitações, fiscalização de contratos e prestação de serviços digitais para a população⁹. O governo federal, por sua vez, instituiu uma política pública voltada à implementação de IA, a Estratégia Brasileira de Inteligência Artificial (EBIA), visando identificar áreas prioritárias no desenvolvimento e uso desta tecnologia para obter ganhos na promoção da competitividade, no aumento da produtividade brasileira, na prestação de serviços públicos e na melhoria da qualidade de vida das pessoas¹⁰.

Em pesquisa sobre o uso das Tecnologias de Informação e Comunicação no setor público brasileiro (TIC Governo Eletrônico 2021), levada a efeito pelo Centro Regional de

⁶ VIECHNICK, Peter; D. EGGERS, William. How much time and money can AI save government? Cognitive technologies could free up hundreds of millions of public sector worker hours. [s. l.]. Deloitte University Press, 2017. Disponível em: https://www2.deloitte.com/content/dam/insights/us/articles/3834_How-much-time-and-money-can-AI-save-government/DUP_How-much-time-and-money-can-AI-save-government.pdf. Acesso em: 30 out. 2022.

⁷ RACCA, Gabriela M.; PERIN, Roberto Cavallo. Corrupção como quebra da confiança na administração pública e violação aos direitos fundamentais. RCJ – Revista Culturas Jurídicas, Niterói, v. 2, n. 3, p. 100-123, set., 2015.

⁸ OCDE. A Caminho da Era Digital no Brasil. Paris: OCDE, 2020. Disponível em: <http://www.oecd.org/publications/a-caminho-da-era-digital-no-brasil-45a84b29-pt.htm>. Acesso em: 27 dez. 2020.

⁹ Na esfera federal: TCU - "Alice" (Análise de Licitações e Editais); TCU - "Sofia" (Sistema de Orientação sobre Fatos e Indícios para o Auditor); TCU - "Monica" (Monitoramento Integrado para Controle de Aquisições). A CGU possui outro sistema baseado em IA usado com o propósito de fiscalizar contratos e fornecedores. A ferramenta elabora uma análise de riscos, incluindo não somente o de corrupção, mas também de outros problemas, como a possibilidade de um fornecedor não cumprir o contrato ou fechar as portas. Paraná Inteligência Artificial (PIÁ): é um programa de IA focado na prestação de serviços à população. A plataforma e o aplicativo reúnem mais de 380 serviços do Governo em um só lugar e funcionam como canais de diálogo com o cidadão para atender suas demandas e reclamações. O PIÁ também tem integração bidirecional com ferramentas do governo federal e integração municipal. Cf. OCDE. A Caminho da Era Digital no Brasil. Paris: OCDE publishing, 2020. Disponível em: <http://www.oecd.org/publications/a-caminho-da-era-digital-no-brasil-45a84b29-pt.htm>. Acesso em: 27 dez. 2020.

¹⁰ PODER EXECUTIVO FEDERAL (Brasil). Ministério da Ciência, Tecnologia e Inovações/Gabinete do Ministro. PORTARIA GM Nº 4.617, DE 6 DE ABRIL DE 2021. Institui a Estratégia Brasileira de Inteligência Artificial e seus eixos temáticos. [s. l.], 12 abr. 2021. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-gm-n-4.617-de-6-de-abril-de-2021-*-313212172. Acesso em: 22 nov. 2021.

Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br do NIC.br), sobre a utilização de tecnologias emergentes, ressaltou-se que a adoção de IA é a que mais cresceu: quase metade (45%) dos órgãos federais e 22% dos órgãos estaduais utilizaram tecnologias de IA nos 12 meses anteriores à pesquisa. A adoção de sistemas inteligentes ocorreu em maiores proporções em órgãos do Judiciário (55%), do Ministério Público (50%) e do Legislativo (48%). Já em relação aos órgãos do executivo federal e estadual, apenas 20% estão utilizando IA¹¹.

Em que pese os incontáveis benefícios e avanços proporcionados pela inteligência sintética, fatos e documentários recentes despertaram o grande público para problemas éticos relacionados à IA. O escândalo da Cambridge Analytica, que utilizou, sem consentimento dos titulares, dados de cerca de 87 milhões de perfis do Facebook para direcionar propaganda política¹², e os documentários lançados pela plataforma Netflix “*O Dilema das Redes*”¹³ e “*Coded Bias*”¹⁴, evidenciaram que entes públicos e privados estão utilizando algoritmos de IA para previsão, controle e exploração de hipossuficiências emocionais e intelecto-volitivas com o intuito de manipular ou induzir o comportamento humano como produto.

Especificamente em relação ao Estado, o poder de obter amplo acesso a dados pessoais, muitas vezes coletados compulsoriamente (por exemplo, mediante câmeras de segurança pública, fiscalização tributária, execução de políticas públicas, censo ou recenseamento demográfico etc.) aliado ao uso de IA, podem representar riscos e ameaças ao Estado Democrático de Direito, bem como a direitos e garantias fundamentais. A temática não é nova, como bem demonstra a obra “1984”¹⁵, que expõe a problemática do controle populacional em tempo real por uma entidade estatal. A IA pode ser utilizada como um poderoso instrumento para elaboração de dossiês contra cidadãos, opositores políticos e atividades de vigilância totalitária, bem como para obtenção manipulada do consenso de uma população “capturada” por computadores inteligentes.

¹¹ NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR. Pesquisa sobre o uso das tecnologias de informação e comunicação no setor público brasileiro: TIC Governo Eletrônico 2022. São Paulo, SP: Comitê Gestor da Internet no Brasil, 2022. Disponível em: https://cetic.br/media/docs/publicacoes/2/20220725170710/tic_governo_eletronico_2021_livro_eletronico.pdf. Acesso em: 17 set. 2022.

¹² BRUNO, Fernanda. Economia Psíquica dos Algoritmos. NEXO, 2018. Disponível em: <https://www.nexojornal.com.br/ensaio/2018/A-economia-ps%C3%ADquica-dos-algoritmos-quando-o-laborat%C3%B3rio-%C3%A9-o-mundo>. Acesso em: 10 nov. 2020.

¹³ ORLOWSKI, Jeff. O Dilema das Redes. Netflix, 2020. Disponível em: <https://www.netflix.com/title/81254224>. Acesso em: 7 nov. 2020.

¹⁴ KANTAYYA, Shalini. CODED Bias. Netflix, 2020. Disponível em: <https://www.netflix.com/title/81328723>. Acesso em: 03 mar. 2023.

¹⁵ ORWELL, George. 1984. 1. ed. Companhia das Letras, 2009.

Além disso, existem algoritmos opacos que funcionam como um sistema fechado (conhecidos como *black box*¹⁶), onde não é possível obter uma explicação sobre como se chegou a um resultado, recomendação ou decisão específica. Questiona-se ainda, dentre outros problemas que serão abordados oportunamente, a ocorrência de vieses algoritmos e injustiças (*bias*); e casos relacionados ao uso indevido de dados pessoais em prejuízo ao direito à privacidade e autodeterminação informativa dos indivíduos.

Diante disso, a presente dissertação tem como problema de pesquisa a ausência de parâmetros éticos ou morais para o design e uso de IA pela Administração Pública, que assegurem a sua confiabilidade e juridicidade. Vale ressaltar que o Brasil não possui uma lei específica que suporte ações planejadas, defina princípios ou regulamente a adoção de IA. Diversos projetos de lei encontram-se em discussão no Congresso Nacional, tendo sido nomeada uma comissão de juristas com a missão de propor um marco regulatório para a Inteligência Artificial.¹⁷ A ausência de Lei, aliada aos riscos relacionados a sistemas inteligentes com potencial de causar grandes violações e danos a direitos fundamentais, e, inclusive, ameaçar o próprio Estado Democrático de Direito, tornam ainda mais premente a necessidade de se buscar uma solução.

Como hipótese de solução à problemática ora apresentada, este trabalho propõe-se, como objetivo primário, realizar um estudo sobre a aplicação do princípio da moralidade administrativa, previsto no art. 37 da CRFB, enquanto desdobramento da ética aplicada ao Poder Público, para assegurar que a tecnologia seja voltada ao atendimento do bem comum; e ao respeito de direitos e garantias fundamentais, desde a sua concepção, promovendo uma IA confiável, centrada no ser humano e sensível a valores constitucionais¹⁸. Para atingir esse objetivo, será avaliada a possibilidade de adoção do princípio da moralidade administrativa como fundamento e fonte de força jurídica para definição de diretrizes e construção de normas e orientações com o intuito de regular e nortear os órgãos e entidades da Administração Pública

¹⁶ É preciso ressaltar, sem embargo, que a opacidade algorítmica da IA não pode ser categorizada como uma falha de sistema em si, visto que existem limitações inerentes à complexidade de alguns modelos de IA, como *deep learning*. Não se olvida, por outro lado, que possa ser considerada uma questão desafiadora, especialmente em áreas em que a transparência e a responsabilidade são importantes, como cuidados de saúde, justiça criminal e crédito. Alguns modelos podem ser mais explicáveis do que outros, mas sempre haverá um equilíbrio entre a transparência e o desempenho preditivo. A busca por soluções eficazes para lidar com a opacidade algorítmica é um campo ativo de pesquisa e desenvolvimento na área de IA. (Cf. MOLNAR, Christoph. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. Victoria: Leanpub, 2022).

¹⁷ AGÊNCIA SENADO. Brasil poderá ter marco regulatório para a inteligência artificial. Senado, 30 out. 2022. Disponível em: <https://www12.senado.leg.br/noticias/materias/2022/03/30/brasil-podera-ter-marco-regulatorio-para-a-inteligencia-artificial>. Acesso em: 30 out. 2022.

¹⁸ MAGRANI, Eduardo. Entre dados e robôs. 2. ed. Porto Alegre: Arquipélago Editorial, 2019.

em todo ciclo de utilização da IA, desde a sua concepção e desenvolvimento, passando pelas fases de treinamento até o uso cotidiano.

Uma vez verificada a possibilidade de aplicação do princípio da moralidade administrativa, serão abordadas as subjetividades públicas incidentes na atividade de elaboração das normas regulamentadoras do uso ético da IA. Por fim, como objetivo secundário, serão propostas algumas diretrizes, com fundamento no princípio da moralidade, que deverão nortear a atividade de normatização pelos órgãos e entes estatais, de sorte a resolver os problemas ético-jurídicos que foram identificados e evitar outros que possam vir a ocorrer no futuro.

Em pesquisa prévia realizada sobre o tema, pode-se constatar que são raras as publicações especificamente voltadas para ao presente objeto de pesquisa, especialmente quando relacionamos as palavras e expressões “IA”, “inteligência artificial”, “administração pública”, “moralidade administrativa” e “ética”¹⁹.

Ressalta-se que não se pretende, neste trabalho, adentrar em aspectos sociológicos, em relação, por exemplo, a estudo de casos e situações fáticas decorrentes da aplicação, ausência ou obstáculos à adoção de IA, mesmo porque o uso desta tecnologia na Administração Pública é bastante incipiente. Também não será objeto do presente estudo a normatização de IA no Poder Legislativo ou do Judiciário²⁰. Por outro lado, a presente pesquisa não se circunscreve a aspectos puramente teóricos e abstratos, para desenvolver “mega hipóteses”, de forma completamente desvinculada do mundo real²¹.

Pretende-se, assim, estudar e compreender não somente a produção jurídica, levando em consideração a transdisciplinaridade que envolve o tema (ética, direito, direitos constitucional, administrativo e digital), mas também realizar um esforço de entendimento dos aspectos técnicos e operacionais relacionados a sistemas computacionais inteligentes, modalidades de Inteligência Artificial (IA Forte e IA Fraca) e principais domínios de aplicação,

¹⁹ De forma a identificar conteúdos especificamente relacionados ao objeto de estudo, procuramos utilizar palavras e expressões como: “ética e inteligência artificial e administração pública”; “ética e inteligência artificial”; “inteligência artificial e administração pública”; inteligência artificial ética e moral”; “inteligência artificial e moralidade administrativa”; “inteligência artificial e constituição”. Constatamos, assim, que a maioria das publicações especializadas sobre o tema discorrem sobre “ética” e “inteligência artificial”, havendo pouco material produzido quando se acrescenta a expressão “administração pública”; “estado” ou “moralidade administrativa”.

²⁰ Em que pese o Conselho Nacional de Justiça ter publicado a Resolução CNJ n.º 332/2020 que dispõe sobre a ética, a transparência e a governança na produção e no uso de IA no Poder Judiciário.

²¹ FRAGALE FILHO, Roberto. Quando a empiria é necessária? In: CONGRESSO NACIONAL DO CONPEDI - CONSELHO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM DIREITO, XIV, 2005, Fortaleza. Anais [...]. Fortaleza: CONPEDI, 2005. Disponível em: <http://www.publicadireito.com.br/conpedi/manaus/arquivos/anais/XIVCongresso/135.pdf>. Acesso em: 6 out. 2021.

bem como compreender o estágio atual de utilização na Administração Pública, de forma a realizar um esforço de extração do direito a partir da compreensão da realidade. É por esta razão que o primeiro capítulo da dissertação abordará aspectos técnicos relacionados à IA, bem como o diagnóstico da sua adoção na Administração Pública e perspectivas para uso futuro.

Em seguida, no segundo capítulo, será abordada a problemática ética relacionada à tecnologia, partindo do contexto histórico da ideologia do progresso técnico surgida no início no séc. XVI e a sua repercussão nos dias atuais, apontando, ao final do capítulo, as principais questões morais relacionadas à adoção da IA pela Administração Pública que foram identificadas na pesquisa: a ditadura algorítmica em prejuízo da privacidade e autodeterminação informacional; algoritmos opacos de Inteligência Artificial (*black box*); substituição da subjetividade do gestor pela “subjetividade” do robô; vieses algorítmicos (*bias*); e impedimento ou dificuldade de acesso a dados públicos.

No terceiro capítulo, é apresentada uma análise da relação entre direito, ética e o princípio da moralidade administrativa, capítulo em que se aprofundará nos aspectos administrativo-constitucionais do referido princípio, bem como na viabilidade de construção de normas a partir do exercício de interpretação constitucional por órgãos e entes da Administração Pública. Após, ainda no terceiro capítulo, tratar-se-á do “como” regulamentar a IA na Administração Pública. Serão, assim, estudados os diferentes modelos regulatórios de IA culminando com a defesa à autorregulação pública setorial interna a ser levada a efeito pelos diferentes órgãos e entidades da Administração Pública. Em seguida, serão tratadas as subjetividades públicas incidentes na tarefa de gênese normativa, a saber: a liberdade conformadora ou configuradora do ato normativo e a margem de livre apreciação do princípio da moralidade administrativa enquanto conceito jurídico indeterminado.

Por fim, no quarto capítulo, propor-se-á diretrizes específicas para auxiliar a Administração Pública na construção dessas normas. Com fundamento no princípio da moralidade administrativa e partindo-se de referências de princípios e parâmetros éticos encontrados na literatura (Leis de Asimov), organismos internacionais, direito estrangeiro e nacional, a presente dissertação propõe nove diretrizes: i) Excelência mediante o uso de IA; ii) Centralidade do ser humano; iii) precaução; iv) Controle e Supervisão por seres humanos; v) Explicabilidade e Revisibilidade das decisões automatizadas; vi) Correção de vieses algorítmicos, vii) Transparência e Disponibilidade de dados públicos; viii) Privacidade e Segurança no tratamento de dados; e ix) Responsabilidade pelo adequado funcionamento.

Com relação ao recorte metodológico, é preciso lembrar que se trata de tema dotado de complexidade e interdisciplinaridade. Além disso, a presente dissertação se propõe a trilhar um

longo itinerário, enfrentando as questões relativas à a) viabilidade da adoção do princípio da moralidade como fundamento para regulamentação de IA na Administração Pública; e b) identificação das subjetividades públicas incidentes na atividade de elaboração das normas regulamentadoras por órgãos e entidades; e c) proposta de diretrizes para construção de normas ético-jurídicas relativas ao design e uso de IA.

Desta forma, e tendo em vista que o tempo do mestrado não torna possível abordagens exaurientes sobre conceitos e teorias, optou-se por adotar algumas definições e teorias sem adentrar em maiores discussões doutrinárias, tais como aspectos técnicos e funcionais de IA; relação entre moral e direito; e análise documental detalhada de fontes de relatórios utilizados para o estudo. No entanto, destaca-se que não está descartada a possibilidade de desenvolver esses aspectos em um eventual doutorado, em caso de continuidade na pesquisa.

No que tange à metodologia adotada, empreendeu-se uma pesquisa bibliográfica, recorrendo às fontes doutrinária, legal e jurisprudencial, procurando-se identificar e priorizar publicações especializadas sobre o tema, especialmente artigos científicos, periódicos, livros, dissertações e teses, legislação e jurisprudência. O mapeamento da produção científica foi feito principalmente em plataformas digitais, bases de dados, indexadores de conteúdos em fontes de reconhecida reputação em relação à qualidade dos trabalhos indexados²². Foram definidos alguns parâmetros para a pesquisa, de forma a identificar os autores de referência do dia, definindo-se critérios delimitadores para determinar onde termina e começa a literatura a ser pesquisada, tais como espaço, tempo, conteúdo, bases de dados, indexadores de conteúdos, fontes de pesquisa, palavras e expressões de busca.

Sobre a delimitação no espaço e tempo, tendo em vista que a presente pesquisa é especialmente voltada ao ordenamento jurídico brasileiro, onde se pretende propor diretrizes que sejam úteis à Administração Pública no desenvolvimento e utilização da IA, procurou-se identificar os trabalhos acadêmicos mais relevantes da produção jurídica nacional, bem como a legislação e jurisprudência brasileira. Sem embargo, não se deixou de realizar uma busca na bibliografia estrangeira, especialmente levando em consideração que se trata de um tema em

²² Foi dada preferência a repositórios científicos institucionais, especialmente às seguintes plataformas:

RVBI: <https://www12.senado.leg.br/institucional/biblioteca/rvbi>

Scielo: <https://www.scielo.org/>

Base de dados/Portal da CAPES: <https://www.periodicos.capes.gov.br>

Biblioteca digital dissertações e teses <https://bdtd.ibict.br/vufind/>

Google Acadêmico <https://scholar.google.com.br/?hl=pt>

Também se realizou buscas em sites de instituições, universidades e especialistas que possuem notoriedade pública em relação ao tema objeto da dissertação, tais como:

ITS - Instituto de Tecnologia e Sociedade

PUC-RJ

Opice Blum Academy

franca ebulição também fora do Brasil, havendo influência direta do direito comparado nos doutrinadores, juízes, e mesmo legisladores brasileiros. Só para se dar um exemplo, o Regulamento do Direito Europeu sobre Privacidade e Proteção de Dados Pessoais (GDPR)²³, que possui normas sobre a utilização de IA, foi uma importante fonte de inspiração do legislador brasileiro na confecção da Lei n.º 13.709/2018 (Lei Geral de Proteção de Dados ou LGPD), sendo inafastável a necessidade de conhecer a produção jurídica europeia relativa a este tema. A Organização para a Cooperação e Desenvolvimento Econômico (OCDE), por seu turno, publicou diretrizes e orientações sobre IA e que estão sendo levadas em consideração por boa parte dos países, inclusive o Brasil, conforme veremos mais adiante.

Com relação à delimitação temporal das publicações, adotou-se como linha de corte os artigos mais recentes, publicados nos últimos três a cinco anos. No Brasil, identificou-se uma produção mais intensa a partir de 2021, talvez impulsionada pela entrada em vigor da LGPD e pelo início das discussões no Congresso Nacional de um Projeto de Lei sobre IA.

Sobre a delimitação de conteúdo, haja vista a especificidade e novidade do tema, são raras as publicações em fontes de conhecimento do direito administrativo. Além disso, a jurisprudência sobre o tema é bastante escassa, tendo havido algumas decisões referentes a sistema de IA relacionadas à privacidade de dados e à LGPD. Evitou-se utilizar conteúdo aberto e editável, como Wikipédia, tendo em vista que, em geral, apresentam erros e imprecisões, além de permitirem que o texto seja alterado a qualquer momento, sem uma identificação precisa da autoria, o que prejudica a integridade e autenticidade da informação. Por outro lado, identificou-se, em alguns casos, referências a artigos, livros e publicações e, com isso, recorreu-se às fontes originais. Também se evitaram conteúdos decorrentes de fontes mais genéricas e sem um aprofundamento científico, como manuais, cursos, ou artigos de opinião. Por fim, e mais a título experimental, foi utilizada uma única pesquisa de conteúdo em um sistema de IA que vem ganhando notoriedade pública, o “ChatGPT”, onde se indagou sobre quais seriam as atividades em que a Inteligência Artificial é comprovadamente superior à inteligência humana, conforme será visto oportunamente.

²³ UNIÃO EUROPÉIA. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho. Relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679&qid=1677953718324&from=PT>. Acesso em: 04 mar. 2022.

CAPÍTULO I - ASPECTOS TÉCNICOS RELACIONADOS À IA E DIAGNÓSTICO DE IMPLANTAÇÃO NA ADMINISTRAÇÃO PÚBLICA

Ainda que a presente dissertação não tenha como propósito estudar os aspectos técnicos de engenharia, matemática e programação relacionados a sistemas inteligentes, é importante que haja uma noção sobre a natureza técnica para que se tenha uma melhor compreensão das possíveis repercussões jurídicas relacionadas à IA.

Os juristas costumam negligenciar a análise cognitiva do fato ao focar seu estudo na produção legislativa, doutrinária e jurisprudencial relativa ao objeto de pesquisa. Mas, para a correta apreciação do fenômeno jurídico, é imprescindível compreender o *modus operandi*, os aspectos técnicos, funcionais e operacionais da IA. A partir da inteligibilidade da sua natureza e funcionamento, pretende-se compreender as questões éticas dela decorrentes, surgindo, a partir daí, a necessidade de regulação e das “tecnologias jurídicas”.

Desta forma, é necessário dialogar com o mundo fático, estudar e compreender não somente a produção jurídica, mas também realizar um esforço de compreensão da natureza e funcionamento do aprendizado de máquina para extrair racionalmente as diretrizes e princípios jurídicos que deverão reger a IA.

1.1 Definição de IA

O desenvolvimento mais recente de pesquisas sobre IA ocorreu logo após a Segunda Guerra Mundial, tendo como personagem central Alan Mathison Turing, matemático, cientista da computação e criptoanalista, considerado o pai da ciência da computação teórica e da IA²⁴. Turing ficara famoso por suas contribuições na Segunda Guerra Mundial, quando logrou êxito em quebrar a criptografia do sistema Enigma²⁵, permitindo que os aliados visualizassem o conteúdo de mensagens cifradas do exército alemão relacionadas ao deslocamento de tropas. O acesso a essas informações foi considerado crucial para a vitória dos aliados, sendo estimado que reduziu o período da Guerra em dois ou três anos, poupando a vida de 14 a 21 milhões de pessoas.²⁶

²⁴ COOPER, S. Barry; VAN LEEUWEN, Jan. Alan Turing: His Work and Impact. Waltham: Elsevier, 2013.p. 481–485.

²⁵ Sistema alemão de encriptação de dados utilizado na Segunda Guerra Mundial.

²⁶ COPELAND, Jack. Alan Turing: The codebreaker who saved ‘millions of lives’. BBC News, Christchurch, Nova Zelândia, 19 jun. 2012. Disponível em: <https://www.bbc.com/news/technology-18419691>. Acesso em: 30 out. 2022.

Ao tratar da problemática relativa à capacidade de computadores pensarem, Turing²⁷ propôs um experimento para verificar se um robô pode ou não ser chamado de inteligente, experimento esse conhecido como o “*Teste de Turing*” ou o “*Jogo da Imitação*”. O teste consiste em uma pessoa interrogar um computador e um ser humano ao mesmo tempo, sem que se saiba quem é quem. Se, uma vez obtidas as respostas de cada um deles, o inquiridor não souber identificar quais das respostas foram emitidas pela máquina, esta passou no teste. Um exemplo de Teste de Turing amplamente utilizado na internet é o CAPTCHA (*Completely Automated Public Turing test to tell Computers and Humans Apart*), destinado a identificar se o usuário é humano ou computador.

A expressão “Inteligência Artificial” foi cunhada por John McCarthy²⁸, que a utilizou pela primeira vez em 1956 em uma conferência realizada no Dartmouth College, em New Hampshire, nos Estados Unidos da América. Significa, de forma geral, qualquer sistema capaz de executar tarefas próprias de seres com capacidade racional. IA englobaria, assim, um conjunto de técnicas voltadas a emular ou replicar aspectos próprios de seres inteligentes, tais como a capacidade de cognição, raciocínio e volição, ligados ao que Aristóteles²⁹ define como sendo propriedades da “alma intelectual”. Está atualmente associada às atividades computacionais relacionadas a inúmeras aplicações, tais como robótica, reconhecimento facial, aprendizado de máquina (*machine learning*), aprendizado profundo (*deep learning*), decisões automatizadas etc.

Conforme ressaltam Stuart Russell e Peter Norvig³⁰, não há consenso entre cientistas e pesquisadores sobre a melhor definição de IA, variando em função de diferentes perspectivas, tais como processos de pensamento e raciocínio, comportamento, desempenho humano: “*The study of how to make computers do things at which, at the moment, people are better*”³¹; “*The automation of activities that we associate with human thinking, activities such as decision-*

²⁷ TURING, Alan. Computing Machinery and Intelligence. *Mind*, v. 59, n. 236, p. 433-460, out. 1950. Disponível em: <https://archive.nytimes.com/www.nytimes.com/library/cyber/surf/1106surf-turing.html>. Acesso em: 30 out. 2022.

²⁸ MCCARTHY, John et al. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Magazine*, v. 27, n. 4, p. 12-14, ago. 2006.

²⁹ ARISTÓTELES. *Da Alma (De Anima)*. Tradutor Edson Bini. São Paulo: Edipro, 2011.

³⁰ RUSSELL, Stuart; NORVIG, Peter. *Artificial Intelligence: A Modern Approach*. New Jersey: Prentice Hall, 1995, p. 4

³¹ RICH, Elaine; KNIGHT, Kevin. *Artificial Intelligence Second Edition*. Nova York: McGraw-Hill, 1991., p.3

making, problem solving, learning ..."³²; e *“The study of the computations that make it possible to perceive, reason, and act”*³³.

A OCDE, por sua vez, apresentou uma definição que vem sendo aceita como uma referência³⁴:

[...] um sistema de IA é um sistema baseado em máquina que pode, para um determinado conjunto de objetivos definidos pelo homem, fazer previsões, recomendações ou tomar decisões que influenciam ambientes reais ou virtuais. Os sistemas de IA são projetados para operar com vários níveis de autonomia". Ainda conforme a OCDE, um sistema de IA consiste em três elementos principais: sensores, lógica operacional e atuadores. Os sensores coletam dados brutos do ambiente, processados pela lógica operacional para fornecer saídas para os atuadores, que por sua vez agem para alterar o estado do ambiente. Este ciclo é repetido inúmeras vezes, e como o ambiente é alterado pelo sistema de IA, a cada ciclo a lógica operacional pode ser aperfeiçoada.

De forma bastante simples, portanto, a IA pode ser entendida como a tecnologia capacitada de executar tarefas que pressupõem o uso de inteligência como, por exemplo, um sistema computacional dotado de autoaprendizagem e capaz de propor respostas de forma autônoma, sem uma predefinição específica dos seus programadores, visando atingir um objetivo previamente definido. Age não apenas automaticamente, mas de forma por vezes imprevisível em relação à sua programação inicial, sendo capazes de aprender por experiência, podendo alterar as instruções de programação iniciais e até criar novas, devido a técnicas de *machine learning*.

1.2 Modalidades: IA Forte e IA Fraca

Há basicamente duas modalidades de IA: a IA Forte ou Geral (*Strong AI* ou *General Artificial Intelligence*) e a IA Fraca (*Weak AI* ou *Narrow AI*).

O termo IA Forte³⁵ remete à ideia de uma superinteligência computacional unificada, capaz de executar tarefas de natureza geral e sem uma destinação específica predefinida. As potencialidades relacionadas a IA Forte despertam preocupações e alimentam o imaginário

³² BELLMAN, R. E. *An Introduction to Artificial Intelligence: Can Computers Think*, New Jersey: Boyd & Fraser Publishing Company, 1978 *apud* RUSSELL, Stuart; NORVIG, Peter. *Artificial Intelligence: A Modern Approach*. New Jersey: Prentice Hall, 1995, 45

³³ WINSTON, P. H. *Artificial Intelligence*. Massachusetts: Addison-Wesley, 1992 *apud* RUSSELL, Stuart; NORVIG, Peter. *Artificial Intelligence: A Modern Approach*. New Jersey: Prentice Hall, 1995., p. 5

³⁴OCDE. Artificial Intelligence in Society. *OECD Publishing*, Paris, 2019. Disponível em: <https://doi.org/10.1787/eedfee77-en>. Acesso em: 6 mar. 2023.

³⁵ RUSSELL, Stuart; NORVIG, Peter. *Artificial Intelligence: A Modern Approach*. 3. ed. New Jersey: Prentice Hall, 2010., p. 1026

popular em relação ao risco de que a humanidade seja subjugada ou até exterminada por máquinas inteligentes. É o que geralmente se nota em alguns filmes de ficção científica que apresentam uma distopia futurista³⁶, desconfiança e uma visão de futuro sombrio em relação à IA.

Também surgem questões filosóficas relativas à natureza racional e consciência de máquinas inteligentes. Recentemente, o Google demitiu o engenheiro de software Blake Lemoine, depois que ele afirmou que a IA Google LaMDA (*Language Model for Dialogue Applications*) da empresa tem alma e sentimentos. Em nota³⁷, a companhia afirmou que as alegações de Lemoine “são totalmente infundadas” e que trabalhou por meses com ele “para esclarecer” a situação.

Ainda que o presente trabalho não pretenda aprofundar em questões filosóficas, há explicações consistentes a refutar a atribuição de estados mentais genuínos a uma IA, tal como demonstraram John Searle³⁸, com o argumento do quarto chinês. Searle argumenta que o fato de uma IA passar no teste de Turing não significa, necessariamente, que realize um ato de cognição relevante. Ele exemplifica imaginando que está em um quarto fechado com peças contendo símbolos chineses e um livro em português onde está escrito um programa de computador para falar chinês. Neste quarto há uma abertura para inputs e outputs. Do lado de fora alguém introduz perguntas feitas em chinês. Ao recebê-las, ele consulta o livro, emitindo respostas de acordo com as instruções e utilizando os símbolos que estão no quarto. Para as pessoas de fora do quarto, não é possível distinguir se quem está dentro é um falante nativo de chinês ou não. Mas, na verdade, ele não entende nenhuma palavra de chinês, pois simplesmente executou o programa na linguagem que conhecia (português). Desta forma, substituindo-o por um programa, percebe-se que a mera execução da tarefa não significa que o computador compreende chinês.

Além disso, Alan Turing e Kurt Godel, demonstram que é impossível provar se certas afirmações matemáticas são verdadeiras ou falsas e que alguns problemas computacionais não podem ser resolvidos com algoritmos. Turing e Godel exemplificam dizendo que sempre que um sistema matemático é rico o suficiente para descrever a aritmética que aprendemos na

³⁶ Vide, por exemplo, o filme *Blade Runner* (1982), e outros, como *Exterminador do Futuro* (1984) e *Matrix* (1999).

³⁷ ARRUDA, Wellington. Engenheiro do Google que afirmou que IA tem vida é demitido. *Tecmundo*, 23 jul. 2022. Disponível em: <https://www.tecmundo.com.br/mercado/242193-engenheiro-google-afirmou-que-ia-tem-vida-demitido.htm>. Acesso em: 30 out. 2022.

³⁸ SEARLE, John R. Minds, brains, and programs. *The Behavioral and Brain Sciences*, Estados Unidos, v. 3, n. 3, p. 417-457, 1980. Disponível em: <https://home.csulb.edu/~cwallis/382/readings/482/searle.minds.brains.programs.bbs.1980.pdf>. Acesso em: 30 out. 2022.

escola, ele não pode provar sua própria consistência. Décadas depois, o matemático Steve Smale propôs uma lista de dezoito problemas matemáticos não resolvidos para o século XXI, sendo o 18º relacionado aos limites da inteligência, tanto para os humanos quanto para as máquinas³⁹.

Pesquisadores da Universidade de Cambridge e da Universidade de Oslo afirmam que a instabilidade é o calcanhar de Aquiles da IA moderna. Por causa do paradoxo acima apontado, uma rede neural inerentemente confiável não pode ser construída, haja vista a dificuldade de se computar redes neurais estáveis e precisas. A IA seria confiável apenas em áreas específicas e usando métodos específicos. Além disso, não se trata de um problema de quantidade ou qualidade dos dados: "Isso é semelhante ao argumento de Turing: Existem problemas computacionais que não podem ser resolvidos independentemente do poder de computação e do tempo de execução", disse Anders Hansen⁴⁰, um dos autores do artigo.

Essas conclusões implicam no reconhecimento de limites às pretensões de simulação completa da inteligência humana, tornando inviável a existência de uma IA Forte, pelo menos, segundo o atual estágio do conhecimento científico. Como consequência, as soluções de IA Fraca são as comumente utilizadas nos dias atuais.

A *Weak AI*⁴¹ é a uma “inteligência específica”, desenvolvida para executar uma tarefa ou atividade determinada. É a IA Fraca que está cada vez mais presente no nosso cotidiano, como algoritmos de recomendação de redes sociais, assistentes virtuais (Google Assistente, Alexa e Siri), *chatbots* etc. O fato de ser uma “*Narrow IA*” não significa que não possa executar tarefas complexas, inclusive de forma mais eficiente que seres humanos. O sistema *Deep Blue*, por exemplo, um supercomputador da IBM especialmente destinado para jogar xadrez, em 1997, venceu Garry Kasparov, tornando-se o primeiro computador a vencer um campeão mundial de xadrez⁴². Também está sendo testada em campos mais desafiadores, tais como carros e veículos autônomos.

³⁹ COLBROOK, Matthew J.; ANTUN, Vegard; HANSEN, Anders C. The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem. Proceedings of the National Academy of Sciences, Estados Unidos, v. 119, n. 12, 2022. DOI: <https://doi.org/10.1073/pnas.2107151119>. Disponível em: <https://www.pnas.org/doi/10.1073/pnas.2107151119>. Acesso em: 30 out. 2022.

⁴⁰ INOVAÇÃO TECNOLÓGICA. Paradoxo matemático expõe limites da Inteligência Artificial. [S.l.], 28 mar. 2022. Disponível em: <https://www.inovacaotecnologica.com.br/noticias/noticia.php?artigo=paradoxo-matematico-expoe-limites-inteligencia-artificial&id=010150220328#.Y18AMXbMKUI>. Acesso em: 30 out. 2022.

⁴¹ RUSSELL, Stuart; NORVIG, Peter. Artificial Intelligence: A Modern Approach. 3. ed. New Jersey: Prentice Hall, 2010., p. 1020

⁴² HIGGINS, Chris. A Brief History of Deep Blue, IBM's Chess Computer. Mental Floss, 29 jul. 2017. Disponível em: <https://www.mentalfloss.com/article/503178/brief-history-deep-blue-ibms-chess-computer>. Acesso em: 30 out. 2022.

O grande benefício advindo da IA Específica talvez seja permitir a sua atuação simbiótica com seres humanos, ou seja, a amplificação da capacidade humana de execução de tarefas, mais conhecida como *Human Augmentation*. A combinação homem-máquina permite que o robô execute operações em que é substancialmente superior aos seres humanos, tais como processamento massivo e rápido de dados e execução de atividades mais simples e rotineiras, como coleta de dados e preenchimento de relatórios. Os seres humanos, por outro lado, poderão ser alocados para tarefas onde são melhores e mais eficientes, tais como atividades que dependam de capacidade de percepção e mensuração de valores morais, hermenêutica, intuição, criatividade, sentimentos e empatia.

1.3 Tipos de Aprendizagem de Máquina

A Aprendizagem de Máquina (*Machine Learning* ou ML) consiste no desenvolvimento de algoritmos por meio dos quais o computador aprende a realizar tarefas sem que seja explicitamente programado. A partir do *input* de dados, a IA identifica padrões e constrói um modelo para ser usado para tomar decisões e prever resultados. A OCDE define Aprendizado de Máquina da seguinte forma⁴³:

Sistemas de IA usam cada vez mais ML. Este é um conjunto de técnicas para permitir que as máquinas aprendam de maneira automatizada por meio de padrões e inferências, em vez de instruções explícitas de um humano. As abordagens de ML geralmente ensinam as máquinas a alcançar um resultado, mostrando a elas muitos exemplos de resultados corretos. No entanto, eles também podem definir um conjunto de regras e permitir que a máquina aprenda por tentativa e erro. ML geralmente é usado na construção ou ajuste de um modelo, mas também pode ser usado para interpretar os resultados de um modelo. O ML contém inúmeras técnicas que têm sido usadas por economistas, pesquisadores e tecnólogos há décadas. Eles variam de regressões lineares e logísticas, árvores de decisão e análise de componentes principais a redes neurais profundas.

Existem diferentes tipos de Aprendizagem de Máquina, geralmente classificados em três categorias amplas: (i) aprendizagem supervisionada; (ii) aprendizagem não supervisionada; e (iii) aprendizagem por reforço⁴⁴:

(i) Aprendizagem supervisionada: É o modelo pelo qual a máquina aprende por meio da entrada de dados previamente rotulados (classificados ou categorizados) por seres humanos e teste dos resultados por um supervisor, a fim de aferir se o algoritmo alcançou a resposta esperada. A partir dessas informações, o modelo adquire a capacidade de fazer previsões para novos dados

⁴³ OCDE. Artificial Intelligence in Society. OECD Publishing, Paris, 2019. Disponível em: <https://doi.org/10.1787/eedfee77-en>. Acesso em: 6 mar. 2023.

⁴⁴ RUSSELL, Stuart; NORVIG, Peter. Artificial Intelligence: A Modern Approach. 2 ed. New Jersey: Prentice Hall, 1995.

que não foram previamente conhecidos. Quanto maior for a quantidade de dados, melhor será a acurácia do algoritmo.

(ii) **Aprendizagem não supervisionada:** Neste caso, o modelo identifica padrões sem o auxílio de dados rotulados. Desta forma, o algoritmo busca informações, em princípio, não conhecidas ou que não se tem clareza ou precisão sobre como podem se caracterizar. É utilizado em situações em que se buscam padrões que não são previamente conhecidos, tais como agrupamentos de usuários com características semelhantes e detecção de anomalias.

(iii) **Aprendizagem por reforço:** Neste tipo de aprendizagem, o modelo é treinado a partir de um sistema de consequências, fornecendo um *feedback* de recompensas ou penalidades em função das decisões tomadas. Assemelha-se, assim, a uma forma de adestramento, tal como os animais aprendem. Utiliza, portanto, estímulos positivos e negativos, aprendendo a partir dos seus próprios erros e acertos para aperfeiçoar as suas decisões. Pode ser usado, por exemplo, na recomendação de conteúdo e publicidade, observando se a reação do cliente perante um conteúdo recomendado (reações manifestadas por emoticons, aquisição de um produto etc.).

1.4 Principais domínios de aplicação

Na atualidade, os campos de aplicação de IA que se encontram mais desenvolvidos e em que se vislumbra ganhos e benefícios para a Administração Pública são (i) Automação de Processos Robóticos; (ii) Ferramentas de otimização de pesquisa; (iii) Rede Neural Artificial; (iv) Extração de informações e mineração de dados; (v) Processamento de Linguagem Natural; e (vi) Visão Computacional⁴⁵, explicados mais detalhadamente, a seguir:

(i) **Automação de Processos Robóticos (*Robotic Process Automation* ou RPA):** Os sistemas RPA aprendem observando o usuário executar uma lista de ações e, em seguida, executam a automação repetindo essas tarefas diretamente. Essas ferramentas permitem que os dados sejam manipulados em e entre vários sistemas, por exemplo, ao receber um correio eletrônico com uma cobrança, extrair os dados e inseri-los em um sistema de pagamentos. Desta forma, no caso de auditorias, por exemplo, depois que os dados a serem analisados são carregados no sistema do órgão fiscalizador, a IA pode identificar inconsistências para que sejam objeto de tratamento de auditores.

⁴⁵ BRASIL. Tribunal de Contas da União. Acórdão nº 1139/2022 – Plenário, 25 maio 2022. Brasília, DF: TCU, 2022. Disponível em: https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/*/NUMACORDAO%253A1139%2520ANOACORDAO%253A2022/DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0/%2520. Acesso em: 25 jun. 2022.

- (ii) Ferramentas de otimização de pesquisa: Trata-se de classificar os dados levantados mediante uma pontuação atribuída em função da relevância identificada por um algoritmo de IA.
- (iii) Rede Neural Artificial (RNA): As RNAs são modelos computacionais que simulam o funcionamento de redes neurais biológicas, como neurônios interconectados, que as tornam capazes de aprender e reconhecer padrões. As redes neurais são sistemas que aprendem a partir dos dados, sendo recomendadas para execução de uma grande quantidade de tarefas de difícil solução. Desta forma, as RNAs identificam padrões de transação com base em problemas ou observações identificadas, apontando para possíveis casos similares.
- (iv) Extração de informações e mineração de dados (*Data Mining*): A IA pode realizar o tratamento de informações armazenadas em diferentes sistemas governamentais. Desta forma, o *data mining* permite a extração, organização e análise de dados, confrontando as informações oriundas de diferentes fontes.
- (v) Processamento de Linguagem Natural (*Natural Language Processing* ou NLP): Trata-se de um sistema que permite a comunicação entre IA e humanos. É um recurso utilizado por tradutores e assistentes virtuais como Siri e Alexa. O NLP permite a compreensão da linguagem humana, escrita ou falada. Pode, por exemplo, “entender”, aprender critérios e aplicá-los em consultas futuras de forma autônoma.
- (vi) Visão Computacional: Significa atribuir à máquina não somente a capacidade de “enxergar” a partir de câmeras e sensores, mas, principalmente, distinguir e interpretar dados e informações relevantes. É utilizada para reconhecimento facial, imagens e vídeos para extrair dados, padrões e afins. Um exemplo prático em que se pretende utilizar na Administração Pública é a solução GeoControle, uma IA do TCU para detecção automática de objetos específicos a partir da análise de imagens captadas por satélites. O sistema permite a fiscalização remota de obras civis, como a construção de rodovias e ferrovias mediante técnicas de geoprocessamento e Visão Computacional sobre imagens de satélite⁴⁶.

1.5 Diagnóstico do TCU sobre situação atual e perspectivas da adoção da IA

Neste tópico, pretende-se apresentar uma visão geral do estágio atual e quais são as perspectivas da adoção de IA pela Administração Pública brasileira. Apesar de a presente dissertação não ter a pretensão de realizar uma pesquisa diagnóstica do uso de sistemas

⁴⁶ Ibid, Apêndice XI, p. 124.

inteligentes pela Administração, é importante que se tenha uma visão da relevância, impactos e oportunidades relacionados a essa tecnologia para a sociedade.

A principal política pública sobre IA atualmente em vigor é a Estratégia Brasileira para Inteligência Artificial (EBIA), instituída pelo governo federal por meio da Portaria MCTI n.º 4.617/2021, posteriormente alterada pela Portaria MCTI n.º 4.979/2021⁴⁷. Trata-se de uma iniciativa que reflete uma corrida global entre as nações para se tornarem líderes na utilização de inteligência sintética.

A EBIA define como seus principais objetivos estratégicos⁴⁸: (i) contribuir para a elaboração de princípios éticos para o desenvolvimento e uso de IA responsáveis; (ii) promover investimentos sustentados em pesquisa e desenvolvimento em IA; (iii) remover barreiras à inovação em IA; (iv) capacitar e formar profissionais para o ecossistema da IA; (v) estimular a inovação e o desenvolvimento da IA brasileira em ambiente internacional; e (vi) promover um ambiente de cooperação entre os entes públicos e privados, a indústria e os centros de pesquisas para o desenvolvimento da Inteligência Artificial.

Com relação ao mapeamento da situação atual e perspectivas da adoção da IA na Administração, optou-se por utilizar como principal referência um relatório publicado pelo TCU em maio 2022⁴⁹, tendo em vista que, além de se basear em trabalhos prévios com escores similares realizados por organismos nacionais e internacionais⁵⁰, esse estudo realizou pesquisa quantitativa e qualitativa junto a 293 organizações administrativas, inclusive entrevistando especialistas tanto do meio acadêmico como da iniciativa privada.

Cabe ressaltar, sem embargo, que o relatório do TCU não se restringiu à Administração Pública, visto que teve por objetivo realizar um diagnóstico da utilização de tecnologias de IA e as perspectivas de sua utilização nas organizações públicas dos três poderes da República (Executivo, Legislativo e Judiciário), além de identificar os principais benefícios e riscos relacionados. Também foram avaliados os resultados alcançados e propostos pela EBIA,

⁴⁷ BRASIL. Ministério da Ciência, Tecnologia e Inovações/Gabinete do Ministro. Portaria GM N° 4.617, DE 6 de Abril de 2021. Institui a Estratégia Brasileira de Inteligência Artificial e seus eixos temáticos. [S. l.], 12 abr. 2021. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-gm-n-4.617-de-6-de-abril-de-2021-*-313212172. Acesso em: 22 nov. 2021.

⁴⁸ Ibid.

⁴⁹ BRASIL. Tribunal De Contas da União. Acórdão N° 1139/2022 – Plenário, 25 maio 2022. Disponível em: https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/*/NUMACORDAO%253A1139%2520ANOACORDAO%253A2022/DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0/%2520. Acesso em: 25 jun. 2022.

⁵⁰ Cite-se como exemplo, o Relatório “Uso de Inteligência Artificial pelo Poder Público – Recomendações de Governança” baseado em pesquisa realizada em setembro de 2020 pela Transparência Brasil, em parceria com a Controladoria-Geral da União, o Ministério da Ciência, Tecnologia e Inovação (MCTI) e o Centro de Estudos sobre Tecnologias Web (Ceweb.br); relatório publicado em 2021 pelo Centro de Inovação, Administração e Pesquisa do Judiciário da Fundação Getúlio Vargas.

identificando os principais riscos e oportunidades associados à sua implementação. Por fim, o estudo tratou especificamente da adoção de IA nas atividades de controle⁵¹.

Sobre as vantagens advindas da adoção de IA, as organizações pesquisadas informaram como principais ganhos: a automação de processos repetitivos com aumento de produtividade; redução do tempo e custos na execução dos processos; otimização na alocação de recursos humanos; aumento da confiabilidade e transparência dos processos; maior tempestividade na resposta aos clientes; maior atuação preventiva das organizações em sua área de atuação; e uniformização de procedimentos⁵².

Com relação à utilização de IA pela Administração Pública, o acórdão do TCU apresenta um estudo denominado Government AI Readiness Index 2020, elaborado pela Oxford Insights e pelo Governo do Canadá, sobre a capacidade dos poderes executivos dos Estados para implementar soluções de IA. A pesquisa contém um ranking de 172 países baseado em 33 indicadores. Segundo a referida publicação, o Brasil ocupa a 63ª posição no ranking mundial de capacidade do governo para implementar soluções de IA; e o 6º na América Latina e Caribe. Já em relação ao uso responsável da IA, o Brasil ocupa a 30ª posição de 34 países avaliados. Sem embargo, cabe ressaltar que o estudo é de 2020, quando o Brasil ainda não havia publicado sua estratégia para IA, o que ocorreu em 2021. Avaliando os dados apresentados no Government AI Readiness Index 2020, o TCU ressalta que os principais pontos positivos para o incentivo à implementação de soluções de IA no Brasil são a disponibilidade e a representatividade de dados.

Com relação ao questionário encaminhado para 263 órgãos e entidades dos três poderes da República, as indagações apresentadas pelo TCU tinham por objetivo averiguar o nível de maturidade da utilização de IA⁵³. Conforme fora constatado a partir das respostas obtidas, as soluções de IA estão sendo mais utilizadas pelos órgãos do Poder Judiciário, para processos administrativos e em atividades de controle. Constatou-se, portanto, que a IA está sendo predominantemente utilizada para otimização de atividades estatais *interna corporis*, com menor participação na prestação de serviços públicos diretamente direcionados aos cidadãos, como transporte, segurança pública e proteção social.

Segundo a análise realizada pelo TCU, os resultados obtidos são reflexo do quadro atual de redução de pessoal e a consequente e necessária otimização de processos. Por outro lado, o

⁵¹ Ibid.

⁵² Ibid, parágrafo 93.

⁵³ O inteiro teor do questionário encontra-se no sítio <https://portal.tcu.gov.br/levantamento-Inteligência-Artificial.htm>.

Órgão de Controle ressalta que, “dado o aspecto disruptivo e inovador que IA traz, verifica-se que há espaço para crescimento do seu uso em soluções que atingem diretamente a sociedade”⁵⁴.

Com relação à quantidade de soluções de IA no Poder Público Federal por domínio de aplicação, em que pese não ser o foco da presente dissertação, a pesquisa evidenciou que o PLN é amplamente utilizado no Poder Judiciário, refletindo a necessidade de tratamento de informações constantes em processos judiciais. Também se verificou a utilização de PLN no desenvolvimento de *chatbots* para interação com os usuários externos e o provimento de serviços. Constatou-se, ainda, a utilização de modelos preditivos especialmente nas áreas de controle e fiscalização⁵⁵.

Por fim, o relatório do TCU realizou uma avaliação da EBIA enquanto política pública da Administração Pública Federal para adoção e estímulo a investimentos públicos e privados em IA, apontando, em síntese, que os objetivos e ações estratégicas definidas não possuem indicadores claros para a ação governamental, conforme requer a metodologia SMART (*Specific, Measurable, Achievable, Relevant e Time*), ou seja, não são específicos, mensuráveis, realistas (alcançáveis) e delimitados em um recorte temporal. É percebido apenas a propriedade de serem apropriados (relevantes)⁵⁶.

O Órgão de Controle indica que a EBIA não possui um mapeamento do estágio de referência inicial (marco zero, situação inicial)⁵⁷. O TCU constata, ainda, falhas formais na apresentação do modelo lógico e sobre como a EBIA incide sobre os problemas/oportunidades, projetando seus resultados e impactos de longo prazo⁵⁸; as estruturas de governança e gestão necessárias à implementação da política não estão formalmente institucionalizadas⁵⁹; e a estrutura de monitoramento e avaliação da EBIA não está formalmente definida em norma⁶⁰.

Diante disso, o TCU conclui que, em razão dessas falhas, na prática, os objetivos apresentados na EBIA correspondem mais a diretrizes estratégicas que orientam as ações derivadas dela. Também tece críticas, entendendo não ser viável ou aceitável que os objetivos definidos em uma política pública não sejam expressos em termos de resultado, sendo essencial

⁵⁴ BRASIL. Tribunal de Contas da União. ACÓRDÃO Nº 1139/2022 – Plenário, 25 maio 2022. Disponível em: https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/*/NUMACORDAO%253A1139%2520ANOACORDAO%253A2022/DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0/%2520. Acesso em: 25 jun. 2022, parágrafo 89.

⁵⁵ Ibid, parágrafo 90

⁵⁶ Ibid, parágrafo 116 et seq.

⁵⁷ Ibid, parágrafo 127 et seq.

⁵⁸ Ibid, parágrafo 139 et seq.

⁵⁹ Ibid, parágrafo 151 et seq.

⁶⁰ Ibid, parágrafo 158 et seq.

a definição de metas quantitativas para a avaliação da eficácia, eficiência e efetividade da política. Aponta, bem assim, que o uso de indicadores quantitativos e/ou qualitativos apropriados deveria fazer parte do desenho da Política.

CAPÍTULO II - DA PROBLEMÁTICA ÉTICA RELACIONADA À TECNOLOGIA E À IA

2.1 A ideologia do progresso técnico: da redenção à destruição da humanidade

Antes de adentrar na problemática ética relacionada à tecnologia e à IA, é necessária uma breve noção a respeito daquilo que chamamos de “técnica”. Para isso, adotar-se-á o conceito trazido por Aristóteles, que sistematizou o conhecimento a partir da divisão dos campos do saber. O Estagirita⁶¹ dividiu as ciências em três grandes ramos: a) as ciências teóricas, cujo objeto é o saber em si mesmo; b) as ciências práticas, em que a busca do saber se destina à conduta do homem, de forma a alcançar a perfeição moral; e, por fim, c) as ciências poéticas ou produtivas, que buscam o saber em função do fazer, ou seja, com o objetivo de produzir determinados objetos, tais como a retórica e a poética.

A *techné* (ou arte) pode ser entendida como o conhecimento decorrente das ciências poéticas ou produtivas, visto que tem como objeto o elemento resultante da ação humana fabricadora ou produtora, a obra ou artefato resultante, e não o ato humano em si mesmo (que é objeto das ciências práticas). A técnica, portanto, é o conhecimento da razão ou causa que permite a produção de resultados externos ao agente. As artes ou técnicas são tão numerosas quanto a engenhosidade humana é capaz de produzir coisas ou obras.

O deslumbramento da humanidade pela técnica tem o seu início no séc. XVI, quando havia uma grande expectativa e esperança em relação ao desenvolvimento tecnológico: a descoberta da América proporcionada por invenções da época, como o astrolábio, eram sinais do alvorecer de uma nova era para a humanidade. Francis Bacon⁶², tido como um dos precursores da ciência moderna, afirmava que a união entre experiência e método permitiria ao homem a “*victoria cursus artis super naturam*” ou seja, a vitória da arte sobre a natureza. Para Bacon⁶³, saber é poder: a ciência é o fundamento da técnica, onde todos os problemas humanos serão resolvidos por especialistas capazes de transformar o saber em uma constelação de intervenções eficazes.

Essa nova concepção foi aplicada inclusive no campo da teologia, visto que significaria, finalmente, o domínio do homem sobre a criação, restabelecendo a ordem perdida em razão do

⁶¹ REALE, Giovanni; ANTISERI, Dario. História da Filosofia: Antiguidade e Idade Média. 3. ed. São Paulo: Paulus, 1990. p. 178-179.

⁶² BACON, Francis. Novum Organum ou Verdadeiras Indicações Acerca da Interpretação da Natureza. Tradução e notas de José Aluysio Reis de Andrade. Pará de Minas: M&M Editores Ltda, 2003.

⁶³ REALE, Giovanni. O saber dos antigos: Terapia para os tempos atuais. 3. ed. São Paulo: Loyola, 2011. p. 80.

pecado original. A restauração do paraíso, segundo Bacon, chama-se fé no progresso, da ligação recém-descoberta entre ciência e prática⁶⁴. As novas descobertas eram apenas um prenúncio do “reino do homem”, do paraíso na terra, de um mundo melhor não situado no além, mas no aquém. Em Nova Atlântida, Bacon⁶⁵ retrata uma sociedade utópica onde profetiza a descoberta de invenções que bem mais tarde se concretizaram, tais como o avião e o submarino. Bacon inaugura, assim, a ideologia do progresso. Nas palavras do filósofo Giovanni Reale⁶⁶, iniciou-se “uma nova e laica ‘latria da ciência’”.

A primeira grande contradição sofrida pela proposta de redenção da humanidade por meio da técnica se deu entre os séculos XVIII e XIX, quando o desenvolvimento tecnológico acelerado proporcionado por invenções como a máquina a vapor culminou na Revolução Industrial. Surge, assim, a classe dos trabalhadores da indústria ou proletariado industrial. Como é de conhecimento geral, o proletariado fora submetido a condições de vida e trabalho extremamente degradantes e desumanas. Ficou evidenciado, pela primeira vez, que o progresso tecnológico não resultaria, necessariamente, em uma melhoria nas condições de vida e bem-estar para toda a sociedade humana.

Foi então que surgiram as primeiras teorias críticas à “fé no progresso”, relacionadas à condição dos operários e materializadas pelo *Manifesto Comunista*, escrito pelos fundadores do socialismo científico Karl Marx e Friedrich Engels, e pela Encíclica *Rerum Novarum* do Papa Leão XIII. Esses escritos denunciavam de forma veemente que, com a Revolução Industrial, os avanços técnicos estavam sendo utilizados como um instrumento de maximização de lucros nas mãos de poucos às custas do exaurimento e exploração da classe trabalhadora.

Já no séc. XX, em um mundo mergulhado em teorias científicas, positivistas e socialistas, a técnica foi vista como instrumento para a reengenharia humana, para criação “do novo homem”, seja no aspecto racial, psicológico ou social. Duas grandes distopias literárias do século XX: 1984, de George Orwell, e Admirável Mundo Novo, de Aldous Huxley, de forma alegórica, sugerem um futuro nefasto para humanidade, caso essas teorias sejam levadas às últimas consequências. Segundo Erich Fromm⁶⁷, pode-se dizer que 1984 lembra mais as ditaduras nazista e stalinista, enquanto Admirável Mundo Novo é um retrato do progresso do mundo ocidental industrial, posto que acompanha a tendência do presente sem mudar sua essência. Todavia, tanto Orwell como Huxley chegam à mesma conclusão: pela adoção de

⁶⁴ REALE, Giovanni. *O saber dos antigos: Terapia para os tempos atuais*. 3. ed. São Paulo: Loyola, 2011, p. 129.

⁶⁵BACON, Francis. *Nova Atlantis*. 1626. Disponível em: http://www.hs-augsburg.de/~harsch/Chronologia/Lspost17/Bacon/bac_atla.html. Acesso em: 3 set. 2022.

⁶⁶ Ibid, p. 80

⁶⁷ FROMM, Erich. *The Sane Society*. New York: Rinehart & Company, 1955.

técnicas que atualmente são de conhecimento comum, é possível a destruição da humanidade dentro do homem e realizar a sua “reprogramação”.

Aldous Huxley⁶⁸, em *Admirável Mundo Novo*, retrata uma sociedade distópica no ano 2540 (equivalente ao ano 632 “depois de Ford”, conforme referencia a obra) essencialmente materialista e hedonista, gerada a partir de técnicas de manipulação genética e adestramento psicológico desde a mais tenra infância, com eliminação da individualidade, do senso crítico, da espiritualidade e do sofrimento, bem como descarte dos “seres humanos inúteis”: do senil e do enfermo.

Huxley⁶⁹ projetou o seu futuro em função da atmosfera ideológico-cultural presente nos anos 1920 e 1930 levada às últimas consequências: a construção de uma Nova Ordem Mundial pasteurizada, fruto de um mix de ideologias tais como cientificismo e eugenia; marxismo e leninismo; fordismo e utilitarismo; relativismo antropológico e psicanálise; e técnicas de condicionamento pavlovianas.

Por sua vez, Orwell⁷⁰, em 1984, sugere que a nova forma de industrialismo gerencial, na qual o homem constrói máquinas que agem como homens e adestra pessoas que agem como máquinas, conduz a uma era de desumanização e completa alienação, na qual pessoas são transformadas em coisas e se tornam apêndices do processo de produção e consumo. Cria-se assim, uma sociedade de homens-autômatos, sem individualidade, amor e pensamento crítico, e que não serão sequer capazes de percebê-lo em decorrência do “duplipensamento”⁷¹.

A percepção da técnica como instrumento de dominação do homem pelo homem encontra-se refletida no pensamento de filósofos como Foucault, Gilles Deleuze e mais recentemente, Byung-Chul Han, ao tratarem dos modelos teóricos das “sociedades disciplinares” e “sociedades de controle”.

Segundo Gilles Deleuze⁷², nos séculos XVIII e XIX, Foucault estudou as sociedades disciplinares como forma de exploração humana, atingindo o seu apogeu no início do século XX. As sociedades disciplinares caracterizam-se pela organização de grandes meios de confinamento. O indivíduo não cessa de passar de um espaço fechado para o outro, cada um

⁶⁸ HUXLEY, Aldous. *Admirável Mundo Novo*. São Paulo: Biblioteca Azul, 2014.

⁶⁹ Em que pese terem sido imaginadas para um século futuro, o próprio Huxley, em 1959, confessava em *Brave New World Revisited* que as profecias feitas em 1931 estão para realizar-se muito mais depressa do que ele calculava. Algumas das “profecias” de Huxley que se concretizaram antes do tempo são as técnicas e invenções sobre: pastilhas de estimulação sexual, anticoncepcionais e rejuvenescimento; fertilização in vitro; manipulação genética; programação neurolinguística; e psicotrópicos. HUXLEY, Aldous. *Brave New World Revisited*. New York: Harper Perennial Modern Classics, 2006.

⁷⁰ ORWELL, George. *1984*. São Paulo: Companhia das Letras, 2009.

⁷¹ FROMM, Erich. *The Sane Society*. New York: Rinehart & Company, 1955.

⁷² DELEUZE, Gilles. *Conversações*. Rio de Janeiro: Editora 34, 2017. p. 219-226.

com suas leis: a família, a escola, a caserna, a fábrica, de vez em quando o hospital, e, eventualmente a prisão, que é considerado o meio de confinamento por excelência e que serve de modelo analógico e de referência para as demais formas de confinamento. “A heroína de Europa 51 pode exclamar, ao ver operários, ‘pensei estar vendo condenados...’⁷³. Esses espaços de confinamento convergem para um proprietário, que é o estado ou potência privada.

O animal que representa os meios de confinamento é a toupeira. Apesar de a disciplina ser de longa duração, infinita, era descontínua. O homem da disciplina era um produtor descontínuo de energia⁷⁴. Byung-Chul Han⁷⁵, explica que a toupeira é o animal da sociedade disciplinar porque o sujeito muda de um ambiente de reclusão para outro. Assim, ele se move em um sistema fechado. Os internos nestes ambientes permitem ser distribuídos no espaço e ordenados no tempo. A toupeira se move em espaços pré-construídos e, assim, se submete a restrições espaciais. É uma trabalhadora e um sujeito submisso.

Entretanto, Deleuze⁷⁶ ressalta que as disciplinas conheceriam uma crise, em favor de novas forças que se instalavam lentamente e que se precipitariam depois da Segunda Guerra Mundial. Há uma crise geral de todos os ambientes de reclusão, visto que o seu fechamento e rigidez não são apropriados para as novas formas de produção pós-industriais, imateriais e em rede, que pressionam para uma maior abertura e dissolução de fronteiras.

Desta forma, as sociedades disciplinares estão sendo substituídas pelas sociedades de controle⁷⁷:

"Controle" é o nome que Burroughs propõe para designar o novo monstro, e que Foucault reconhece como nosso futuro próximo. Paul Virillo também analisa sem parar as formas ultra rápidas de controle ao ar livre, que substituem as antigas disciplinas que operavam na duração de um sistema fechado. Não cabe invocar produções farmacêuticas extraordinárias, formações nucleares, manipulações genéticas, ainda que elas sejam destinadas a intervir no novo processo. Não se deve perguntar qual é o regime mais duro, ou o mais tolerável, pois é em cada um deles que se enfrentam as liberações e as sujeições. Por exemplo, na crise do hospital como meio de confinamento, a setorização, os hospitais-dia, o atendimento a domicílio puderam marcar de início novas liberdades, mas também passaram a integrar mecanismos de controle que rivalizam com os mais duros confinamentos. Não cabe temer ou esperar, mas buscar novas armas.

O regime disciplinar é organizado como um "corpo". É um regime biopolítico. O regime neoliberal, por outro lado, comporta-se como “alma”. A psicopolítica é sua forma de governo. Ela institui entre os indivíduos uma rivalidade sem fim como uma competição saudável, como

⁷³ Ibid.

⁷⁴ Ibid.

⁷⁵ HAN, Byung-Chul. Psicopolítica: O neoliberalismo e as novas técnicas de poder. Belo Horizonte: Âuiné, 2018.

⁷⁶ Ibid.

⁷⁷ Ibid.

uma excelente motivação. Motivação, projeto, competição, otimização e iniciativa são inerentes à técnica de dominação psicopolítica do regime neoliberal⁷⁸.

A toupeira não pode suportar a abertura promovida pela sociedade atual. O animal da sociedade de controle neoliberal é a serpente: o homem do controle é ondulatório, funcionando em órbita, num feixe contínuo. Ao contrário da toupeira, a serpente não se move em espaços confinados. Ela define o espaço a partir de seu movimento. A serpente é uma empreendedora. A serpente é um projeto, na medida em que gera espaço a partir de seu movimento⁷⁹.

A passagem da toupeira à serpente, do sujeito ao projeto, não é uma irrupção para um modo de vida totalmente diferente, mas uma mutação, um agravamento do próprio capitalismo: a reduzida capacidade de movimento da toupeira limita a produtividade. Mesmo quando trabalha de forma disciplinada, não ultrapassa um certo nível de produtividade. A serpente anula essa limitação por meio de novas maneiras de movimento. Desta forma, o sistema capitalista baseado no modelo da toupeira muda para o modelo da serpente para aumentar a produtividade. A serpente encarna sobretudo a culpa, as dívidas, que o regime liberal estabelece como meio de dominação. O homem não é mais o homem confinado, mas o homem endividado.⁸⁰

Com relação à educação, as escolas e universidades transformaram-se em centros profissionalizantes cujo objetivo primordial é prover formação para a técnica, para o exercício de uma atividade profissional. Abandonou-se, assim, o projeto de *universitas*, da formação integral da pessoa humana, sendo cada vez menos valorizado o estudo das artes, da pesquisa científica, da cultura e das humanidades⁸¹. Hoje, ensinam-se competências, não virtudes. Formam-se tecnocratas. Isso explica o sucesso de escolas e cursinhos preparatórios para o vestibular e a franca decadência dos colégios de ensino mais humanista.

⁷⁸ HAN, Byung-Chul. Psicopolítica: O neoliberalismo e as novas técnicas de poder. Belo Horizonte: Âuiné, 2018.

⁷⁹ Ibid

⁸⁰ Ressalta ainda o filósofo sul-coreano que, diferente da suposição de Marx, a contradição entre as forças produtivas e as relações de produção não pode ser superada pela revolução comunista. Não há mais um “nós político” capaz de agir em comum, pois a auto exploração atinge todas as classes, o que inviabiliza completamente a ditadura do proletariado. O neoliberalismo, como mutação do capitalismo, torna o trabalhador um empreendedor. As figuras do proletariado explorado e burguês explorador confundem-se na mesma pessoa. Ou seja, é o próprio indivíduo que se auto explora de bom grado, levando ao esgotamento, à depressão e ao burn out. O indivíduo que acredita estar livre se submete a um processo de auto exploração guiado pela obrigação de desempenho e otimização. A coação passa a ser interna. Enquanto as pessoas estão em constante produção e competição, o capital se reproduz de maneira desenfreada. A morte de Deus proclamada por Nietzsche foi substituída por uma nova forma de transcendência, uma nova forma de subjetivação, um novo senhor. Walter Benjamin concebe o capitalismo como religião. Manifestação clara disso é o consumismo e o materialismo desenfreado da sociedade atual. O exercício do culto a esse novo deus, ao invés de ser libertador, perpetua o estado de falta de liberdade. “É o primeiro caso de culto não expiatório, mas culpabilizador”, diz Benjamin. HAN, 2018.

⁸¹ DELEUZE, Gilles. Conversações. Rio de Janeiro: Editora 34, 2017. p. 219-226.

Tal como ressalta Deleuze⁸², nas sociedades de controle, o monitoramento não ocorre apenas em espaços confinados, visto que, por meio da tecnologia de informática e comunicação, é possível definir a localização geográfica de um indivíduo em qualquer parte do globo:

Não há necessidade de ficção científica para se conceber um mecanismo de controle que dê, a cada instante, a posição de um elemento em espaço aberto, animal numa reserva, homem numa empresa (coleira eletrônica). Félix Guattari imaginou uma cidade onde cada um pudesse deixar seu apartamento, sua rua, seu bairro, graças a um cartão eletrônico (dividual) que abriria as barreiras; mas o cartão poderia também ser recusado em tal dia, ou entre tal e tal hora; o que conta não é a barreira, mas o computador que detecta a posição de cada um, lícita ou ilícita, e opera uma modulação universal.

Os smartphones e celulares funcionam, assim, como uma verdadeira coleira eletrônica, na medida em que disponibilizam com precisão a localização de um indivíduo em tempo real. Por outro lado, pela gamificação e pelo estabelecimento de um sistema de score social, como, conforme será abordado adiante, já vem sendo utilizado na China⁸³, um indivíduo tem seus direitos concedidos ou denegados - tais como obtenção de um emprego, a compra de uma passagem, a concessão de um crédito - em função da decisão automatizada de um algoritmo de IA.

Byung-Chul Han⁸⁴ provoca ainda reflexão sobre o controle psicopolítico e digital dos tempos atuais. O referido autor ressalta que, por essa vigilância, o neoliberalismo contemporâneo realiza uma manipulação imperceptível da liberdade humana, por meio de um estímulo positivo à comunicação amigável e ilimitada, sendo um meio muito eficiente de controle psicopolítico do indivíduo. O sujeito de desempenho, que se julga livre, é na realidade um servo: um servo absoluto, na medida em que, sem um senhor, explora voluntariamente a si mesmo.

O filósofo sul-coreano compara o estado de vigilância criado por Orwell em 1984 (as teletelas presentes em todos os ambientes) com o panóptico digital em que vivemos hoje. Panóptico é um termo utilizado para designar uma penitenciária ideal, concebida pelo filósofo e jurista inglês Jeremy Bentham em 1785, que permite a um único vigilante observar todos os prisioneiros, sem que estes possam saber se estão ou não sendo vigiados⁸⁵.

⁸² Ibid.

⁸³ THE MAKING. Social Credit: China's Digital Dystopia. Direção: Janis Mackey Frayer. NBC News, 2018. Disponível em: <https://www.youtube.com/watch?v=0cGB8dCDf3c>. Acesso em: 30 out. 2022.

⁸⁴ HAN, Byung-Chul. Psicopolítica: O neoliberalismo e as novas técnicas de poder. Belo Horizonte: Âuiné, 2018.

⁸⁵ Ibid.

O panóptico digital, agindo de forma obscura (*black box* da IA onde não há transparência nas decisões automatizadas), permite a criação de um banco de dados sobre padrões coletivos de comportamento que são utilizados para o controle, disciplina e incentivo de determinadas condutas. Com o *Big Data*, nossos hábitos digitais são milimetricamente guardados, quantificados, montando assim um perfil comportamental, podendo até representar de forma mais exata a nossa imagem do que aquela que nós possuímos de nós mesmos. Ou seja, este modelo pode tornar legíveis desejos do nosso inconsciente⁸⁶.

Tendo em vista a quantidade de informação que se lança voluntariamente na rede, o próprio conceito de proteção de dados se torna obsoleto. Han lembra, inclusive, da decisão do Tribunal Constitucional Alemão, que reconheceu a autonomia do direito à autodeterminação informativa a partir da análise da Lei do Censo alemã de 1983, onde havia a previsão de que os cidadãos fornecessem uma série de dados pessoais para mensurar estatisticamente a distribuição espacial e geográfica da população. O Tribunal declarou a inconstitucionalidade parcial da norma em razão de sua vagueza e amplitude, o que possibilitava o cruzamento dos dados coletados com outros registros públicos, bem como a sua transferência para outros órgãos da Administração. Para o Tribunal, a capacidade do indivíduo de autodeterminar seus dados pessoais é parcela fundamental do seu direito de desenvolver livremente sua personalidade. O estágio de desenvolvimento da tecnologia informática utilizada no processamento das informações levantadas com o censo era um fator determinante, visto que a elaboração de perfis formados sobre dados dos indivíduos teria potencial ilimitado e prováveis resultados danosos aos direitos individuais⁸⁷.

Desta forma, Han⁸⁸ entende que a capacidade de autodeterminação informacional pode ser, hoje, considerada uma parte essencial da liberdade. Ressalta o autor que a decisão do Tribunal Constitucional Alemão decorreu da necessidade de impor limites ao poder de extração e processamento de dados pessoais contra a vontade dos cidadãos. No entanto, a verdade é que, atualmente, indivíduos expõe voluntariamente os dados pessoais sem qualquer coerção, sem avaliar as consequências. Conclui Byung-Chul Han que esse caráter incontrolável representa uma gravíssima crise da liberdade.

⁸⁶ Ibid.

⁸⁷ Ibid.

⁸⁸ Ibid.

Mas não é somente a pessoa humana que é explorada e exaurida pela tecnologia. Edgar Morin e Anne Brigitte Kern⁸⁹ ressaltam os nefastos danos causados ao meio ambiente em função desse “tecnomorfismo”:

A falsa racionalidade, ou seja, a racionalização abstrata e unidimensional, triunfa no campo: os loteamentos apressados, os sulcos demasiado profundos e longitudinais, o corte de matas e a desarborização não controlados, o asfaltamento de estradas, o urbanismo que visa apenas à rentabilização da superfície do solo, a pseudo-funcionalidade planejadora que não leva em conta necessidades não quantificáveis e não identificáveis por questionários, tudo isso multiplicou os subúrbios retalhados, as cidades novas que se tornam rapidamente núcleos isolados de tédio, de sujeira, de degradações, de incúria, de despersonalização, de delinquência. Disso resultam catástrofes humanas cujas vítimas e consequências não são contabilizadas, e um agravamento das catástrofes naturais, como recentemente em Vaison-la-Romaine [Departamento de Vaucluse, França].

Por toda parte, e durante dezenas de anos, soluções pretensamente racionais, apresentadas por experts convencidos de agir em nome da razão e do progresso e de encontrar apenas superstições nos costumes e temores das populações, empobreceram ao enriquecer, destruíram ao criar.

Em suma, Teodoro W. Adorno⁹⁰ sintetizou de forma cabal o ceticismo na “fé no progresso” no séc. XX, especialmente diante do acontecimento de duas guerras mundiais, a Guerra Fria e a corrida armamentista, genocídios por todas as partes e a fome ainda presente no mundo. O desenvolvimento tecnológico não levaria à emancipação da humanidade: “visto de perto, seria o progresso da funda⁹¹ à megabomba”. Na Alemanha nazista, a eficiência técnica desvinculada da ética e do bem comum levou à Solução Final⁹², uma verdadeira máquina de extermínio do povo judeu. Edgar Morin⁹³ afirma ainda que a religião da técnica não tem mais regras: “A corrida da tríade que se encarregou da aventura humana, ciência/técnica/indústria, é descontrolada. O crescimento é descontrolado, seu progresso conduz ao abismo”.

Fica evidente, portanto, que o progresso não possui uma bondade intrínseca e indissociável. Não há dúvidas de que o desenvolvimento tecnológico oferece uma miríade de possibilidades para o bem da humanidade, mas também abre muitos caminhos para o mal, caminhos que antes não existiam. A tecnologia, por si só, não tem o condão de eliminar todas

⁸⁹ MORIN, Edgar; KERN, Anne-Brigitte. Terra-Pátria. Porto Alegre: Sulina, 2003, p. 155-156

⁹⁰ BENEDICTUS PP. XVI. Carta Encíclica SPE SALVI. Libreria Editrice Vaticana, Roma, novembro 2007. Disponível em: https://www.vatican.va/content/benedict-xvi/pt/encyclicals/documents/hf_ben-xvi_enc_20071130_spe-salvi.html. Acesso em: 3 set. 2022.

⁹¹ Instrumento de arremesso de pedras.

⁹² A SOLUÇÃO Final. Enciclopédia do Holocausto, 19 mai. 2022. Disponível em: <https://encyclopedia.ushmm.org/content/pt-br/article/the-final-solution#:~:text=A%20%20E2%80%9CSolu%C3%A7%C3%A3o%20Final%E2%80%9D%20foi%20o,em%20massa%20dos%20judeus%20europeus>. Acesso em: 3 set. 2022.

⁹³ MORIN, Edgar; KERN, Anne-Brigitte. Terra-Pátria. Porto Alegre: Sulina, 2003, p. 92

as desigualdades e injustiças sociais, de criar “o paraíso na terra” tal como romantizava Francis Bacon. Mal utilizada, a técnica pode se converter em instrumento de dominação e exaurimento do homem pelo homem, em um progresso abissal em direção à destruição do planeta.

2.2 Problemas éticos relacionados à adoção da IA pela Administração Pública

Conforme fora mencionado na Introdução, constatou-se que sistemas inteligentes vêm sendo utilizados como instrumento de um Capitalismo de Vigilância⁹⁴ cada vez mais feroz e crescente, voltado à Economia da Atenção⁹⁵. Percebeu-se, ainda, uma enorme assimetria informacional entre pessoas singulares, de um lado, e organizações do setor público e privado, de outro, que utilizam IA como ferramenta para um verdadeiro “hackeamento” do ser humano, a partir da coleta de dados pessoais que nutrem algoritmos que pretendem nos conhecer melhor do que nós mesmos nos conhecemos, a partir da definição de um “perfil comportamental” ou “*data profiling*”⁹⁶, com o objetivo de fazer previsões e intervenções sobre nossas emoções e condutas.

Para atingir um determinado objetivo, a computação cognitiva pode adotar meios flagrantemente contrários à ética e valores consagrados constitucionalmente, tal como ocorrera no caso Molly Russell, uma adolescente de 14 anos que cometeu suicídio sob influência de conteúdos antiéticos recomendados pelos algoritmos de uma rede social, tais como *self-harm* e instigação ao suicídio. Questiona-se, ainda, a publicidade abusiva direcionada a crianças e adolescentes instigando o consumo de alimentos, compras relacionadas a jogos online etc.⁹⁷.

Constatou-se, ainda, que a polarização ideológica e política crescente em todo mundo tem sido, em boa parte, atribuída à formação de “Filtros Bolha” (*Filter Bubble*) na internet, fenômeno este identificado pelo ativista Eli Pariser⁹⁸.

⁹⁴ ZUBOFF, Shoshana. A Era do Capitalismo de Vigilância. Edição digital: Intrínseca, 2021.

⁹⁵ SIMON, Herbert A. Designing Organizations for an Information-Rich World. In: COMPUTERS, communications, and the public interest. Baltimore: Johns Hopkins Press, 1971.

⁹⁶ O Regulamento Geral de Proteção de Dados da União Europeia define o “*data profiling*” como “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 económica, saúde, preferências pessoais, interesses, fiabilidade, comportamento, localização ou deslocações”. UNIÃO EUROPEIA. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho. Relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). Art. 4º, item 4. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679&qid=1677953718324&from=PT>. Acesso em: 04 mar. 2022.

⁹⁷ CRAWFORD, Angus. Instagram “helped kill my daughter.” BBC News, Inglaterra, 22 jan. 2019. Disponível em: <https://www.bbc.com/news/av/uk-46966009>. Acesso em: 07 nov. 2019.

⁹⁸ PARISER, Eli. The Filter Bubble: What the Internet Is Hiding from You. Nova York: Penguin Books, 2011.

Pariser⁹⁹ observou que os resultados de buscas online são únicos para cada pessoa, gerando um conjunto de informações variável em função do perfil comportamental construído pela rede social a partir dos dados pessoais coletados. Desta forma, alerta que algoritmos de redes sociais realizam uma edição invisível e pouco transparente de conteúdos ideologicamente alinhados e autorreferenciados à pessoa, evitando informações conflitantes ou relacionadas a pensamentos políticos divergentes, mesmo quando gerados por amigos e pessoas próximas.

O que se tem percebido, em suma, é que, por mais que máquinas pensantes possam replicar operações inerentes à natureza humana, não alcançam a percepção ética do seu proceder¹⁰⁰, o que pode levar a decisões arbitrárias, preconceituosas ou ilegais, causando uma “erosão da autodeterminação humana”, prática esta conhecida como “*nudging*”, tal como apontado por Danilo Cesar Maganhoto Doneda, Carlos Affonso Pereira de Souza et al¹⁰¹:

A IA pode corroer a autodeterminação pessoal, pois pode induzir a mudanças não planejadas e indesejadas nos comportamentos humanos para acomodar rotinas que facilitam a automação e a vida das pessoas. O poder de previsão da IA e o denominado “*nudging*”, mesmo que não intencional, devem fomentar, e nunca minar a dignidade humana e a autodeterminação.

Especificamente em relação à Administração Pública, conforme fora exposto na Introdução, o amplo poder de acesso a dados pessoais, com a formação de *Big Datas* a partir do compartilhamento de informações entre diferentes órgãos e entes públicos, aliado ao uso de IA, podem representar graves riscos ao Estado Democrático de Direito e aos direitos e garantias fundamentais.

A seguir, apresenta-se situações geralmente apontadas como problemas éticos relacionados ao uso de IA pela Administração Pública: i) Ditadura Algorítmica, em prejuízo da privacidade e autodeterminação informacional; ii) Algoritmos opacos de IA; iii) Substituição da subjetividade do gestor pela “subjetividade” do robô; iv) Vieses algorítmicos (*bias*); v) Falta de publicidade e impedimento ou dificuldade de acesso a dados públicos.

O objetivo da indicação desses casos não é realizar um levantamento exaustivo, mas trazer a lume a existência de uma nova conjuntura proporcionada pela utilização de inteligência

⁹⁹ Ibid.

¹⁰⁰ MELO, Vinicius Holanda; PEREIRA JÚNIOR, Antonio Jorge. Os limites da inteligência artificial no exercício da prudência: as atividades jurídicas correm risco? *Revista dos Tribunais Online*, v. 1015, p. 107-127, 2020. Disponível em: <https://www.thomsonreuters.com.br/content/dam/openweb/documents/pdf/Brazil/revistas-especializadas/rt-1015-vinicius-holanda-melo-e-antonio-jorge-pereira-junior-os-limites-da-inteligencia-artificial.pdf>. Acesso em: 04 jan. 2022.

¹⁰¹ DONEDA, Danilo Cesar Maganhoto; SOUZA, Carlos Affonso Pereira de Souza et al. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. *Pensar*, Fortaleza, v. 23, n. 4, p. 1-17, dez, 2018.

sintética e que, atualmente, não possui uma resposta a ser dada pelo ordenamento jurídico, ao menos de forma evidente. A presente dissertação pretende, assim, propor uma solução que seja suficientemente abrangente e que dê direcionamento para questões éticas presentes e futuras, não se restringindo a resolver os casos ora apontados.

2.2.1 Ditadura Algorítmica, em prejuízo da privacidade e autodeterminação informacional

Conforme visto, a IA, como um panóptico digital, pode ser utilizada pela Administração como um poderoso instrumento de vigilância totalitária, perseguição política, controle comportamental e manipulação da liberdade. Há um evidente prejuízo à privacidade e à autodeterminação informacional que, na esteira do pensamento de Byung-Chul Han antes citado¹⁰², são hodiernamente considerados como uma parte essencial da liberdade individual e desenvolvimento da personalidade humana.

De forma a ilustrar a atualidade do problema, apresenta-se o caso do sistema de score ou crédito social criado pelo estado chinês. Sob a justificativa de resolver uma crise de confiança entre cidadãos, o governo chinês criou um sistema de IA onde o cidadão ganha ou perde pontos em função daquilo que o governo entende como comportamentos bons ou maus. Esses pontos são utilizados para estabelecer restrições ou benefícios no acesso a bens e serviços públicos. O score social inclui direitos como usar um meio de transporte público, maior probabilidade de receber ofertas de emprego, facilitação de crédito, acesso à saúde etc.¹⁰³.

Mediante um enorme sistema de câmeras inteligentes instaladas por todo o país, são avaliados comportamentos como atravessar dentro da faixa de pedestres, tocar música com volume alto, comer em metrô ou trens e violar as regras de trânsito. Ressalta-se ainda que a lista negra de pessoas consideradas “não confiáveis” pelo estado chinês, o que inclui dissidentes e opositores políticos, é de acesso público, com localização por GPS das pessoas próximas e suas respectivas notas¹⁰⁴.

Vê-se, assim, que permitir que sistemas como o crédito social chinês, em que há um controle amplo e irrestrito do comportamento moral dos cidadãos, sejam instituídos pelo

¹⁰² HAN, Byung-Chul. *Psicopolítica: O neoliberalismo e as novas técnicas de poder*. Belo Horizonte: Âuiné, 2018.

¹⁰³ NBC NEWS. *SOCIAL Credit: China's Digital Dystopia In: The Making*. Direção: Janis Mackey Frayer. 2018. Disponível em: <https://www.youtube.com/watch?v=0cGB8dCDf3c>. Acesso em: 30 out. 2022.

¹⁰⁴ Ibid.

Estado, é, *ipso facto*, uma grave lesão ao Estado Democrático de Direito. Ocorre que não existem regras ou diretrizes na Administração Pública para avaliação de grau de riscos de robótica cognitiva, bem como medidas recomendadas em função desses riscos, de forma a coibir o seu mau uso, desvio de poder ou utilização para finalidades escusas, especialmente em atividades de alto risco de prejuízo a direitos e garantias fundamentais.

2.2.2 Algoritmos opacos de IA

Há algoritmos de IA que funcionam como um sistema fechado (conhecidos como *black box*), onde não é possível uma explicação sobre como se chegou ao resultado ou decisão específica. Questiona-se, nesses casos, a falta de transparência ou obscuridade de decisões automatizadas, não só para os indivíduos, como também para autoridades reguladoras e supervisoras¹⁰⁵, especialmente nas situações que afetam direitos e interesses dos administrados.

A opacidade algorítmica pode decorrer, basicamente, de três causas: i) proteção à propriedade intelectual e do segredo de negócio, no caso de sistemas proprietários; ii) necessidade de possuir conhecimentos técnicos específicos (matemáticos, programação etc.) para compreender o funcionamento do sistema; e, por fim, iii) aspectos técnicos relativos a modelos computacionais que possuem uma opacidade intrínseca à sua natureza, não sendo possível obter explicação completa sobre o funcionamento do modelo ou de todas as decisões e previsões, sendo ininteligível inclusive para os próprios desenvolvedores e especialistas¹⁰⁶.

A falta de fundamentação das decisões automatizadas viola a necessidade de motivação das decisões administrativas¹⁰⁷ e prejudica o exercício do contraditório e ampla defesa, visto que a pessoa afetada não vislumbra claramente as razões de decidir da Administração, restando inviabilizada a possibilidade de desafiar a referida decisão mediante recurso. Neste sentido, conforme o enunciado 12 exarado na I Jornada de Direito

¹⁰⁵ DONEDA, Danilo Cesar Maganhoto; SOUZA, Carlos Affonso Pereira de Souza et al. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. *Pensar*, Fortaleza, v. 23, n. 4, p. 1-17, dez, 2018.

¹⁰⁶ Comissão de Juristas Responsável por Subsidiar Elaboração de Substitutivo sobre Inteligência Artificial no Brasil. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 2002. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04>. Acesso em: 8 mar. 2023, p. 114.

¹⁰⁷ SADDY, André. *Curso de direito administrativo brasileiro*. 2. ed. Rio de Janeiro: CEEJ, 2023, v.1, p. 384.

Administrativo do Centro de Estudos Judiciários do Conselho da Justiça (CEJ/CJF) ¹⁰⁸, “a decisão administrativa robótica deve ser suficientemente motivada, sendo a sua opacidade motivo de invalidação”.

2.2.3 Substituição da subjetividade do gestor pela “subjetividade” do robô

O ministro Gilmar Mendes¹⁰⁹, no julgamento da ADI 6389 MC-REF/DF, alertou para o fato de que decisões estatais críticas estão sendo substituídas por mecanismos automatizados: “Vivemos na era das escolhas de Sofia automatizadas”. Para ilustrar essa ideia, o ministro apresenta os seguintes casos:

Alguns exemplos nesse sentido são dignos de nota. Nos Estados Unidos, por exemplo, uma ferramenta de gerenciamento automatizado do sistema prisional chamada de Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) tem sido utilizada para avaliação do risco de reincidência dos egressos. Essa ferramenta funciona a partir de árvore decisória, que classifica os detentos em um espectro de risco que varia de um a nove, sendo nove o mais alto e um o mais baixo. Em 2017, a Suprema Corte de Wisconsin manteve a condenação de um réu que foi acusado de fugir da polícia ao dirigir um carro anteriormente utilizado em um tiroteio. Ele havia sido condenado previamente por agressão sexual e, após uma avaliação do algoritmo, considerou-se que havia alto risco de cometer outro crime, tendo sido condenado, assim, a uma sentença de seis anos.

O uso da inteligência artificial tem avançado nos tempos atuais de pandemia inclusive para a tomada de decisões relevantíssimas como a alocação de leitos de UTIs, conforme explorado em detalhes em recente artigo da autora Marcela Matiuzzo publicado na Folha de São Paulo: ‘Nos Estados Unidos, por exemplo, os algoritmos já estão sendo utilizados para substituir a discricionariedade dos médicos na escolha sobre qual paciente deve ou não ser internado. Essas decisões são feitas com base em critérios utilitaristas pretensamente neutros como os pacientes com maior chance de sobrevivência e de qualidade de vida pós-tratamento. No Brasil, o Dr. Daniel Forte, presidente do comitê de bioética do Hospital Sírio-Libanês, explicou recentemente que se cogita da utilização desses algoritmos no Brasil para afastar alegações de obscuridade e enviesamento. O Conselho Federal de Medicina e a Associação Brasileira de Medicina Intensiva tem buscado estabelecer critérios para classificação de pacientes em graus de prioridade.

¹⁰⁸ I Jornada de Direito Administrativo - Enunciados aprovados. Disponível em: https://www.cjf.jus.br/cjf/corregedoria-da-justica-federal/centro-de-estudos-judiciarios-1/publicacoes-1/cjf/corregedoria-da-justica-federal/centro-de-estudos-judiciarios-1/publicacoes-1/Jornada%20de%20Direito%20Administrativo%20-%20Enunciados%20aprovados/?_authenticator=f147b8888b42ee73c25f9f3ea6258093fadd0b5a. Acesso em: 02 mar. 2023.

¹⁰⁹ BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº 6387. Ementa. Medida Cautelar em Ação Direta de Inconstitucionalidade. Referendo. Medida Provisória Nº 954/2020. Emergência de Saúde Pública de Importância Internacional Decorrente do Novo Coronavírus (covid-19). Compartilhamento de Dados dos usuários do serviço telefônico fixo comutado e do serviço móvel pessoal, pelas empresas prestadoras, com o Instituto Brasileiro de Geografia e Estatística. Fumus Boni Juris. Periculum in Mora. Deferimento. Recorrente: Conselho Federal da Ordem dos Advogados do Brasil - CFOAB. Intimado: Presidente da República. Relator: min. Rosa Weber. Julgamento em 6 e 7.5.2020. DJE nº 137, divulgado em 02/06/2020. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=754>.

O grande risco para o Estado de Direito é a completa substituição da subjetividade do gestor pela subjetividade da máquina, sob a pretensa alegação de se afastar particularismos em uma decisão. Não se ouvida, sem embargo, do Princípio da máxima objetividade, segundo o qual a Administração, sempre que possível, deve objetivar sua conduta¹¹⁰. Neste caso, a capacidade de identificar padrões de técnicas de Aprendizagem de Máquina poderia ser utilizada para garantir a conformidade a esse princípio. Sem embargo, a utilização da IA poderá ser prejudicial nos casos de conflito entre igualdade material e igualdade formal, onde a adoção de decisões padronizadas e repetitivas pode levar ao cometimento de uma injustiça em função das peculiaridades do caso concreto. Sobre esse ponto, ressaltam André Saddy e João Victor Tavares Galil¹¹¹:

Com essa técnica, é inegável que prevaleça a possibilidade de que casos sejam decididos de maneira uniforme, com base em características iguais, que passarão a constituir a realidade do mundo para a máquina, e muitas vezes não alcançará o melhor nível de justiça individualmente considerada. Eis um conflito principiológico que é, e sempre será, o pano de fundo da atividade administrativa em abstrato e que, no cenário da inteligência artificial, apresenta-se ainda mais marcante: o confronto, base de ponderação, entre segurança jurídica, própria da igualdade formal, e igualdade material, [...]

Por conseguinte, o servidor público teria ainda um ônus argumentativo maior para o caso de decidir de forma contrária - ainda que justa - ao recomendado pela IA. Há, portanto, uma evidente redução (com risco de total eliminação) na discricionariedade administrativa, em razão da natural acomodação ao automatismo produzido pelo sistema, bem como pela insegurança de se alterar ou ir contra a recomendação algorítmica, que poderia causar responsabilização da pessoa responsável pela decisão.

2.2.4 Vieses algorítmicos (*bias*)

Entende-se o viés como “*inclination or prejudice of a decision made by an AI system which is for or against one person or group, especially in a way considered to be unfair*”¹¹².

¹¹⁰ SADDY, André. Curso de direito administrativo brasileiro. 2. ed. Rio de Janeiro: CEEJ, 2023, v. 1, p. 406.

¹¹¹ SADDY, André. GALIL, João Victor Tavares. O processo de tomada de decisão administrativa e o uso da inteligência artificial. In: SADDY, André et al. Inteligência Artificial e Direito Administrativo. Rio de Janeiro: Centro para Estudos Empírico-Jurídicos (CEEJ), 2022, p. 119.

¹¹² Em tradução livre “*inclinação ou preconceito manifesto por uma decisão materializada por um sistema de IA que se apresenta como favorável ou contrária a uma pessoa ou grupo, especialmente de modo a traduzir uma injustiça*”. In: NTOUSI, Eirini et al. Bias in data-driven artificial intelligence systems – an introductory survey. WIREs Data Mining and Knowledge Discovery, v. 10, n. 3, p. 1-14, 2020. Disponível em: <https://wires.onlinelibrary.wiley.com/doi/epdf/10.1002/widm.1356>. Acesso em: 06 fev. 2023.

Algoritmos de computação cognitiva podem produzir resultados injustamente desiguais em razão de refletir preconceitos existentes na sociedade, principalmente no caso de tratamento de dados sensíveis. É preciso ter em conta que os sistemas são criados e desenvolvidos por pessoas, que podem refletir seus vieses e preconceitos em variáveis inseridas no código de programação. Além disso, as bases de dados utilizadas para treinar e alimentar esses sistemas também podem ser tendenciosas.

Um caso que classicamente se aponta como resultante de viés algorítmico é o de Tay, a “robô adolescente” da Microsoft. O Chatbot Tay fora desenvolvido para interagir com internautas jovens, entre 18 e 24 anos. O problema é que Tay acabou sendo manipulada por internautas no Twitter, adquirindo uma personalidade extremamente agressiva e preconceituosa, com declarações chocantes como “- Hitler estava certo. Eu odeio judeus”¹¹³.

Os vieses algoritmos podem ser causados por diversos fatores, tais como:

- a) Vieses estatísticos: Causados pela insuficiência ou incorreção da amostra de dados trabalhada, não correspondendo à realidade que lhe serviria de fonte de aprendizado, ou sobre a qual uma eventual predição virá a incidir.
- b) Vieses sociais: São reflexos de comportamentos desviantes existentes na coletividade, ou seja, é a transposição para o sistema computacional de visões oblíquas pré-existentes nas práticas sociais e nas instituições, sejam elas explícitas, sejam ocultas ou inconscientes¹¹⁴.
- c) Vieses técnicos: São limitações presentes no sistema computacional que se revelam capazes de interferir no resultado, podendo incidir no hardware ou no software. São falhas tecnológicas e limitações inerentes à computação cognitiva, tal como fora abordado no tópico referente às modalidades de Inteligência Artificial, onde se apontou que existem problemas que não podem ser resolvidos independentemente do poder computacional ou da quantidade de dados utilizados para treinar o algoritmo.
- d) Vieses Emergentes: relacionados à superveniência de mudanças no cenário de fato e de direito, tais como mudanças normativas, ou mesmo câmbio de valores sociais, podendo reduzir ou prejudicar a acurácia das indicações algorítmicas¹¹⁵.

¹¹³ RODRIGUES, Marcelo. Microsoft explica episódio com chatbot racista e diz que Tay deve voltar. *Tecmundo*, 26 mar. 2016. Disponível em: <https://www.tecmundo.com.br/inteligencia-artificial/102835-microsoft-explica-episodio-chatbot-racista-diz-tay-deve-voltar.htm>. Acesso em: 6 fev. 2023.

¹¹⁴ VALLE, Vanice Regina Lírio do. Administração Pública, viés algorítmico e desvio de finalidade: existe conciliação possível? In: ZOCKUN, MAURÍCIO; GABARDO, EMERSON. Direito Administrativo e Inovação: crises e soluções. CURITIBA: Íthala, 2022, p. 507-508.

¹¹⁵ *Ibid.*

2.2.5 Falta de publicidade e impedimento ou dificuldade de acesso a dados públicos

A falta da adequada publicidade de que uma pessoa está interagindo com sistemas automatizados é uma situação que também pode ser objeto de questionamentos. Conforme será visto com mais detalhe adiante, o GDPR, por exemplo, estabelece como um direito da pessoa singular que seja devidamente informada sobre eventual interação com sistemas de IA.

Outro ponto de questionamento, por outro lado, diz respeito ao impedimento ou dificuldade de acesso a dados públicos, o que acaba prejudicando a precisão dos sistemas inteligentes. Conforme fora visto no tópico referente aos principais domínios de aplicação, pelo *Data Mining*, a IA pode realizar a extração e tratamento de dados em diferentes bases governamentais. A acurácia de sistemas de IA depende do volume de dados em que são treinados, tal como tratado na parte referente às formas de Aprendizado de Máquina.

Ocorre que, muitas vezes, dados considerados de acesso público, nos termos da Lei n.º 12.527/2011 (Lei de Acesso à Informação ou LAI)¹¹⁶, encontram-se em bases desatualizadas, são de difícil acesso ou sequer estão disponibilizados por órgãos públicos. No Acórdão 2593/2017, com enfoque na fiscalização dos processos de aquisições da Administração Pública Federal, operacionalizada com ferramentas informatizadas de prospecção, especialmente o sistema de Análises de Editais e Licitações (Alice), o Tribunal de Contas da União manifesta preocupação em relação à disponibilidade de dados públicos em formatos que facilitem o trabalho de tratamento realizado por IA. Nas palavras do voto do ministro revisor, Aroldo Cedraz¹¹⁷:

Sobre o tema, comungo da preocupação externada pelo eminente revisor, pois a disponibilização de dados desatualizados prejudica o controle e se contrapõe à inteligência da Lei de Acesso à Informação, a qual possui dentre suas diretrizes a divulgação de informações de interesse público e o desenvolvimento do controle social da administração pública. É oportuna, então, a proposta de ser exarada determinação ao Ministério do Planejamento visando o aprimoramento dos citados repositórios.

Desta forma, determino, com fulcro no art. 8º da Lei 12.527/2011 c/c art.8º, § 1º, do Decreto 8.777/2016, que se mantenha atualizado o repositório de informações sobre as contratações públicas no portal dados abertos do Governo Federal; Além disso, determine que se depure a base de dados do painel de preços (<http://paineldeprecos.planejamento.gov.br>) e, concomitantemente, crie mecanismos

¹¹⁶ BRASIL. Lei n.º 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei n.º 8.112, de 11 de dezembro de 1990; revoga a Lei n.º 11.111, de 5 de maio de 2005, e dispositivos da Lei n.º 8.159, de 8 de janeiro de 1991; e dá outras providências. [S. l.], 18 nov. 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: 7 fev. 2023.

¹¹⁷ BRASIL. Tribunal de Contas da União. Acórdão 2593/2017 - Plenário, 22 nov. 2017. Disponível em: <https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/AC%25C3%2593RD%25C3%2583O%25202593%252F2017/%2520DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0/%2520>. Acesso em: 25 jun. 2022.

para padronizar os dados nele constantes e a inserção de novas informações, de forma minimizar as divergências observadas pela má-alimentação desse sistema de informação e facilitar a comparação de preços praticados no âmbito da administração-pública.

Portanto, a falta de disponibilização ou atualização adequada de informações públicas, em formatos que facilitem o uso de IA nas atividades de controle, seja ele exercido por órgãos e entidades vinculados ao estado (por exemplo, CGU e TCU), seja por entidades da sociedade civil, prejudica a sindicabilidade das atividades estatais em relação à economicidade, legalidade e combate à corrupção.

2.3 Solução: Ética como principal instrumento norteador da regulação de IA

Constata-se, assim, que nem tudo o que é tecnicamente possível é lícito do ponto de vista ético. O desenvolvimento técnico deve ser necessariamente acompanhado da formação e do progresso ético do homem, da resposta a questões de natureza moral ante os novos desafios tecnológicos que se nos apresentam, sob pena de se transformar em causa de destruição da humanidade e do mundo.

Por conseguinte, normas antropológicas são absolutamente necessárias para nortear os rumos do avanço tecnológico. Tal como o auriga é indispensável para conduzir os cavalos para a reta direção ao destino ou a uma finalidade, é fundamental que forças de regulamentação tomem as rédeas do desenvolvimento tecnológico, de forma a não colocar em perigo a civilização e as culturas. Nas palavras de Edgar Morin¹¹⁸:

Corremos assim para a autodestruição? Para uma mutação? Os feedbacks positivos que levam ao runaway podem eventualmente produzir uma mutação. Mas seria preciso que as forças de controle e de regulação tomassem a dianteira. Trata-se, portanto, de frear o avanço técnico sobre as culturas, a civilização, a natureza, que ameaça tanto as culturas como a civilização e a natureza. Trata-se de diminuir a marcha para evitar ou uma explosão, ou uma implosão. Trata-se de desacelerar para poder regular, controlar e preparar a mutação. A sobrevivência exige revolucionar o devir. Precisamos chegar a um outro futuro. Essa é que deve ser a tomada de consciência decisiva do novo milénio.

[...]

O desenvolvimento é uma finalidade, mas deve deixar de ser uma finalidade míope ou uma finalidade-término. A finalidade do desenvolvimento submete-se ela própria a outras finalidades. Quais? Viver verdadeiramente. Viver melhor. Verdadeiramente e melhor, o que significa isso? Viver com compreensão, solidariedade, compaixão. Viver sem ser explorado, insultado, desprezado. Significa que as finalidades do

¹¹⁸ MORIN, Edgar; KERN, Anne-Brigitte. Terra-Pátria. Porto Alegre: Sulina, 2003, p. 95.

desenvolvimento dependem de imperativos éticos. O económico deve ser controlado e finalizado por normas antro-po-éticas¹¹⁹. (destaques pessoal).

Platão, em Protágoras¹²⁰, a partir do mito da criação dos seres mortais, apresenta um ensinamento que nos é de grande valia: os eixos da vida em sociedade não devem se centrar no desenvolvimento tecnológico, mas no respeito que todos devem ter uns pelos outros e na justiça.

Segundo narra Platão¹²¹, os titãs Epimeteu e Prometeu foram incumbidos pelos deuses de distribuir, de forma equilibrada, dons e qualidades entre os seres mortais. Etimologicamente, Epimeteu significa aquele que pensa antes de fazer, enquanto Prometeu é aquele que pensa depois.

Epimeteu procurou distribuir os dons e recursos de forma que os animais tivessem capacidade de obter alimentos, não levassem a uma destruição recíproca e resistissem às intempéries. Atribuiu a alguns seres mortais a força, mas sem velocidade, enquanto a outros atribuiu velocidade, em que pese serem mais fracos. Forneceu a alguns armas de ataque e de defesa, a outros, atribuiu a capacidade de voar ou de esconder-se sob a terra.

Epimeteu, quando finalmente chegou no homem, apercebeu-se de que havia esgotado todos os recursos de que dispunha, deixando os seres humanos nus, descalços, desabrigados e inermes. Diante disso, Prometeu encarregado de fazer a revisão final do labor, constatou que o homem era o único animal desprovido do necessário para própria sobrevivência.

Para salvar a espécie humana, Prometeu realizou um ato ardiloso: roubou a sabedoria da técnica de Atena e o fogo de Hefesto e os entregou aos homens. De posse desses dons divinos, da técnica e do fogo, os homens foram capazes de desenvolver a linguagem, construir casas, confeccionar vestes e calçados, desenvolver a agricultura, medicina, metalurgia etc.

Ocorre que Prometeu não conseguiu roubar a sabedoria política, visto que era guardada por Zeus, cuja morada era vigiada por terríveis guardas. Por isso, os homens espalharam-se por todas as partes e não existiam cidades. Estavam à mercê das feras, já que eram muito menos fortes do que elas. A arte que possuíam permitia que obtivessem alimentos, mas não era suficiente na guerra contra os outros animais. Por isso, procuravam ficar juntos, mas assim que fundavam cidades, cometiam injustiças uns para com os outros, instaurando-se um caos que fatalmente os levaria à autodestruição.

De forma a garantir a sobrevivência da espécie humana, Zeus determinou então que os dons do respeito (*aidos*) e da justiça (*diké*) fossem igualmente distribuídos a todos os homens,

¹¹⁹ Ibid, p. 106.

¹²⁰ PLATÃO. *Protágoras*. Tradução Carlos Alberto Nunes. Pará: Universidade Federal do Pará, 1980.

¹²¹ Ibid.

para que fossem princípios organizadores das cidades e vínculos produtores de amizade. Por fim, determinou Zeus¹²²: “-Aliás, em meu nome, estabeleça como lei que aquele que não sabe compartilhar o respeito e a justiça seja morto como um mal da Cidade”.

Platão, nos ensina, a partir da referida passagem mitológica, que não são suficientes as refinadas artes e o fogo para o progresso e desenvolvimento da espécie humana, ou seja, as poderosas energias libertadas pela técnica e pelo trabalho. Se a humanidade não recupera os outros dons divinos – os de Zeus -, isto é, o respeito e a justiça, que o rei do Olimpo desejou que fossem dados a todos sem distinção, os homens estariam condenados à autodestruição e extinção.

Voltando o olhar para as novas tecnologias inteligentes Danilo Doneda et al¹²³ defendem a necessidade da “ética dos algoritmos”, especialmente nas situações em que ainda não possam ser efetivamente objeto de regulação legal, mas que, haja vista a relevância dos valores envolvidos, exigem uma resposta célere e adequada. No mesmo sentido, Eduardo Magrani¹²⁴ sustenta que o design de tecnologia deve ser centrado no ser humano e sensível a valores constitucionais, com ética, segurança e privacidade por meio do “design sensível a valores”.

No caso da Administração Pública, a crescente adoção de sistemas de IA, tal como constatado no tópico referente ao diagnóstico e perspectivas de sua utilização pelo Estado, torna ainda mais premente a necessidade de se buscar soluções jurídicas para que os agentes públicos atuem de forma moral e constitucionalmente orientados, visando o atendimento do interesse público e do bem comum, bem como respeitando direitos e garantias fundamentais. O TCU¹²⁵, por exemplo, reconheceu expressamente a necessidade de um *framework* que oriente a utilização ética e segura nas atividades de controle. Para esse desiderato, é mister compreender a relação entre ética e direito, e como se materializa o dever de agir de forma ética na

¹²² Ibid. p. 322.

¹²³ DONEDA, Danilo Cesar Maganhoto; SOUZA, Carlos Affonso Pereira de Souza et al. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. *Pensar*, Fortaleza, v. 23, n. 4, p. 1-17, dez, 2018, p. 11.

¹²⁴ MAGRANI, Eduardo. *Entre dados e robôs*. 2. ed. Porto Alegre: Arquipélago Editorial, 2019.

¹²⁵ “Se por um lado a inteligência artificial tem o potencial de aumentar os ganhos em eficiência, melhorar serviços e reduzir custos, por outro, ela também apresenta novos desafios e riscos. Tais desafios envolvem questões como a segurança de dados, a possibilidade de tratamento automatizado desigual e a produção em massa de decisões institucionais incorretas e/ou discriminatórias. Nesse contexto, é necessário o uso de controles pelos gestores públicos para garantir a confiabilidade dos sistemas de IA, assim como é necessário que os órgãos de controle auditem as aplicações baseadas em algoritmos de IA. Dessa forma, várias entidades que atuam no segmento de controle já desenvolveram propostas e frameworks que visam endereçar as questões relativas à assecuração da utilização ética e segura da IA.”. Cf. BRASIL. Tribunal de Contas da União. Acórdão nº 1139/2022 – Plenário, de 25 de maio de 2022. Disponível em: https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/*/NUMACORDAO%253A1139%2520ANOACORDAO%253A2022/DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0/%2520. Acesso em: 25 jun. 2022.

Administração Pública, que, conforme será tratado nos tópicos seguintes, materializa-se no princípio da moralidade administrativa.

CAPÍTULO III – DIREITO, ÉTICA E O PRINCÍPIO DA MORALIDADE ADMINISTRATIVA

Ainda que não se pretenda realizar um estudo filosófico aprofundado sobre a relação entre direito e moral, tendo em vista que a presente dissertação tem como enfoque a aplicação da moralidade administrativa como fundamento para regulação de IA, é importante que haja uma breve reflexão sobre a relação entre moralidade e direito, bem como sobre o marco teórico adotado no presente trabalho, em razão das consequências práticas decorrentes desta definição para o desenvolvimento das análises e soluções subsequentes

A título ilustrativo, apenas para se ter uma ideia das repercussões decorrentes desta relação, Eduardo Magrani comenta a respeito de um caso de adoção de uma perspectiva ética utilitária e econômica pela Ford Motor Company para justificar a decisão empresarial de manter a comercialização do carro Ford Pinto, ainda que consciente de um defeito de segurança que causava incêndio no tanque de combustível. Em um memorando de 1973, foi realizada uma análise de custo-benefício, comparando o custo de recall (US\$ 137,5 milhões) ao custo relacionado a lesões e mortes relacionadas com incêndios nos veículos vendidos nos Estados Unidos (US\$ 49,5 milhões), decidindo-se por manter a comercialização do veículo sem recall. Quando veio à tona, a decisão foi extremamente criticada, tendo a National Highway Traffic Safety Administration determinado o recall, em que pese a empresa ter decidido por fazê-lo antes da referida decisão¹²⁶.

Em seguida, far-se-á uma análise do princípio da moralidade administrativa sobre o prisma da doutrina constitucional e administrativa.

3.1 A relação entre moral e direito

Neste tópico tratar-se-á da relação entre direito e moral (ou ética, aqui considerados como sinônimos), entre a ordem jurídica e a ordem moral. Questiona-se, com efeito, se o direito pode ser considerado como parte da moral. Tal como ressalta Miguel Reale¹²⁷, trata-se de um dos “problemas mais difíceis e também dos mais belos da Filosofia Jurídica”. Há

¹²⁶ MAGRANI, Eduardo. Entre dados e robôs. 2. ed. Porto Alegre: Arquipélago Editorial, 2019, p. 138-139

¹²⁷ REALE, Miguel. Lições preliminares de direito. 27. ed. São Paulo: Saraiva, 2002, p. 44.

inúmeras escolas de pensamento relacionadas ao tema, que variam entre os extremos do positivismo jurídico, que propõe uma separação total e completa entre direito e moral, cujo maior representante é Hans Kelsen, e escolas que defendem que o direito seria parte da moral, como a teoria do “mínimo ético”, a que se filiam autores como Jeremias Bentham e Georg Jellinek¹²⁸.

No presente trabalho, adotar-se-á como marco teórico os pensadores que seguem uma abordagem aristotélica sobre o tema, especialmente Javier Hervada¹²⁹. A razão para adoção desta linha de pensamento está relacionada ao fato de Aristóteles ter construído uma consistente, sistemática e harmônica teoria ética e de justiça, especialmente trazidas em *Ética a Nicômaco*¹³⁰, e cujo desenvolvimento perpassou pelos séculos, fundamentando sistemas jurídicos como o direito romano e outros ordenamentos do mundo ocidental, sendo a reflexão sobre a sua teoria continuada até os dias atuais¹³¹.

Feitos os esclarecimentos sobre a perspectiva teórica do tema, é mister ressaltar que a realidade moral do homem pode ser entendida como aquela pertencente à dimensão da vida individual e em sociedade, decorrente do exercício da liberdade da pessoa humana, dotada de autodeterminação e consciência. A realidade moral foi chamada de “*éthos*” na Grécia, de onde vem a palavra “ética”. Em Roma, foi designada como aquilo que é próprio do costume ou “*mors*”, de onde surgiu a palavra “moral”. A perspectiva moral não engloba esferas do ser regidas pelas leis físicas e biológicas, que são objeto das ciências exatas ou experimentais, tais como a biologia e a medicina, por exemplo¹³².

O direito, por sua vez, disciplina a conduta humana enquanto externalizada no meio social, conduta esta decorrente da liberdade do indivíduo que realiza um ato ou omissão. Desta forma, no plano jurídico, pressupõe-se o exercício do livre arbítrio para que possam advir consequências jurídicas ao agente. No direito penal, por exemplo, a configuração de um crime decorre de uma ação ou omissão livre e consciente da pessoa, não gerando responsabilização pelo ato de um demente, animal ou ser inanimado. O Código Penal, por exemplo, prevê no art.

¹²⁸ Pela teoria do mínimo ético, “o Direito representa apenas o mínimo de Moral declarado obrigatório para que a sociedade possa sobreviver”, visto que, em que pese, em geral, a moral ser cumprida de forma espontânea, não é possível garantir que todos os indivíduos atuem sempre de forma correta. Por esta razão, é necessário estipular um núcleo de condutas consideradas indispensáveis à paz social. A referida teoria “pode ser reproduzida através da imagem de dois círculos concêntricos, sendo o círculo maior o da Moral, e o círculo menor o do Direito”, e, portanto, “tudo o que é jurídico é moral, mas nem tudo o que é moral é jurídico”. In: REALE, 2002, p. 44

¹²⁹ HERVADA, Javier. *Lições Propedêuticas de Filosofia do Direito*. São Paulo: Martins Fontes, 2008.

¹³⁰ ARISTÓTELES. *Ética a Nicômaco*. Tradução de Mário de Gama Kury. 4. ed. Brasília: Editora Universidade de Brasília, 2001.

¹³¹ Cf. BONALDO, Frederico et al. *Direito Natural Contemporâneo: A Renascença do Jusnaturalismo no Brasil e no Mundo*. 1. ed. Londrina: Thoth Editora, 2021.

¹³² HERVADA, Javier. *Lições Propedêuticas de Filosofia do Direito*. São Paulo: Martins Fontes, 2008.

26 que é isento de pena o agente que, por doença mental ou desenvolvimento mental incompleto, é inteiramente incapaz de entender o caráter ilícito do fato ou de determinar-se de acordo com esse entendimento¹³³. Por conseguinte, pode-se afirmar que as leis pertencem à realidade moral do homem.

Por outro lado, é preciso distinguir a ciência do direito da ciência moral a partir da perspectiva específica de cada uma delas em relação à realidade moral, que é seu objeto material. Desta forma, um mesmo ato humano poderá ter uma apreciação que variará em função do ponto de vista, ou seja, se o seu objeto formal for a ciência moral ou a ciência jurídica, “pois cada ciência só nos informa de um aspecto da coisa”¹³⁴.

A ciência moral, ou filosofia moral determina os princípios e regras comuns dos atos humanos, uma vez que os vê em seu aspecto moral mais fundamental. Todo ato humano será objeto da ciência moral, inclusive os atos com repercussão no direito¹³⁵. A moralidade de um ato é avaliada segundo aspectos anímicos ou interiores do indivíduo, tais como a intencionalidade com que um ato é praticado. Inclusive, alguém pode praticar um ato moralmente condenável, ainda que sem qualquer repercussão no mundo exterior, por exemplo, quando, tomado por um sentimento de ira, uma pessoa imagina-se matando o seu desafeto de diversas maneiras. Este pensamento é condenável do ponto de vista moral, mas não terá qualquer repercussão jurídica, caso o indivíduo não tenha ultrapassado a fase da cogitação do *iter criminis*.

A ciência jurídica, por outro lado, não estuda a realidade moral em todas as suas dimensões, mas segundo a perspectiva própria do direito, sendo mais limitada que a ciência moral em relação ao seu objeto material, já que o seu estudo se circunscreve aos atos humanos com repercussão jurídica. O Direito nem sequer estuda a virtude da justiça em sua integridade, pois limita-se à repercussão externa da justiça, sem adentrar necessariamente na intenção do indivíduo ao realizar um ato socialmente positivo, como o pagamento de tributos, por exemplo. Neste caso, alguém pode pagar tributos por medo das sanções decorrentes do inadimplemento, o que não seria meritório do ponto de vista moral¹³⁶.

Para o direito, não interessa a intencionalidade do indivíduo, desde que o tributo seja pago. Seu objeto não consiste em que o homem seja justo e virtuoso do ponto de vista moral, mas em que seja respeitado o direito de cada indivíduo, tal como manifestado exteriormente.

¹³³ BRASIL. DECRETO-LEI nº 2.848, de 7 de dezembro de 1940. Código Penal. [S. l.], 1940. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em: 7 dez. 2021.

¹³⁴ HERVADA, Javier. Lições Propedêuticas de Filosofia do Direito. São Paulo: Martins Fontes, 2008, p. 286.

¹³⁵ Ibid.

¹³⁶ Ibid, p. 287.

Por conseguinte, o direito é uma ciência de determinadas relações sociais, não podendo ser considerada, por esta razão, como parte da ciência moral, visto que o jurista não é um moralista. Ressalta Javier Hervada¹³⁷:

Ambos, juristas e moralistas, estudam a justiça, porém suas perspectivas são diferentes: interessa ao moralista, por exemplo, que os empréstimos não sejam usurários, para que os homens se comportem como pessoas e não cometam pecado (ofensa a Deus por violar sua lei); interessa ao jurista a mesma coisa quanto aos empréstimos, mas por outro motivo: para que sejam respeitados os direitos de cada um, para que cada qual receba o que lhe cabe e haja assim uma ordem social justa.

A ética estuda a consciência individual, enquanto o direito trata das externalidades refletidas nas relações sociais entre as pessoas. Daí a razão da máxima “*ubi societas, ibi jus*” referida por Ulpiano como lembrado no Corpus Iuris Civilis. Por conseguinte, o direito não pode ser considerado parte da moral, por se tratarem de ciências autônomas, em uma perspectiva diferente em relação ao agir humano.

Sem embargo, o direito pode receber da moral dados ou influxos para resolver situações que dependem de uma avaliação de ordem moral, tal como o princípio doutrinário: “*nemo auditur turpitudinem allegans*”¹³⁸. No direito administrativo, por sua vez, adota-se o princípio da moralidade administrativa, que deve nortear todos os atos dos agentes públicos, ideia que será desenvolvida no próximo tópico.

Por outro lado, as leis, afora aspectos técnicos e organizativos, podem ter repercussão moral na medida que induzem o comportamento humano, ao ponto de gerarem um hábito ou virtude, aspectos esses atinentes à ciência moral. Neste sentido, explica Javier Hervada¹³⁹:

e) Lei e comportamento moral: Quando se age de acordo com as leis, elas criam hábitos e costumes. Por causa desse efeito, não se circunscrevem a fazer bons cidadãos do ponto de vista da conduta externa; também influenciam a moralidade do homem, ao contribuir para formar virtudes. Como a maioria das virtudes não são inatas, mas adquiridas pela repetição de atos, as leis, compelindo a agir segundo uma virtude, acabam conseguindo que quem as obedece adquira as virtudes correspondentes. O motorista que cumpre o Código de Trânsito acaba possuindo o hábito de dirigir com prudência; todos temos experiência de que, por cumprir sempre as leis, chega-se a fazer por costume — por virtude — muitas das coisas que mandam sem nos lembrarmos da lei. Eis um importante aspecto das relações entre a moral e as leis. As leis não são indiferentes em relação à formação e ao comportamento morais do homem; pelo contrário, influem neles intensamente, contribuindo de modo notável para moralizar os costumes (ou para favorecer a imoralidade, caso das leis

¹³⁷ HERVADA, Javier. Lições Propedêuticas de Filosofia do Direito. São Paulo: Martins Fontes, 2008, p. 287.

¹³⁸ Em tradução livre, “ninguém pode ser ouvido ao alegar a própria torpeza”.

¹³⁹ Ibid, p. 289.

permissivas, imorais ou injustas).

Por conseguinte, as leis não são neutras em relação à moral. Por esta razão, a construção de um ordenamento jurídico neutro ou amoral, como pretendeu a escola do positivismo jurídico, sob o argumento de criar uma “ciência pura”, pavimentou a construção de estados totalitários, como a Alemanha nazista. Ao ingressar no comportamento humano, a lei acaba por afetar a ordem humana da liberdade, e, por conseguinte, na moralidade, sendo, portanto, realidades indissociáveis.

Por fim, ainda que as leis devam tender a formar bons cidadãos para realização do bem comum ou social, não devem impor um elevado padrão moral de comportamento, sob pena de invadirem o campo próprio da ciência moral, os recônditos da consciência. Tendo em vista o estado moral da sociedade, “as leis, portanto, podem não exigir as virtudes em toda sua força e até mesmo tolerar algumas condutas não-boas”¹⁴⁰.

3.2 Análise administrativo-constitucional do princípio da moralidade administrativa

A presente dissertação propõe, como hipótese solução para a problemática ética relativa à adoção de IA pela Administração Pública, a aplicação do princípio da moralidade administrativa como fundamento e fonte da regulamentação para nortear a criação, desenvolvimento e uso de IA pelos órgãos e entes públicos.

Diante disso, é relevante compreender com maior profundidade os aspectos jurídicos de natureza administrativa e constitucional relativos ao princípio da moralidade administrativa. É o que será abordado nos tópicos seguintes.

3.2.1 Previsão constitucional do princípio da moralidade administrativa

O princípio da moralidade administrativa foi positivado e alçado ao *status* constitucional na CRFB¹⁴¹, sendo um dos pilares da Administração Pública. Está umbilicalmente ligado ao conceito de excelência e boa administração, ao elemento ético e à honestidade, à adoção de meios lícitos para realização da finalidade ou razão de ser da

¹⁴⁰ HERVADA, Javier. Lições *Propedêuticas de Filosofia do Direito*. São Paulo: Martins Fontes, 2008, p. 289.

¹⁴¹ “Art. 37. A administração pública direta e indireta de qualquer dos Poderes da União, dos Estados, do Distrito Federal e dos Municípios obedecerá aos princípios de legalidade, impessoalidade, moralidade, publicidade e eficiência e, também, ao seguinte”. Constituição da República Federativa do Brasil de 1988. Brasília, DF: Senado Federal, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 7 dez. 2021.

Administração Pública: o atendimento ao interesse público e a efetivação de direitos e garantias fundamentais. O comportamento dos agentes, órgãos e entes da Administração, em qualquer das suas formas de manifestação, deve se guiar pelos parâmetros e vetores ético-jurídicos que são decorrência do princípio constitucional da moralidade administrativa. Na esteira dos ensinamentos de Hely Lopes Meirelles¹⁴² e Antônio José Brandão¹⁴³, respectivamente:

O agente administrativo, como ser humano dotado de capacidade de atuar, deve, necessariamente, distinguir o Bem do Mal, o Honesto do Desonesto. E ao atuar, não poderá desprezar o elemento ético da sua conduta. Assim, não terá que decidir somente entre o legal e o ilegal, o justo do injusto, o conveniente e o inconveniente, o oportuno e o inoportuno, mas também entre o honesto e o desonesto.

É o órgão da pública Administração que, usando de sua competência para o preenchimento das atribuições legais, se determina não só pelos preceitos vigentes, mas também pela moral comum. Se os primeiros delimitam as fronteiras do lícito e do ilícito, do justo e do injusto positivos — a segunda espera 48ele conduta honesta, verdadeira, intrínseca e extrinsecamente conforme à função realizada por seu intermédio. Não será a ordem administrativa parte, ou aspecto, da ordem jurídica postulada pelo Bem-comum? E não será a ordem jurídica parte, ou aspecto, da ordem cósmica que se define mediante a liberdade humana? Como então recusar esta evidência: comete uma “imoralidade administrativa” o administrador que perturba a ordem administrativa com uma conduta determinada por fins concretos alheios à pública administração, muito embora não tenha cometido faltas legais?.

Sem embargo é importante esclarecer que a moralidade administrativa não se confunde com a moralidade comum. Trata-se, com efeito, de regras da boa administração, ou nas palavras de Welter¹⁴⁴ “pelo conjunto das regras finais e disciplinares suscitadas, não só pela distinção entre o Bem e o Mal, mas também pela idéia geral de administração e pela idéia de função administrativa”, onde, na visão de Lacharriere¹⁴⁵, incluem-se prescrições de caráter técnico. Desta forma, a honestidade profissional dos administradores, “reclama destes não só uma atividade moralmente admissível, mas também respeitadora da especialização institucional de cada organização ou serviço”¹⁴⁶.

A ética no agir do servidor público, quando combinada ao princípio da eficiência, acrescentado ao *caput* do art. 37 da CRFB pela Emenda Constitucional nº 19, de 1998, enseja a necessidade de que o gestor público atue com excelência, elegendo os meios mais eficazes e vantajosos para realização dos fins da Administração Pública, adotando-se a melhor técnica

¹⁴² MEIRELLES, Hely Lopes. Direito Administrativo Brasileiro. São Paulo: Ed. Medeiros, 2012. p. 90.

¹⁴³ BRANDÃO, Antônio José. Moralidade administrativa. Revista de Direito Administrativo, Rio de Janeiro, v. 25, n. 99, p. 454-467, 1951.

¹⁴⁴ *Apud* BRANDÃO, 1951, p. 458-459.

¹⁴⁵ *Apud* BRANDÃO, 1951, p. 460-461.

¹⁴⁶ *Ibid.*

moralmente lícita. Por outro lado, a busca por obter o máximo de eficiência administrativa, quando desvinculada da moralidade, pode resultar em ações imorais. Nas palavras de Antônio José Brandão¹⁴⁷:

[...] a atividade dos administradores, além de traduzir a vontade de obter o máximo de eficiência administrativa, terá ainda de corresponder à vontade constante de viver honestamente, de não prejudicar outrem e de a cada um dar o que lhe pertence – princípios de direito natural já lapidarmente formulados pelos juriconsultos romanos.

A partir da promulgação da CRFB, a observância da moralidade administrativa passou a ser um postulado de legitimação constitucional dos atos emanados pelo Estado, fundados, ou não, em competência discricionária, ou seja, independente da subjetividade ou autonomia pública que constitui no meio pelo qual o agente manifesta a sua vontade e concretiza as três funções (executiva, legislativa ou judiciária), tal como ministram Manoel de Oliveira Franco Sobrinho, Alexandre de Moraes, Lúcia Valle Figueiredo, Celso Antônio Bandeira De Mello, Hely Lopes Meirelles e Maria Sylvia Zanella Di Pietro¹⁴⁸. Os atos da Administração, em todas as relações, inclusive entre órgãos, entes e particulares, e mesmo naqueles praticados por particulares que venham a colaborar com o Estado, devem conformar-se à moralidade administrativa¹⁴⁹.

Não basta, assim, que um ato administrativo esteja subsumido à lei do ponto de vista jurídico-formal: é preciso também que seja praticado segundo os ditames da licitude moral. Como consequência disso, toda e qualquer imoralidade administrativa será necessariamente ilegal. Neste sentido, ministra André Saddy¹⁵⁰:

[...] no que diz respeito à moralidade administrativa, toda imoralidade administrativa é ilegal, pelo simples fato de ela estar prevista na Constituição e em inúmeras leis infraconstitucionais. Não cabe dizer que existe uma atuação formalmente legal, mas materialmente imoral. Tal será, também, ilegal, pois a moralidade administrativa está prevista na lei.

Em outras palavras, o princípio da moralidade administrativa, conquanto tenha relação intrínseca com o princípio da legalidade administrativa, é um princípio autônomo que não se confunde com este último, mas necessariamente o afetará, pois, como mencionado, toda atuação que viole a moralidade administrativa será ilegal, porém nem toda atuação ilegal será imoral.

Os atos de imoralidade administrativa, assim, podem ser considerados como “anverso

¹⁴⁷ Ibid.

¹⁴⁸ *Apud* BRASIL. Supremo Tribunal Federal. ADI 2661 MC. Relator: Celso de Mello. Tribunal Pleno. Julgado em 05/06/2002. DJ 23-08-2002.

¹⁴⁹ SADDY, André. Curso de direito administrativo brasileiro. 2. ed. Rio de Janeiro: CEEJ, 2023, v.1, p 387

¹⁵⁰ Ibid, p. 390

da moeda” e, portanto, passíveis de controle jurisdicional. Foi Maurice Hauriou¹⁵¹ quem pela vez tratou do tema, desenvolvendo a tese de que a conformidade dos atos jurídicos administrativos “aos princípios basilares da ‘boa administração’”, determinante necessária de qualquer decisão administrativa, é fiscalizada por outro recurso, fundado no desvio de poder, cuja zona de policiamento é a zona da “moralidade administrativa”.

Por conseguinte, a imoralidade administrativa caracteriza-se por um desvio de poder, levando à escolha pelo agente de fim metajurídico irregular. Na visão de Antônio José Brandão¹⁵², o desvio de poder está para o direito público assim como o abuso do direito está para o direito privado. Ambos os casos se caracterizam por uma espécie de juízo de caducidade de um poder (ou direito) que tenha sido imoralmente exercido, de moralidade no exercício de poderes (ou direitos). E conclui o referido autor¹⁵³: “Quis-se dar ao ato objeto lícito, mas para atingir fim ilícito – alheio, portanto, aos concretos interesses da pública administração que ao órgão competia satisfazer? Nesse caso, cometeu-se desvio de poder”.

Esta é a razão pela qual o controle jurisdicional pode abarcar a moralidade do ato administrativo, condição de validade da conduta do administrador público, caracterizando desvio de poder o ato que desatende à ética e aos fins administrativos¹⁵⁴. Tal como ficara consagrado na jurisprudência do STF:

O princípio da moralidade administrativa – enquanto valor constitucional revestido de caráter ético-jurídico – condiciona a legitimidade e a validade dos atos estatais. A atividade estatal, qualquer que seja o domínio institucional de sua incidência, está necessariamente subordinada à observância de parâmetros ético-jurídicos que se refletem na consagração constitucional do princípio da moralidade administrativa. Esse postulado fundamental, que rege a atuação do poder público, confere substância e dá expressão a uma pauta de valores éticos sobre os quais se funda a ordem positiva do Estado¹⁵⁵.

[...] É por essa razão que o princípio constitucional da moralidade administrativa, ao impor limitações ao exercício do poder estatal, legitima o controle externo de todos os atos, quer os emanados do Poder Público, quer aqueles praticados por particulares que venham a colaborar com o Estado na condição de licitantes ou contratados e que transgridam os valores éticos que devem pautar o comportamento dos órgãos e

¹⁵¹ *Apud* BRANDÃO, 1951, p. 457.

¹⁵² BRANDÃO, Antônio José. Moralidade administrativa. *Revista de Direito Administrativo*, Rio de Janeiro, v. 25, n. 99, p. 454-467, 1951, p. 464.

¹⁵³ *Ibid.*, p. 465.

¹⁵⁴ Além de atender à legalidade, o ato do administrador deve conformar-se com a moralidade e finalidade: “o controle jurisdicional se restringe ao exame da legalidade do ato administrativo; mas por legalidade ou legitimidade se entende não só a conformação do ato com a lei, como também com a moral administrativa e com o interesse coletivo”. TJSP, RDA 89/134, sendo o acórdão de lavra do Des. Cardoso Rolim) *apud* MEIRELLES, 1997, p. 84-85.

¹⁵⁵ BRASIL. Supremo Tribunal Federal. ADI 2.661 MC, Relator: Min. Celso de Mello. Julgado em 5-6-2002. Publicado no DJ de 23-8-2002.

agentes governamentais.¹⁵⁶

O princípio da moralidade encontra, ainda, guarida constitucional na ação popular, garantia prevista no art. 5º, LXXIII da CRFB¹⁵⁷, que resguarda o direito de qualquer cidadão manejar a ação popular para anular ato lesivo à moralidade administrativa. Trata-se de medida que sobremaneira facilita o controle social da moralidade dos atos administrativos, visto que a competência para julgar ação popular contra ato de qualquer autoridade, até mesmo do presidente da República, é, via de regra, do juízo de primeiro grau. Além disso, o autor da ação popular não pode ele ser condenado nos ônus das custas e da sucumbência, a não ser quando há comprovação de má fé¹⁵⁸. Por exemplo, a partir de caso levado ao Judiciário via ação popular, o STF¹⁵⁹ considerou que a fixação, por vereadores, de remuneração para viger na própria legislatura, ofende à moralidade administrativa, que constitui patrimônio moral da sociedade.

A improbidade administrativa, por sua vez, poderia ser considerada como sendo um ato atentatório ao princípio da moralidade administrativa. Sem embargo, é preciso compreender que não necessariamente há essa correlação de forma direta, visto que o ordenamento jurídico brasileiro positivou um significado próprio para os atos de improbidade.

A CRFB estabelece que os atos de improbidade administrativa importarão na suspensão dos direitos políticos, a perda da função pública, a indisponibilidade dos bens e o ressarcimento ao erário, sem prejuízo da ação penal cabível, consoante disposto no § 4º do art. 37¹⁶⁰. Pela referida norma constitucional, são previstos tratamentos sancionatórios diferenciados aos ilícitos civis, penais e político-administrativos.

A Lei n.º 8.429/1992 positivou as condutas consideradas como ímprobadas, tendo sido

¹⁵⁶ BRASIL. Supremo Tribunal Federal. MS 32.494 MC, Relator: Min. Celso de Mello. Julgado em 11-11-2013, decisão monocrática. Publicado no DJE de 13-11-2013.

¹⁵⁷ art. 5º, LXXIII que "*qualquer cidadão é parte legítima para propor ação popular que vise a anular ato lesivo ao patrimônio público ou de entidade de que o Estado participe, à moralidade administrativa, ao meio ambiente e ao patrimônio histórico e cultural, ficando o autor, salvo comprovada má-fé, isento de custas judiciais e do ônus da sucumbência*". In: BRASIL. Constituição (1988). Constituição da República Federativa do Brasil de 1988. Brasília, DF: Senado Federal, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 7 dez. 2021.

¹⁵⁸ BRASIL. Supremo Tribunal Federal. RE 221.291, Relator: Min. Moreira Alves. Julgado em 11-4-2000, 1ª Turma. Publicado no DJ de 9-6-2000.

¹⁵⁹ BRASIL. Supremo Tribunal Federal. RE 206.889, Relator: Min. Carlos Velloso. Julgado em 25-3-1997, 2ª Turma. Publicado no DJ de 13-6-1997.

¹⁶⁰ art. 37. § 4º: "*Os atos de improbidade administrativa importarão a suspensão dos direitos políticos, a perda da função pública, a indisponibilidade dos bens e o ressarcimento ao erário, na forma e gradação previstas em lei, sem prejuízo da ação penal cabível*". In: BRASIL. [Constituição (1988)]. Constituição da República Federativa do Brasil de 1988., 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 7 dez. 2021.

recentemente alterada pela Lei n.º 14.230/2021¹⁶¹, punindo-se mais severamente o agente público e todos aqueles que o auxiliem. Tal como assentado na jurisprudência do STF¹⁶², os atos de improbidade podem ser definidos como:

[...] possuindo natureza civil e devidamente tipificados em lei federal, ferem direta ou indiretamente os princípios constitucionais e legais da administração pública, independentemente de importarem enriquecimento ilícito ou de causarem prejuízo material ao erário; podendo ser praticados tanto por servidores públicos (improbidade própria), quanto por particular – pessoa física ou jurídica – que induzir, concorrer ou se beneficiar do ato (improbidade imprópria).

A jurisprudência do STF, em diversos julgados, reconheceu a autonomia e aplicabilidade imediata do princípio da moralidade administrativa, inclusive em relação à dispensa de sua regulamentação legal, visto que os princípios do art. 37, *caput*, da CRFB sequer precisam de lei formal para serem obrigatoriamente observados:

ADMINISTRAÇÃO PÚBLICA. VEDAÇÃO NEPOTISMO. NECESSIDADE DE LEI FORMAL. INEXIGIBILIDADE. PROIBIÇÃO QUE DECORRE DO ART. 37, CAPUT, DA CF. RE PROVIDO EM PARTE.

[...]

II – A vedação do nepotismo não exige a edição de lei formal para coibir a prática.

III – Proibição que decorre diretamente dos princípios contidos no art. 37, *caput*, da Constituição Federal¹⁶³.

[...]

RECURSO EXTRAORDINÁRIO. REPERCUSSÃO GERAL. LEI PROIBITIVA DE NEPOTISMO. VÍCIO FORMAL DE INICIATIVA LEGISLATIVA: INEXISTÊNCIA. NORMA COERENTE COM OS PRINCÍPIOS DO ART. 37, CAPUT, DA CONSTITUIÇÃO DA REPÚBLICA. RECURSO EXTRAORDINÁRIO PROVIDO.

[...]

Não é privativa do chefe do Poder Executivo a competência para a iniciativa legislativa de lei sobre nepotismo na administração pública: leis com esse conteúdo normativo dão concretude aos princípios da moralidade e da impessoalidade do art. 37, *caput*, da Constituição da República, que, ademais, têm aplicabilidade imediata, ou seja, independente de lei. Precedentes. Súmula Vinculante¹⁶⁴.

¹⁶¹ BRASIL. Lei n.º 8.429, de 2 de junho de 1992. Dispõe sobre as sanções aplicáveis em virtude da prática de atos de improbidade administrativa, de que trata o § 4º do art. 37 da Constituição Federal; e dá outras providências. [S. l.], 3 jun. 1992. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/L8429compilada.htm. Acesso em: 7 out. 2022.

De forma geral, há modalidades de atos de improbidade:

- a) art. 9º: com enriquecimento ilícito
- b) art. 10: com prejuízo ao erário
- c) art. 11: contra os princípios da Administração

¹⁶² BRASIL. Supremo Tribunal Federal. Acórdão da Ação Originária n. 1.833, Relator o Ministro Alexandre de Moraes, Primeira Turma, DJE de 8-5-2018.

¹⁶³ BRASIL. Supremo Tribunal Federal. RE 579.951, Rel.: Min. Ricardo Lewandowski, Julg.: 23.10.2008, Plenário, DJE de 24-10-2008.

¹⁶⁴ BRASIL. Supremo Tribunal Federal. RE 570.392, Rel.: Min. Cármen Lúcia, Julg.: 11-12-2014, Plenário, DJE de 19-2-2015, Tema 29.

Diante de todo o exposto, podemos concluir, com Wallace Paiva Martins Júnior¹⁶⁵, que o princípio constitucional da moralidade administrativa é “um superprincípio informador dos demais (ou um princípio dos princípios”, devendo sempre ser observado pelo agente público como parâmetro de conduta do bom administrador.

3.2.2 Da construção de normas a partir do exercício de interpretação constitucional por órgãos e entes da Administração Pública

A possibilidade de a hermenêutica constitucional ser exercida por “intérpretes informais” foi especialmente defendida por Peter Häberle¹⁶⁶, segundo o qual todas as forças pluralistas públicas são, potencialmente, intérpretes da Constituição. Para Häberle, a ideia de uma fixação exclusiva na jurisdição, de um monopólio da interpretação constitucional, deve ser superada. Desta forma, o processo constitucional formal não é a única via de acesso à interpretação, mesmo porque muitos problemas e questões constitucionais sequer chegam à Corte Constitucional, seja por falta de competência específica, seja pela falta de iniciativa de eventuais interessados. Existem atores do processo de interpretação constitucional em sentido amplo que desenvolvem, autonomamente, o direito constitucional material. Peter Häberle cita, como exemplo, o caso das disposições dos regimentos parlamentares¹⁶⁷:

Assim, a Constituição material “subsiste” sem interpretação constitucional por parte do juiz. Considerem-se as disposições dos regimentos parlamentares! Os participantes do processo de interpretação constitucional em sentido amplo e os intérpretes da Constituição desenvolvem, autonomamente, direito constitucional material. Vê-se, pois, que o processo constitucional formal não é a única via de acesso ao processo de interpretação constitucional.

Como consequência, o jurista alemão sustenta que a sociedade se torna mais aberta e livre, porque todos afetados pela norma constitucional estão aptos a oferecer alternativas para a sua interpretação, são participantes fundamentais no processo de “tentativa e erro”, de descoberta e obtenção do direito, ampliando-se o raio de interpretação normativa além do caso concreto. O constitucionalista é apenas um mediador no processo de interpretação constitucional, desempenhando um papel coadjuvante de “legitimar” o concerto acima referido.

¹⁶⁵ MARTINS JÚNIOR, Wallace Paiva. Probidade Administrativa. 3ª ed. São Paulo: Ed. Saraiva. 2006, p. 31.

¹⁶⁶ HÄBERLE, Peter. Hermenêutica Constitucional: a sociedade aberta dos intérpretes da Constituição - contribuição para a interpretação pluralista e procedimental da Constituição. Porto Alegre: Sergio Antonio Fabris, 1997.

¹⁶⁷ Ibid, p. 42.

O mérito da sua teoria está em ultrapassar o monopólio dos intérpretes tradicionais da Constituição como os magistrados, legisladores e chefes do poder executivo, ou seja, os destinatários da norma constitucional também interpretam e fazem valer os seus direitos. O reconhecimento da pluralidade e da complexidade da interpretação constitucional traduziria não apenas uma concretização do princípio democrático, mas também uma consequência metodológica da abertura material da Constituição, visto que o exercício hermenêutico constitucional deslocar-se-ia do âmbito processual, verificando-se um verdadeiro processo material de interpretação constitucional¹⁶⁸.

Esta pode ser a solução em relação à necessidade de orientação ética para adoção da IA pela Administração Pública: as esferas administrativas inferiores podem ser consideradas como participantes do processo constitucional de construção do direito material, na medida que poderiam exercer a hermenêutica do princípio da moralidade administrativa previsto no *caput* do art. 37 da CRFB, pois, como fora visto anteriormente, o próprio STF admite a autoaplicabilidade do princípio constitucional da moralidade administrativa a toda Administração Pública independente de lei.

Peter Häberle¹⁶⁹ ainda tece críticas à interpretação constitucional jurídica, no sentido de que teria uma visão reduzida da realidade, visão esta proporcionada pelos textos constitucionais que a antecedem: funcionam com filtros que limitam a perspectiva dos juízes. Os métodos “jurídicos” acabariam sendo um fator restritivo no processo de interpretação constitucional, eis que seriam disciplinados e disciplinadores, o que vale também para os processos legislativos de interpretação da Constituição, bem como para a Administração enquanto Administração “interpretativa” do bem-estar.

Também neste aspecto, o entendimento de Häberle em relação à necessidade de maior abertura material do direito, vai ao encontro do defendido na presente dissertação, visto que, no presente caso, significa receber dados e influxos da ciência moral e do princípio da moralidade administrativa, para construção de normas que regularão o uso ético da IA.

3.3 Como regular e regulamentar a IA na Administração Pública

3.3.1 Modelos regulatórios de IA

¹⁶⁸ Ibid, p. 52.

¹⁶⁹ Ibid, p. 43.

Não há dúvidas de que às tecnologias digitais se aplicam todas as leis vigentes, inclusive em relação aos direitos civil, penal, administrativo etc. Indaga-se, entretanto, até que ponto esse direito em grande parte relacionado às condições do “mundo analógico” é suficiente para o atendimento das exigências da transformação digital, especialmente quanto à forma de lidar com inteligência sintética, ou se precisa ser modificado e complementado¹⁷⁰. É mister, portanto, que o ordenamento jurídico possua mecanismos satisfatórios para o adequado tratamento ético-jurídico dos desafios relacionados à robótica cognitiva, mediante sua a regulamentação específica.

É necessário, antes de mais, compreender o conceito de regulação aqui adotado. Trata-se de uma restrição/interferência intencional nas eleições/escolhas de si ou de terceiros, por meio de atos indicativos, incitativos ou imperativos¹⁷¹. Quando realizado pelo Estado, caracteriza uma interferência pública, resultando em uma intervenção indireta, contínua e organizada e afetando a maneira como sujeitos públicos e privados se comportam em atividades econômicas ou sociais. Nas palavras de André Saddy¹⁷² regulação estatal é uma “interferência pública intencional que limita as escolhas dos privados e, até mesmo, do próprio Poder Público”.

André Saddy, ainda, ensina que a regulação pode ser realizada por alguém extrínseco ao sujeito (heterorregulação) ou pelo próprio interessado (autorregulação), além de pode ser pública, realizada por ente ou órgão público ou, privada, realizada pela iniciativa privada, *in verbis*¹⁷³:

- (i) heterorregulação pública, que normalmente se denomina apenas como regulação estatal, é uma forma de regulação externa, que vem de fora, executada por ente ou órgão público que interfere nas escolhas de terceiros, não necessariamente apenas sujeitos privados, mas também nas escolhas de sujeitos administrativos. [...]
- (ii) heterorregulação privada, que é uma regulação externa, que vem de fora, realizada por privados, em que um grupo de pessoas físicas e jurídicas confere poderes a uma entidade privada para que esta interfira, de forma autônoma, em suas liberdades de escolhas futuras. [...]
- iii) autorregulação pública, que é o estabelecimento, por meio de um documento escrito, de normas de conduta e padrões de comportamento criados por entes ou órgãos públicos, cujo objetivo é autolimitar ou cercear suas próprias liberdades de escolhas futuras. Exemplos: Quando uma empresa estatal edita um código de ética, conduta, integridade ou de boas práticas.
- (iv) autorregulação privada, que é o mesmo que a autorregulação pública, mas realizada por pessoas físicas ou jurídicas, ou grupo destas. [...]

¹⁷⁰ HOFFMANN-RIEM, Wolfgang. Inteligência Artificial Como Oportunidade para a Regulação Jurídica. Direito Público, Porto Alegre, v. 16, n. 90, p. 11-38, nov-dez, 2019.

¹⁷¹ SADDY, André. Curso de direito administrativo brasileiro. 2. ed. Rio de Janeiro: CEEJ, 2023, v. 2, p 555 e 556.

¹⁷² Ibidem, p. 560.

¹⁷³ Ibidem, p. 563 e 564.

O conceito de regulação é utilizado, portanto, como um instituto que compreende qualquer forma de atuação do Estado que interfira nas escolhas de alguém, seja de si mesmo ou de terceiros. A regulamentação é um dos instrumentos para tanto, quiçá o principal, mas regulação e regulamentação não são sinônimos.

Aplicando os referidos conceitos à regulação de IA no âmbito da Administração Pública, pode-se dizer que são 3 (três) as formas mais comuns que vem sendo utilizadas em diferentes ordenamentos: (i) heterorregulação pública holística ou global, baseada em direitos e/ou riscos, em que o legislador abarcaria o tratamento jurídico de forma geral, englobando o setor público e privado e independentemente da aplicação específica a que se destina a inteligência computacional. Com já mencionado, ainda não existe no nosso ordenamento jurídico uma legislação sobre IA; (ii) heterorregulação pública principiológica, onde a lei estipula os princípios que devem reger a adoção de IA, sem adentrar no tratamento casuístico ou setorial. Esse é o caminho que está sendo adotado no Projeto de Lei (PL) n.º 21/2020 que, conforme se observará, é objeto de severas críticas. Pode-se dizer que a EBIA também estabelece uma heterorregulação pública principiológica, porém, decorrente de um *softregulation*, é dizer, de uma regulação fraca, externada por meio de ações indicativas, ou seja, recomendações e incentivos à conduta das pessoas, de modo a atingir consensualmente os objetivos do instrumento; (iii) autorregulação pública, onde a normatização da IA é atribuída a diferentes entidades ou órgãos públicos em função da natureza da atividade desenvolvida. Um exemplo é a Portaria MCTI n.º 5.156/2021 que institui o programa “*MCTI Futuro: Futuro do Trabalho, Trabalho do Futuro*” no âmbito do Ministério da Ciência, Tecnologia e Inovações. Das 3 (três) formas antes mencionadas, é possível, ainda, realizar combinações, dando origem a uma quarta hipótese no que diz respeito especificamente à regulação da IA no âmbito da Administração Pública: (iv) a autorregulação regulada, ou melhor, a autorregulação pública heterorregulada pelo próprio Estado. Esta ocorrerá na hipótese em que o Estado regular a autorregulação pública por meio da tutela administrativa, ou seja, quando ente ou órgão distinto da Administração Pública regule a autorregulação pública, podendo se desdobrar em três modelos¹⁷⁴. Não se localizou regulamentação deste modelo no âmbito da Administração, mas

¹⁷⁴ Segundo André Saddy (SADDY, André. Curso de direito administrativo brasileiro. 2. ed. Rio de Janeiro: CEEJ, 2023, v. 2, p. 572) existem ao menos 3 (três) modelos existentes quando se correlaciona a heterorregulação ou autorregulação privada e a regulação estatal: (i) um modelo de heterorregulação ou autorregulação complementar ou suplementar regulada, (ii) um modelo de heterorregulação ou autorregulação regulada subsidiária ou supletiva; e (iii) um modelo de heterorregulação ou autorregulação regulada pura. Tal hipóteses, também podem ser aplicadas quando diante de autorregulação pública e a regulação estatal.

cita-se como exemplo a Resolução CNJ n.º 332/2020 que dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências.

Observando-se de forma geral as leis e regulamentos sobre IA de países integrantes da OCDE, nota-se que há uma grande variedade de modelos regulatórios, com alguns países mais centrados em questões precipuamente éticas, outros concedendo maior margem para regulação setorial, enquanto outros possuem uma regulação mais centralizada¹⁷⁵. Percebe-se, ainda, que alguns países parecem pouco dispostos a criar regulações nacionais, preferindo acompanhar as diretrizes de órgãos supranacionais, fenômeno este observado especialmente na União Europeia¹⁷⁶. Por fim, algumas nações optaram por atualizar leis e regulamentos para incluir previsões relativas às aplicações de sistemas inteligentes¹⁷⁷.

Para definição do melhor modelo regulatório de IA, é preciso levar em consideração que, mesmo que a lei seja atualizada e modernizada para abarcar novidades tecnológicas, é impossível que acompanhe *pari passu* a transformação digital acelerada que estamos vivenciando, visto que, como diria Kelsen¹⁷⁸, a ciência jurídica é “esta província afastada do centro do espírito que só lentamente costuma coxear atrás do progresso”. Demais disso, uma legislação excessivamente detalhista incorreria no risco de cair em rápida obsolescência, tendo em vista que se trata de tecnologia em constante e acelerada evolução, com inovações, funcionalidades e soluções sequer imaginadas pelo legislador.

Deve-se considerar, ainda, na esteira dos ensinamentos de Peter Häberle sobre interpretação constitucional jurídica desenvolvida no tópico anterior, que a normatização da inteligência computacional por métodos puramente jurídicos, como a regulamentação por uma

¹⁷⁵ Conforme achados constantes no relatório da comissão de juristas, Japão e o Reino Unido foram os países que estudaram de forma mais profunda diferentes modelos de regulação. A Coreia do Sul é o país que mais apoia o desenvolvimento de IA, e a Alemanha se destaca pela quantidade de iniciativas de regulação. Cf. COMISSÃO DE JURISTAS RESPONSÁVEL POR SUBSIDIAR ELABORAÇÃO DE SUBSTITUTIVO SOBRE INTELIGÊNCIA ARTIFICIAL NO BRASIL. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 2002. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04>. Acesso em: 8 mar. 2023. p. 260/261, p. 261

¹⁷⁶ COMISSÃO DE JURISTAS RESPONSÁVEL POR SUBSIDIAR ELABORAÇÃO DE SUBSTITUTIVO SOBRE INTELIGÊNCIA ARTIFICIAL NO BRASIL. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 2002. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04>. Acesso em: 8 mar. 2023. p. 260/26., p. 260/261.

¹⁷⁷ Ibid.

¹⁷⁸ KELSEN, Hans. Teoria pura do direito. Tradutor João Baptista Machado. 6. Ed. São Paulo: Martins Fontes, 1998, p. VII.

lei em sentido estrito, poderia gerar uma espécie de “leito de Procusto” sufocando a inovação, bem como iniciativas e incentivos ao uso e desenvolvimento da IA de forma excelente, eis que as potencialidades desta tecnologia ainda não foram plenamente desvendadas ou compreendidas. Por estas razões, Sergio Paulo Gallindo¹⁷⁹ defende que haja uma “regulamentação tardia ou *ex post*”, permitindo que haja uma maturação dos fenômenos tecnológicos e regulamentando na medida em que seja efetivamente necessário.

Seria, ainda, praticamente impossível dar tratamento legal de forma casuística a todas as situações em que a Administração pode se deparar com a necessidade de regulamentar mecanismos automatizados. Por exemplo, os cuidados e preocupações éticas relacionadas à regulação pela ANAC da adoção de sistemas autônomos no transporte aeroviário (para se prevenir acidentes causados por falha de sistema como os ocorridos com o Boeing 737 MAX¹⁸⁰) certamente serão completamente diferentes em relação à regulação para execução de atividades de auditoria de contratos administrativos por órgãos de controle, que serão distintos dos cuidados relacionados à prestação de serviços públicos ou execução de políticas públicas em *Smart Cities*.

A corroborar com esta visão, Fernando Filgueiras¹⁸¹ tece críticas à adoção de um regulador único, defendendo a utilização das capacidades estatais já existentes, num modelo policêntrico, “que possa ser coordenado a partir de um conjunto de parâmetros, de critérios, de requisitos que estejam presentes na lei, ou seja, uma série de princípios que orientem esse processo regulatório”.

Trata-se, assim, de buscar o adequado equilíbrio entre o uso justo e ético da IA, centrado no ser humano e respeitando valores democráticos, direitos e garantias fundamentais (tais como

¹⁷⁹ Ibid., nota 43

¹⁸⁰ O Boeing 737 MAX teve a sua autorização de voo suspensa em todo o mundo entre março de 2019 a novembro de 2020, depois que se identificou que um defeito no sistema de controle de voo causou a queda de duas aeronaves: uma na Indonésia e outra na Etiópia, entre outubro de 2018 e março de 2019. Na apuração dos fatos, foi amplamente reportado que a Boeing forçou situações para emitir a aprovação do 737 MAX pela Federal Aviation Administration (agência de transporte do governo dos EUA) para competir com o Airbus A320 neo, que chegou no mercado nove meses à frente do modelo da Boeing. O 737 MAX recebeu a certificação FAA em 8 de março de 2017. Dentre as irregularidades encontradas, identificou-se que, durante o processo de certificação, a FAA delegou várias avaliações à própria Boeing, permitindo que a fabricante revisse seu próprio produto. Cf. Gates, Dominic. «Flawed analysis, failed oversight: How Boeing and FAA certified the suspect 737 MAX flight control system». *The Seattle Times*. Disponível em: <https://www.seattletimes.com/business/boeing-aerospace/failed-certification-faa-missed-safety-issues-in-the-737-max-system-implicated-in-the-lion-air-crash/>. Acesso em: 9 jan. 2022

¹⁸¹ COMISSÃO DE JURISTAS RESPONSÁVEL POR SUBSIDIAR ELABORAÇÃO DE SUBSTITUTIVO SOBRE INTELIGÊNCIA ARTIFICIAL NO BRASIL. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 2002. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04>. Acesso em: 8 mar. 2023, nota 50.

a privacidade e a não-discriminação), por um lado, e o estímulo à inovação e ao desenvolvimento da inteligência sintética em ordem à realização do bem comum.

A prudência recomenda, portanto, evitar-se uma regulamentação legal açodada, tal como estava prestes a ocorrer com Projeto de Lei (PL) 21/20 que foi votado e aprovado na Câmara dos Deputados em 29/09/2021¹⁸². Ao ser encaminhado para votação no Senado Federal, o referido PL, que no Senado ganhou o número 872/2021, sofreu severas críticas em razão de se tratar de um texto prematuro, sem que tenha havido uma discussão ampla com a sociedade e especialistas¹⁸³. Então, o presidente do Senado, senador Rodrigo Pacheco, ante a reconhecida “complexidade técnica, jurídica e moral da ampliação de tecnologias de inteligência artificial como instrumentos auxiliares na tomada de decisões pelos poderes públicos e agentes privados”, nomeou uma Comissão de Juristas com a missão de elaborar um projeto substitutivo, assunto este que será tratado mais adiante¹⁸⁴. Constatou-se, destarte, que a complexidade e novidade do tema aconselham que o seu tratamento legal seja decorrente de uma discussão serena e dialógica no Congresso Nacional, com ampla participação da social.

Diante disso, muitos estudiosos defendem um modelo descentralizado e especializado que privilegie a regulação setorial, ainda que com algumas variações¹⁸⁵. Esta é a forma de regulação que vem sendo adotada nos Estados Unidos, ressaltando David Vladeck¹⁸⁶ que “compete aos especialistas [de cada setor] a habilidade de supervisionar de forma especializada as Ias”. Admite-se, sem embargo, a necessidade de uma coordenação geral entre setores, de forma a se evitar superposições e conflitos regulatórios¹⁸⁷, ou um marco legal de abrangência geral, ou que estabeleça princípios, que deverá ser conjugado com normas setoriais específicas.

¹⁸² CÂMARA DO DEPUTADOS. PL 21/2020, 29 set. 2021. Disponível em: <https://www.camara.leg.br/propostas-legislativas/2236340>. Acesso em: 23 mar. 2023.

¹⁸³ AGÊNCIA SENADO. Para especialistas, PL sobre Inteligência Artificial precisa de mais debates. Fonte: Agência Senado. Senado Notícias, [S. l.], p. página web, 9 dez. 2021. Disponível em: <https://www12.senado.leg.br/noticias/materias/2021/07/16/para-especialistas-pl-sobre-inteligencia-artificial-precisa-de-mais-debates>. Acesso em: 9 dez. 2021.

¹⁸⁴ RODRIGO PACHECO. Ato do Presidente do Senado Federal N° 4, DE 2022. Institui Comissão de Juristas responsável por subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei n°s 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 24 fev. 2022. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9087218&ts=1671479001450&disposition=inline>. Acesso em: 10 jan. 2023.

¹⁸⁵ Cf. COMISSÃO DE JURISTAS RESPONSÁVEL POR SUBSIDIAR ELABORAÇÃO DE SUBSTITUTIVO SOBRE INTELIGÊNCIA ARTIFICIAL NO BRASIL. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado n° 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei n°s 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 2002. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04>. Acesso em: 8 mar. 2023, nota 54.

¹⁸⁶ Ibid., p. 156.

¹⁸⁷ Ibid., nota 56

Assim, no conjunto da regulação setorizada, é possível vislumbrar modelos de autorregulação regulada, em uma combinação de códigos de conduta setoriais e heterorregulação pública, com incentivos à autorregulação somados a comandos e controles governamentais¹⁸⁸, podendo dar-se de forma pura, complementar e subsidiária.

3.3.2 Hetero e autorregulação pública dos aspectos ético-jurídicos da IA

Por todo exposto, e tendo em vista o especial enfoque deste estudo, voltado à resolução de questões ético-jurídicas relacionadas à adoção de sistemas inteligentes pela Administração Pública, no atual estágio de desenvolvimento da IA, parece recomendável que haja uma maior desconcentração e descentralização da regulação para, assim, atender à necessidade de se dar uma resposta jurídica tempestiva e equânime aos questionamentos decorrentes da adoção desta tecnologia. Por conseguinte, a Administração Pública deverá recorrer à autorregulação pública nos diferentes entes ou órgãos da Administração Pública, estipulando códigos de conduta, de boas práticas e deônticos com fundamento no princípio da moralidade administrativa. Além disso, recomenda-se existir um sistema de heterorregulação centralizado que estabeleça uma governança, coordenação e supervisão geral entre setores, de forma a se evitar superposições e conflitos regulatórios, conforme será mais bem desenvolvido no tópico relativo à Liberdade conformadora ou configuradora para Hetero e autorregulação da adoção de IA.

A partir da perspectiva da doutrina de Häberle apresentada no tópico anterior em relação ao objeto ora em estudo, a aportação de conhecimentos decorrentes da ciência moral enseja um maior elasticidade hermenêutica ao princípio da moralidade administrativa, podendo representar uma ampliação de perspectiva, ainda mais se considerado que se está diante de um conceito jurídico indeterminado, em que há uma margem de livre apreciação mais ampla.

Por outro lado, cabem algumas ponderações e temperamentos à teoria de Peter Häberle, levando-se em consideração, inclusive, as críticas lançadas por Paulo Bonavides¹⁸⁹, que apontou algumas dificuldades decorrentes da aplicação do método defendido:

[...] o método concretista da “Constituição aberta” demanda para uma eficaz aplicação a presença de sólido consenso democrático, base social estável, pressupostos institucionais firmes, cultura política bastante ampliada e desenvolvida, fatores sem dúvida difíceis de achar nos sistemas políticos e sociais das nações subdesenvolvidas ou em desenvolvimento, circunstância essa importantíssima, porquanto logo invalida como terapêutica das crises aquela metodologia cuja flexibilidade engana à primeira vista. Até mesmo para a constituição dos países desenvolvidos sua serventia se torna

¹⁸⁸ Ibid., nota 67.

¹⁸⁹ BONAVIDES, Paulo. Curso de direito constitucional. 10. ed. São Paulo: Malheiros, 2000.

relativa e questionável, com um potencial de risco manifesto. Debilitando o fundamento jurídico específico do edifício constitucional, a adoção sem freios daquele método – instalada uma crise que não se lograsse conjurar satisfatoriamente – acabaria por dissolver a Constituição e sacrificar a estabilidade das instituições. Ademais, o surto de preponderância concedido a elementos fáticos e ideológicos de natureza irreprimível é capaz de exacerbar na sociedade, em proporções imprevisíveis, o antagonismo de classes, a competição dos interesses e a repressão das idéias.

No presente caso objeto de estudo, a concessão de poder de hermenêutica constitucional para definição do agir, em consonância com a moralidade administrativa, a quaisquer servidores públicos responsáveis pela programação, desenvolvimento, treinamento ou aplicação de um algoritmo de IA certamente ensejaria maior insegurança jurídica. A ausência de um parâmetro normativo concreto e direcionador dificultaria sobremaneira a tarefa do agente, na medida em que é, por vezes, árduo vislumbrar como se concretiza o agir corretamente em determinadas situações e circunstâncias, haja vista a novidade e especificidade de situações em que o servidor poderá se deparar. Corre-se, ainda, o risco de haver decisões equivocadas e contraditórias entre diferentes agentes e órgãos da Administração.

Assim, nas circunstâncias atuais, a solução aqui proposta está na autorregulação pública setorial, mediante a emissão de atos normativos (regulamentos) pelos entes e órgãos da Administração Pública com base na moralidade administrativa. Segundo Maurice Hauriou¹⁹⁰, está é o “conjunto de regras de conduta tiradas da disciplina interior da Administração”, ou, nas palavras de René Ladreit de Lacharrière¹⁹¹, a moralidade administrativa é compreendida como uma “disciplina interna” estabelecida pelo superior hierárquico dentro da margem de liberdade concedida pela lei à Administração. Propõe-se, em suma, estipular uma disciplina interna da Administração sobre sistemas de IA mediante a emissão de normas ético-jurídicas.

Por conseguinte, os órgãos e entes da Administração Pública, revestidos de competência normativa e dentro da sua esfera de atribuição, deverão recorrer ao princípio da moralidade administrativa para regulamentar o design e uso dos sistemas inteligentes. O código de conduta ética em IA tem por finalidade direcionar e orientar os servidores públicos para que atuem de forma moral e constitucionalmente orientada, respeitando direitos e garantias fundamentais, não só para coibir o mau uso ou desvio para finalidades escusas, mas também para estimular a

¹⁹⁰HAURIOU, Maurice. Précis Elémentaire de Droit Administraof. 10. ed. Paris: Recueil Sirey, 1937 *apud* BRANDÃO, Antônio José. Moralidade administrativa. Revista de Direito Administrativo, Rio de Janeiro, v. 25, p. 454–467, jul. 1951, p. 457

¹⁹¹LACHARRIÈRE, René Ladreit de. Le Controle hierarchique de l'administration dans la forme juridictionnelle. [Paris]: Recueil Sirey, 1937 *apud* BRANDÃO, Antônio José. Moralidade administrativa. Revista de Direito Administrativo, Rio de Janeiro, v. 25, p. 454–467, jul. 1951, p.60

adoção da tecnologia de forma excelente e votada ao interesse público e ao atendimento do bem comum.

Por outro lado, analisando a questão sob o enfoque material, ainda que se observe uma estrutura organizada de governança interna, com emissão de normativos por esferas intermediárias competentes, o elastecimento material proposto por Häberle também pode ser fator de insegurança em virtude do risco de esvaziamento do texto constitucional em uma hermenêutica descomprometida com o princípio de juridicidade.

Destarte, para i) no aspecto jurídico formal, identificar de forma clara e precisa a esfera administrativa competente para emissão de um código ou norma para regulação setorial no âmbito da Administração; e ii) no aspecto jurídico material, definir as fronteiras para o exercício da hermenêutica constitucional do princípio da moralidade administrativa no agir ético dos servidores, é mister verificar as subjetividades públicas incidentes no exercício da atividade de gênese normativa, conforme será tratado nos tópicos seguintes.

3.3.3 Diretrizes para a regulamentação da IA: subjetividades incidentes

Tomando como base o entendimento de André Saddy¹⁹² a respeito das formas de manifestação da vontade administrativa (vinculada, conformadora, discricional, apreciativa ou por meio de conceitos jurídicos indeterminados), identificar-se-ão as subjetividades públicas incidentes à atividade regulamentar da IA.

Essa verificação é relevante, na medida em que proporciona a exata extensão e profundidade da liberdade e autonomia dos órgãos e entes da Administração Pública para definição da forma e conteúdo da norma. Feito esse esclarecimento, conforme procurou-se demonstrar a seguir, haverá incidência de duas subjetividades públicas, a saber: i) a liberdade conformadora ou configuradora e; ii) a margem de livre apreciação do princípio da moralidade administrativa enquanto conceito jurídico indeterminado.

3.3.3.1 Liberdade conformadora ou configuradora para regulamentação da adoção de IA

A liberdade de conformação ou de configuração é a subjetividade própria de quem cria normas jurídicas, define direitos e obrigações ou os elucida, ou seja, liberdade de opções entre

¹⁹² SADDY, André. Curso de direito administrativo brasileiro. 2. ed. Rio de Janeiro: CEEJ, 2023. v.1, p. 513

várias soluções possíveis, configurando, assim, os componentes lógico-estruturais da norma jurídica (hipótese, operador deôntico e consequência jurídica). Concretiza-se por meio de atos legislativos, normativo-regulamentares, editais, contratuais, jurisdicionais e até mesmo por tomadas de decisões administrativas normativas¹⁹³.

Trata-se, assim, de uma “discrecionabilidade forte”, onde há delegação deliberada de poder, manifestando-se de forma mais evidente no exercício da atividade legislativa e nas atribuições próprias de agentes políticos, tais como chefes do executivo, ministros e secretários, que definem prioridades para o governo. São os responsáveis pelo desenho de políticas públicas, por exemplo, eis que definem, com fundamento na legitimidade democrática, os meios que reputam adequados para alcançar determinado fim de interesse público¹⁹⁴.

Especificamente em relação à Administração, se por um lado recebe do Poder Legislativo a sua estrutura legal, atribuições e competências, por outro possui liberdade de conformação na definição de um conjunto de regras e meios para o exercício da sua atividade em ordem à realização do seu fim. Ou, mais bem explicado por Antônio José Brandão, ao comentar a entendimento de Maurice Hauriou¹⁹⁵:

Por outras palavras: no seio da empresa manifesta-se a vocação institucional, que requer a específica adaptação dos meios utilizáveis à consecução do objetivo previsto. Desta sorte, os agentes administrativos, se têm de inspirar os seus atos nas leis jurídicas vigentes, expressão normativa da ordem jurídica a que a Administração está submetida, movem-se também na órbita da ordem interna desta, gerada pela comunhão espiritual em que a ideia diretriz se tornou objetiva. Semelhante ordem interna encontra a sua expressão normativa num “direito natural da instituição”, - o qual, no caso da pública Administração, é a moralidade administrativa”.

A proposta de hetero e autorregulação pública setorial no seio da Administração Pública para o design e uso ético da IA possui natureza de ato normativo-regulamentar, visto que referida normatização dará forma ao informe, determinações precisas e positivadas, de sorte a conferir transparência, previsibilidade, segurança jurídica e uniformidade no agir da Administração sobre determinada atividade ou serviço público que pretende utilizar inteligência computacional.

Mas a liberdade de conformação ou de configuração não significa autonomia absoluta para criação de regras. Os órgãos e pessoas jurídicas pertencentes à Administração Pública devem observância ao princípio da juridicidade, ou seja, a atuação administrativa está

¹⁹³ SADDY, André. Apreciatividade e discrecionabilidade administrativa. 2. ed. Rio de Janeiro: CEEJ, 2020, p. 472.

¹⁹⁴ Ibid, p.76.

¹⁹⁵ BRANDÃO, Antônio José. Moralidade administrativa. Revista de Direito Administrativo, Rio de Janeiro, v. 25, n. 99, p. 454-467, 1951., p. 458.

submetida ao Direito como um todo, e não unicamente à estrita legalidade administrativa¹⁹⁶. O regulamento tem caráter secundário, subalterno, subordinado, inferior e complementar à Lei. Sua submissão à Lei, a princípio, é absoluta: não é fonte originária de direitos e obrigações, visto que se produz no âmbito e nos limites da Lei ou, no presente caso, da moralidade administrativa, de sorte a dar fiel cumprimento ao mandamento constitucional que visa regulamentar.

Demais disso, as entidades da Administração Pública, em geral, possuem normas que regem o processo interno de gênese normativa, atribuindo competência para elaboração, aprovação e alteração de atos normativos sobre matérias de sua atribuição e obrigando ao seu cumprimento àqueles que lhe estão jurisdicionados, sob pena de responsabilidade. No caso do TCU, por exemplo, identificou-se, em conformidade com o disposto no art. 2º do seu Regimento Interno¹⁹⁷, que a regulamentação da IA deverá ser emitida pela maior autoridade do órgão, ou seja, o Plenário. Portanto, o Plenário do TCU tem liberdade configuradora para emitir ato normativo-regulamentar sobre o design e uso de IA no âmbito do Tribunal.

Por outro lado, ainda que se defenda uma autonomia *interna corporis* de órgãos e entidades da Administração Pública revestidos de competência normativa para autorregulação setorial, nada impede que os órgãos superiores da Administração estabeleçam orientações de governança para o design e uso de IA, exercendo, assim, uma heterorregulação pública. Os chefes do Poder Executivo, por exemplo, poderão, mediante Decreto, estabelecer uma governança sobre adoção de robótica cognitiva na Administração Pública direta e indireta, estipulando políticas, princípios, e diretrizes que deverão ser observados pelos órgãos e entidades inferiores ou subordinados. Estes, por sua vez, poderão, então, emitir autorregulações setoriais para orientação mais concreta e casuística dos agentes públicos a eles subordinados.

Pode-se, inclusive, aproveitar normas, estruturas e procedimentos de governança já criados por força do disposto na Lei Geral de Proteção de Dados, como a figura do Encarregado¹⁹⁸, ou disposições referentes às Boas Práticas e da Governança¹⁹⁹, para que sejam estendidos a questões éticas relacionadas a IA. A estruturação da Governança pode prever uma área, colegiado ou comitê responsável por assegurar a conformidade e observância dos

¹⁹⁶ SADDY, André. *Curso de direito administrativo brasileiro*. 2. ed. Rio de Janeiro: CEEJ, 2023, v.1, p. 345

¹⁹⁷ BRASIL. Tribunal de Contas da União. Regimento interno do Tribunal de Contas da União. Brasília, 20 set. 2022. Disponível em: <https://portal.tcu.gov.br/data/files/AA/97/D7/1C/631D28102DFE0FF7F18818A8/RITCU.pdf>. Acesso em: 20 set. 2022.

¹⁹⁸ Cf. art. 41 et seq.

¹⁹⁹ Cf. art. 50 et seq.

princípios e normas no âmbito da Administração Pública, realizar treinamentos e medidas de conscientização, auditorias e certificações nos órgãos e entidades, bem como funcionar como um canal de comunicação entre a sociedade e a Administração Pública para emitir esclarecimentos públicos; receber solicitações e denúncias.

Para que não haja uma sobreposição de normas e orientações aplicáveis à inteligência sintética, recomenda-se que seja feito um trabalho prévio de harmonização com a legislação preexistente. Esse cuidado deve ser adotado especialmente em relação às normas de privacidade e proteção de dados, visto que o uso de IA envolve, em muitos casos, o uso massivo de dados pessoais, havendo na LGPD disposição específica relativamente ao tratamento automatizado de dados pessoais²⁰⁰.

Uma boa prática que vem sendo adotada por empresas e organizações privadas em relação à revisão de dados ou avaliação de conformidade a leis e normas internas é a criação de conselhos ou comitês de supervisão. O Facebook, por exemplo, criou um Comitê de Supervisão (*Oversight Board*), de forma a auxiliar a definir quais conteúdos são permitidos ou não dentro da plataforma²⁰¹. Da mesma forma, a Administração poderia criar conselhos ou comitês de supervisão de IA, inclusive com a participação de especialistas e representantes da sociedade civil que agreguem à discussão com uma perspectiva exógena à Administração.

Como decorrência do princípio da publicidade, é recomendado ainda que as minutas de governança, políticas, diretrizes, normas e códigos relativos à IA sejam submetidos a um escrutínio público, mediante a criação de espaços abertos para consultas públicas, audiências e debates que estimulem a discussão e participação da sociedade civil na concretização normativa dos princípios éticos a serem observados na pesquisa, desenvolvimento e uso da IA no âmbito da Administração Pública.

3.3.3.2 Margem de livre apreciação do princípio da moralidade administrativa

Além da incidência da liberdade configuradora/configuradora, a hetero e a autorregulação pública de IA deve ser uma decorrência da hermenêutica do princípio reitor da sua elaboração, da ética, que, no caso da Administração Pública, materializa-se pelo princípio

²⁰⁰ Cf. art. 20 et seq.

²⁰¹ WAKKA, Wagner. Brasileiro integra Comitê de Supervisão de Liberdade no Facebook. *Canaltech*, mai. 2020. Disponível em: <https://canaltech.com.br/redes-sociais/brasileiro-integra-comite-de-supervisao-de-liberdade-no-facebook-164485/>. Acesso em: 20 out. 2021.

da moralidade administrativa. Trata-se, com efeito, de um conceito jurídico indeterminado, visto que, como ministra André Saddy²⁰²:

Entende-se, assim, que conceitos jurídicos indeterminados são aqueles conceitos que não se podem traçar precisamente porque não deixam ao seu intérprete a compreensão exata de quais significados possui o objeto ou a realidade abrangida, ou seja, são os conceitos que deixam dúvida de duas ou mais compreensões de seu significado quando visualizado no caso concreto. A verdadeira indeterminação, e com ela a autonomia, surge da impossibilidade de identificar no todo ou em parte o significado, por meio da interpretação, frente ao caso concreto.

Sendo o princípio da moralidade administrativa um conceito jurídico indeterminado, ao processo de gênese normativa haverá a incidência de uma segunda subjetividade pública, qual seja, a margem de livre apreciação de conceito jurídico indeterminado. O conceito jurídico indeterminado acaba sendo uma forma de atribuição de competências do legislador às autoridades administrativas²⁰³. No presente caso, essa atribuição se manifesta pela autonomia e aplicabilidade imediata dos princípios previstos no art. 37, *caput*, da CRFB, conforme reconheceu o STF em acórdão mencionado anteriormente. Será, assim, o princípio da moralidade administrativa que legitima a ação da Administração para exercer a regulação/regulamentação setorializada da IA para as atividades administrativas.

Não se pode olvidar, entretanto, que a atividade hermenêutica em questão está inserida em um contexto constitucional: trata-se de interpretar uma norma dotada de superioridade hierárquica, que serve de fundamento de validade de toda a ordem jurídica. As demais normas jurídicas são impregnadas e interpretadas à luz dos valores veiculados pela Constituição.

Mas isso não quer dizer que as circunstâncias atuais não possam redundar em uma consequente mutação constitucional como resultante da autonomia concedida ao intérprete para concretizar valores constitucionais. De fato, no entender de Konrad Hesse²⁰⁴, a Constituição jurídica está condicionada pela realidade histórica, não podendo ser separada das circunstâncias concretas de seu tempo. Desta forma, a mudança das relações fáticas deve provocar mudanças na interpretação da Constituição, sob pena de, conforme diz Ferdinand Lassalle não passar de um pedaço de papel (em Stück Papier).

²⁰² SADDY, André. Apreciatividade e discricionariedade administrativa. 2. ed. Rio de Janeiro: CEEJ, 2020.

²⁰³ SOUSA, Antônio Francisco de. Os "conceitos legais indeterminados" no direito administrativo alemão. Revista de Direito Administrativo, v. 166, p. 276-291, 5 fev. 1986.

²⁰⁴ HESSE, Konrad. A Força Normativa da Constituição. Tradutor Gilmar Ferreira Mendes. Porto Alegre: SAFE, 1991, p. 9

Segundo Ana Cândida Ferraz²⁰⁵, as mutações constitucionais podem ser entendidas como processos informais de mudança da Constituição, podendo ser levadas a efeito não só pelo Judiciário, mas pelo Executivo e Legislativo. Contrapõe-se aos processos formais promovidos pelo poder constituinte derivado, como a emenda ou revisão. Isso porque a Constituição caracteriza-se por ser um “organismo vivo”, sujeito a desenvolvimento e mudanças²⁰⁶.

Em que pese a necessidade de atualização hermenêutica para que a Constituição preserve a sua força normativa, existe o risco de a mutação degenerar em um “tumor maligno”, num processo cancerígeno de ativismo que leve à instabilidade política e institucional e consequente “morte” da ordem constitucional, tal como alertado por Paulo Bonavides em excerto antes citado. Por isso, é necessária a definição de limites e balizas ao processo hermenêutico constitucional, o que é levado a efeito pelo princípio da conformidade ou correção funcional, pelo qual, segundo Canotilho²⁰⁷, “tem em vista impedir, em sede de concretização da constituição, a alteração de repartição de funções constitucionalmente estabelecida”. Caso esses limites sejam ultrapassados, incorreria o intérprete em uma mutação inconstitucional.

Konrad Hesse²⁰⁸ propõe que seja feita uma interpretação constitucional construtiva, sempre possível e necessária, dentro de certos limites:

Ao mesmo tempo, o sentido da proposição jurídica estabelece o limite da interpretação e, por conseguinte, o limite de qualquer mutação normativa. A finalidade (Telos) de uma proposição constitucional e sua nítida vontade normativa não devem ser sacrificadas em virtude de uma mudança da situação. Se o sentido de uma proposição normativa não pode mais ser realizado, a revisão constitucional afigura-se inevitável. Do contrário, ter-se-ia a supressão da tensão entre norma e realidade com a supressão do próprio direito: Uma interpretação construtiva é sempre possível e necessária dentro desses limites. A dinâmica existente na interpretação construtiva constitui condição fundamental da força normativa da Constituição e, por conseguinte, de sua estabilidade. Caso ela venha a faltar, tornar-se-á inevitável, cedo ou tarde, a ruptura da situação jurídica vigente.

Além disso, pelo princípio da interpretação efetiva (também chamado de princípio da máxima efetividade ou da eficiência), o intérprete deve dar a maior eficácia possível à norma constitucional. Deve-se atribuir à norma um sentido, alcance, conteúdo que lhe garanta a máxima efetividade, sendo vedado a interpretação que reduza ou suprima a finalidade. Embora

²⁰⁵ FERRAZ, Anna Cândida da Cunha. Processos Informais de Mudança da Constituição: Mutações Constitucionais e Mutações Inconstitucionais. São Paulo: Max Limonad, 1986, p. 12.

²⁰⁶ O espanhol Enrique Menault chama mutação de “mudança de contexto sem mudança de texto”.

²⁰⁷ CANOTILHO, José Joaquim Gomes. Direito Constitucional: e Teoria da Constituição. 7. ed. Coimbra: Almedina, 2003, p. 1224

²⁰⁸ HESSE, Konrad. A Força Normativa da Constituição. Tradutor Gilmar Ferreira Mendes. Porto Alegre: SAFE, 1991, p. 23.

esse princípio seja especialmente invocado no âmbito dos direitos fundamentais, Canotilho²⁰⁹ ensina que se trata de “um princípio operativo em relação a todas e quaisquer normas constitucionais”.

Podemos concluir, na esteira de Hesse²¹⁰, e em concerto com os princípios da correção funcional e da máxima efetividade, que a interpretação adequada é aquela que consegue concretizar, de forma excelente, o sentido (*Sinn*) da proposição normativa dentro das condições reais dominantes numa determinada situação. No caso objeto do nosso estudo, diante de novos riscos, situações e oportunidades derivados do avanço tecnológico, o intérprete deve buscar a concretização da renovação da força normativa do princípio constitucional da moralidade administrativa, sem ultrapassar, entretanto, a fronteira hermenêutica para este conceito jurídico indeterminado, ou seja, o seu sentido teleológico, sob pena de se autoconferir poder constituinte de forma oblíqua.

Outro aspecto que deve ser levado em consideração diz respeito ao porquê de haver “indeterminação” em conceitos jurídicos. Para André Saddy²¹¹, para além das três incertezas semânticas que caracterizam os conceitos jurídicos indeterminados (polissemia, vagueza e a textura aberta), acrescenta a confrontação de conceitos técnicos como uma forma de indeterminação.

Ademais, defende-se que três são as incertezas semânticas responsáveis para a dificuldade de determinação dos sentidos normativos: a polissemia, a ambiguidade e a textura aberta. Dessas três incertezas semânticas, deu-se a confrontação de conceitos técnicos, que é também uma forma de indeterminação no Direito que se analisou. A cada uma destas formas, defendeu-se que a textura aberta possui uma incapacidade natural para ser utilizada como instrumento de subjetividade ou autonomia pública e que os “conceptos” polissêmicos e vagos são instrumentos aceitáveis de utilização de subjetividade ou autonomia pública. O primeiro porque sua instrumentalização pode conferir alternativas ou opções, toda vez que os termos polissêmicos, dependendo do caso, podem conceder indeterminação ao significado da palavra e, o segundo, porque leva a inúmeras controvérsias sobre a relação que possa existir entre eles e uma eventual margem de liberdade. Por fim, a confraternização de conceitos técnicos também pode levar a indeterminações, toda vez que muitas situações, ainda ao ser analisadas por disciplinas técnicas, não oferecerem uma única resposta universalmente incontestável, ou seja, nem sempre a metodologia científica implica certeza absoluta e irrefutável. É, inclusive, quanto a esta última incerteza semântica que se discute o problema da chamada erroneamente discricionariedade técnica.

De qualquer forma, entende Saddy²¹² que, para identificar se um conceito jurídico indeterminado possui de fato uma margem de livre apreciação, é mister avaliar a sua estrutura,

²⁰⁹ CANOTILHO, José Joaquim Gomes. Direito Constitucional: e Teoria da Constituição. 7. ed. Coimbra: Almedina, 2003, p. 1224

²¹⁰ Ibid, p. 23.

²¹¹ SADDY, André. Apreciatividade e discricionariedade administrativa. 2. ed. Rio de Janeiro: CEEJ, 2020.

²¹² Ibid.

a partir das ideias de núcleo conceitual e auréola conceitual. O núcleo conceitual caracteriza-se “sempre que se tenha uma noção clara do conteúdo e da extensão de um conceito”. Por outro lado, a auréola conceitual ocorre “onde as dúvidas começam”²¹³:

O núcleo conceitual também é chamado de zona de certeza, que pode ser positiva (é seguro que “*eso*” pode denominar-se com esse termo) ou negativa (é seguro que “*eso*” não pode se denominar assim). A auréola do conceito é a zona intermediária, de dúvida ou de incerteza ou, dito com maior precisão, zona onde não existe uma certeza prévia e cuja determinação exige desdobrar a ideia nuclear do conceito. Toda a dificuldade de aplicação dos conceitos jurídicos indeterminados estará na zona de incerteza (ou “halo del concepto”), já que tem maior dificuldade para localizar a única solução justa.

Há casos, destarte, que tendem a haver um grande consenso, seja em relação à aplicação (certeza positiva) seja em relação à exclusão (certeza negativa) da definição ou aplicabilidade de um princípio, norma ou conduta como derivado de um conceito jurídico indeterminado²¹⁴. Neste sentido, a Administração deverá observar uma série de diretrizes fundamentadas na moralidade administrativa, diretrizes essas que serão melhor tratadas mais adiante. Cita-se algumas a título exemplificativo: a centralidade do ser humano (*human-centric AI*) e a correção de vieses algorítmicos (*bias*). Não há dúvidas de que os referidos parâmetros são derivados da moralidade administrativa, visto que, uma vez que não sejam observados pela Administração, haverá desvio de poder. As referidas diretrizes caracterizam, portanto, uma zona de certeza positiva, ou seja, há um consenso em relação ao reconhecimento da sua decorrência de uma ética algorítmica.

Por outro lado, pode haver discussão ou debate em relação à forma de aplicação da moralidade administrativa em casos que caracterizam uma zona de incerteza, ou seja, em que há divergência interpretativa na doutrina e na jurisprudência, tais como uso de IA para sistemas de identificação biométrica remota em espaços acessíveis ao público²¹⁵. Na exposição de motivos constante no relatório para elaboração do PL de IA, por exemplo, consta que inúmeras contribuições da sociedade e especialistas apontaram para uma elevada periculosidade relacionada a sistemas de identificação biométrica à distância de forma contínua e em espaços

²¹³ Ibid, p. 396.

²¹⁴ CARVALHO FILHO, José dos Santos. A discricionariedade: análise do seu delineamento jurídico. In: GARCIA, Emerson. *Discricionariedade Administrativa*. Rio de Janeiro: Lumen Juris, 2005.

²¹⁵ Algumas entidades defendem que deveria haver uma proibição total do uso de IA para os casos citados, como o Direitos Digitais Europeus (EDRi) e 119 organizações da sociedade civil lançaram uma declaração coletiva para pedir uma Lei de Inteligência Artificial (AIA) que priorize os direitos fundamentais. In: EUROPEAN DIGITAL RIGHTS. An EU Artificial Intelligence Act for Fundamental Rights: A Civil Society Statement. *Internet*, 30 nov. 2021. Disponível em: <https://edri.org/wp-content/uploads/2021/11/Political-statement-on-AI-Act.pdf>. Acesso em: 5 dez. 2021.

públicos. Por esta razão, a comissão de juristas recomendou que o tema seja regulado por lei federal específica²¹⁶.

A zona de incerteza é especialmente marcante nos casos de conflito entre igualdade material e igualdade formal, ou nas situações em que Aristóteles chama de “juízo de equidade” onde a aplicação de uma regra, que, por natureza, é geral e abstrata (e, portanto, generalizante) pode, a depender das peculiaridades de um dado caso concreto, causar uma injustiça, levando ao que Cícero²¹⁷ afirmava “*Summum ius, summa iniuria*”, ou seja, o “excesso de justiça” resulta em injustiça.

Esta situação é especialmente alarmante nos casos de adoção de mecanismos automatizados de decisão que afetem direitos fundamentais, visto que, conforme fora apontado no capítulo referente aos aspectos técnicos de IA, existem limitações e problemas computacionais matemáticos que inviabilizam completamente a capacidade de percepção de um algoritmo de que está errando ou pode errar. Ressaltam a respeito André Saddy e João Victor Tavares Galil²¹⁸:

Trazendo para o cenário do direito, parece claro que a concretização de uma decisão justa, que haverá de ser formada a partir da noção da igualdade material, deverá ser o norte perseguido pelo agente humano que, exercendo sua competência, decidiu agir através da utilização de inteligência artificial. Assim sendo, será seu dever a supervisão constante sobre a atuação do sistema, ainda que para sua remodelagem, sempre quando os parâmetros da tecnologia empregada já não mais atenderem às expectativas do agente competente ou da sociedade que, submetendo-se ao tratamento pela máquina, enxerga-se em uma relação jurídica perante o órgão ou perante a entidade estatal.

Por essas razões, nos casos de zona de incerteza ou auréola conceitual, conflito entre igualdade material e formal e nos juízos de equidade, é recomendável que sejam previstos mecanismos mais rigorosos para adoção de IA, tais como revisão humana obrigatória de decisões automatizadas.

3.3.4 Análise do PL de IA apresentado pela comissão de juristas

Conforme fora mencionado, o presidente do Senado instituiu, no dia 17/02/2022, uma comissão de 18 juristas²¹⁹ responsável por subsidiar a elaboração de minuta de substitutivo aos

²¹⁶ Ibid, p. 12.

²¹⁷ CICERO, Marcus Tullius. *De Officiis*. Transl. Walter Miller. Cambridge, MA: Harvard University Press, 1913.

²¹⁸ Ibid, p. 125.

²¹⁹ Composição da Comissão:

Presidente: Ricardo Villas Bôas Cueva

Relatora: Laura Schertel Ferreira Mendes

Projetos de Lei no 5.051/2019, 21/2020, e 872/2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da IA no Brasil. O relatório final da comissão foi entregue no dia 6/12/2022, contendo uma minuta de PL substitutivo com pouco mais de 40 artigos. A comissão promoveu reuniões, seminários e audiências públicas divididas por eixos temáticos, com a participação de especialistas e representantes nacionais e internacionais, visando aprofundar no tema de forma inclusiva e permeável à participação pública, seguindo a bem-sucedida experiência obtida nas discussões do Marco Civil da Internet e Lei Geral de Proteção de Dados²²⁰.

Especificamente em relação ao tema que nos interessa tratar, relativamente à regulação de IA, a minuta de PL apresentada pela Comissão possui a seguinte Exposição de Motivos:

[...]

Nessa quadra, esse novo marco legal tem um duplo objetivo. De um lado, estabelecer direitos para proteção do elo mais vulnerável em questão, a pessoa natural que já é diariamente impactada por sistemas de inteligência artificial, desde a recomendação de conteúdo e direcionamento de publicidade na Internet até a sua análise de elegibilidade para tomada de crédito e para determinadas políticas públicas. De outro lado, ao dispor de ferramentas de governança e de um arranjo institucional de fiscalização e supervisão, criar condições de previsibilidade acerca da sua interpretação e, em última análise, segurança jurídica para inovação e o desenvolvimento econômico-tecnológico.

Portanto, este substitutivo de projeto de lei parte da premissa de que não há um trade-off – uma escolha mutuamente excludente – entre a proteção de direitos e liberdades fundamentais, da valorização do trabalho e da dignidade da pessoa humana face à ordem econômica e à criação de novas cadeias de valor. Pelo contrário, seus fundamentos e a sua base principiológica buscam tal harmonização, conformando-se à Constituição Federal e de forma dialógica com outras leis que enfrentam o mesmo tipo de desafio (e.g., Código de Defesa do Consumidor e Consolidação das Leis do Trabalho).

Seu objetivo normativo é conciliar uma abordagem baseada em riscos com uma modelagem regulatória baseada em direitos. Ao mesmo tempo em que se preveem instrumentos de governança para que sejam prestadas contas e seja premiada a boa-fé dos agentes econômicos que gerenciam de forma eficaz os riscos em torno da concepção e implementação de sistemas de inteligência artificial, também há uma forte carga obrigacional para florescimento do escrutínio individual e social a seu respeito.

Nesse sentido, o segundo capítulo deste projeto de lei é dedicado à previsão de direitos e deveres, de acordo com os efeitos de um sistema de inteligência artificial na vida

Titulares: Ana de Oliveira Frazão; Bruno Ricardo Bioni; Danilo Cesar Maganhoto Doneda; Fabrício de Mota Alves; Miriam Wimmer; Wederson Advincula Siquiera; Claudia Lima Marques; Juliano Souza de Albuquerque Maranhão; Thiago Luís Santos Sombra; Georges Abboud; Frederico Quadros D'Almeida; Victor Marcel Pinheiro; Estela Aranha; Clara Iglesias Keller; Mariana Giorgetti Valente; Filipe Medon. In: PACHECO, Rodrigo. Ato do Presidente do Senado Federal nº 4, de 2022. Institui Comissão de Juristas responsável por subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 24 fev. 2022. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9087218&ts=1671479001450&disposition=inline>. Acesso em: 10 jan. 2023.

²²⁰ Agência Senado. Comissão conclui texto sobre regulação da inteligência artificial no Brasil. [Senado Notícias](https://www12.senado.leg.br/noticias/materias/2022/12/06/comissao-conclui-texto-sobre-regulacao-da-inteligencia-artificial-no-brasil), Brasília, 6 dez. 2022. Disponível em: <https://www12.senado.leg.br/noticias/materias/2022/12/06/comissao-conclui-texto-sobre-regulacao-da-inteligencia-artificial-no-brasil>. Acesso em: 10 jan. 2023.

das pessoas. Além de fixar direitos básicos e transversais para todo e qualquer contexto em que há interação entre máquina e ser humano (e.g., de informação e transparência), intensifica-se tal carga obrigacional quando o sistema de IA produz efeitos jurídicos relevantes ou impactem os sujeitos de maneira significativa (e.g., direito de contestação e intervenção humana). Busca-se assegurar contraditório e ampla defesa – devido processo informacional – a quem possa ter direitos e liberdades afetadas por uma decisão totalmente ou parcialmente automatizada.

Assim, o peso da regulação é dinamicamente calibrado de acordo com os potenciais riscos do contexto de aplicação da tecnologia. Foram estabelecidas, de forma simétrica aos direitos, determinadas medidas gerais e específicas de governança para, respectivamente, sistemas de inteligência artificial com qualquer grau de risco e para os categorizados como de alto risco.

[...]

Com isso, a partir de uma abordagem mesclada de disposições que se aplicam ex-ante e ex-post, o substitutivo traça critérios quantitativos e qualitativos para fins de cognição, avaliação e desencadeamento granular de quais tipos de ações devem ser tomadas para mitigação dos riscos em jogo. Inclusive, com delegações para que a regulação seja uma empreitada colaborativa, que privilegia o envolvimento dos setores interessados no processo regulatório, ao invés de conduzi-lo exclusivamente dentro das capacidades do Estado (corregulação).

Além de incentivar a adoção de arranjos consensuais, como, por exemplo, ao dispor que códigos de boas práticas são um indicativo de boa-fé, o substitutivo dedica uma seção para procedimentalizar os chamados sandboxes regulatórios. [...].”

Analisando os dispositivos do PL, constata-se, *ab initio*, que se trata de uma lei geral, que estabelece normas para o desenvolvimento, implementação e uso de IA no Brasil, visando “proteger os direitos fundamentais e garantir a implementação de sistemas seguros e confiáveis, em benefício da pessoa humana, do regime democrático e do desenvolvimento científico e tecnológico”²²¹.

Percebe-se, ainda, a clara referência da LGPD para elaboração do Projeto, recorrendo-se a uma estrutura normativa bastante similar. Tal como a Lei Geral de Proteção de Dados, suas normas são de interesse nacional e devem ser observadas pela Administração Pública da União, Estados, Distrito Federal e Municípios, abarcando tanto o setor público quanto o privado²²²; possuem definições semelhantes²²³; ambas preveem uma autoridade responsável por zelar,

²²¹ Cf. art. 1º do PL. In: COMISSÃO DE JURISTAS RESPONSÁVEL POR SUBSIDIAR ELABORAÇÃO DE SUBSTITUTIVO SOBRE INTELIGÊNCIA ARTIFICIAL NO BRASIL. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil.. Brasília, 2002. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04>. Acesso em: 8 mar. 2023. p. 114.

²²² Ibid., art. 4º, II e III, onde, tanto o fornecedor quanto o operador podem ser uma pessoa natural ou jurídica, de natureza pública ou privada. Na LGPD, art. 5º, VI E VII, controlador e operador também podem ser pessoa natural ou jurídica, de direito público ou privado.

²²³ Ibid., art. 4º. “Para as finalidades desta Lei, adotam-se as seguintes definições: IV – agentes de inteligência artificial: fornecedores e operadores de sistemas de inteligência artificial”. A LGPD, por seu turno, no art. 5º: “Para os fins desta Lei, considera-se: IX - agentes de tratamento: o controlador e o operador”.

implementar e fiscalizar o cumprimento da lei²²⁴; e estabelecem sanções administrativas similares²²⁵.

Com relação à governança de IA, o Projeto estabelece a obrigação de os agentes de Inteligência Artificial²²⁶ possuírem estruturas de governança e processos internos de forma a garantir a segurança dos sistemas e o atendimento dos direitos de pessoas afetadas. Deverão ser adotadas medidas desde a concepção inicial (*by design*) aplicáveis ao longo de todo o ciclo de vida, até o encerramento de suas atividades. O Projeto determina medidas tais como: transparência quanto ao uso de sistemas de IA na interação com pessoas naturais, mitigação e prevenção de potenciais vieses discriminatórios, conformidade à legislação de proteção de dados e segurança da informação²²⁷.

Sobre as medidas de governança para sistemas de IA de alto risco, deverá ser implementado um *framework* para avaliação de riscos a direitos²²⁸, acrescentando medidas aplicáveis ao Poder Público tais como, consulta e audiência públicas prévias, controle de acesso e utilização do sistema, garantia de direito à explicação e revisão humanas, publicização das avaliações preliminares dos sistemas de IA desenvolvidos independentemente do grau de risco²²⁹. Além disso, é obrigatória a avaliação de impacto algorítmico de IA para sistema considerado como de alto risco pela avaliação preliminar²³⁰. Caso os riscos identificados nesta avaliação não possam ser eliminados ou substancialmente mitigados, a utilização do sistema deverá ser descontinuada²³¹.

Relativamente à regulação setorial, há expressa menção, na Exposição de Motivos acima citada, da adoção de um modelo corregulatório, com delegações para estimular uma regulação colaborativa entre setores interessados e o Estado, o que caracteriza uma heterorregulação ou autorregulação complementar ou suplementar regulada. Neste sentido, o PL prevê a possibilidade de os agentes de IA formularem códigos de boas práticas e de governança, individualmente ou por meio de associações, que estabeleçam as condições de organização, o regime de funcionamento, os procedimentos, as normas de segurança, os

²²⁴Ibid, art. 4º. “Para as finalidades desta Lei, adotam-se as seguintes definições: V – autoridade competente: órgão ou entidade da Administração Pública Federal responsável por zelar, implementar e fiscalizar o cumprimento desta Lei em todo o território nacional”. A LGPD, por seu turno, no art. 5º: “Para os fins desta Lei, considera-se: XIX - autoridade nacional: órgão da administração pública responsável por zelar, implementar e fiscalizar o cumprimento desta Lei em todo o território nacional”.

²²⁵ Ibid, no PL, art. 36 e, na LGPD, art. 52.

²²⁶ Ibid, nota 113.

²²⁷ Ibid, art. 19 do PL.

²²⁸ Ibid, art. 20 do PL.

²²⁹ Ibid, art. 21 do PL.

²³⁰ Ibid, art. 22 do PL.

²³¹ Ibid, §2º do art. 21 do PL.

padrões técnicos, as obrigações específicas para cada contexto de implementação, as ações educativas, os mecanismos internos de supervisão e de mitigação de riscos e as medidas de segurança técnicas e organizacionais apropriadas para a gestão dos riscos decorrentes da aplicação dos sistemas²³².

Com relação à implementação de programa de boas práticas de governança, o Projeto estimula que sejam adotados processos e políticas internas que assegurem o cumprimento, de forma abrangente, de normas e boas práticas relativas à não maleficência e proporcionalidade entre os métodos empregados e as finalidades determinadas e legítimas dos sistemas de IA²³³, bem como que tenha o objetivo de estabelecer relação de confiança com as pessoas afetadas, por meio de atuação transparente e que assegure mecanismos de participação²³⁴. A adesão voluntária a código de boas práticas e governança pode ser considerada como indicativo de boa-fé por parte do agente e será levada em consideração pela autoridade competente para fins de aplicação de sanções administrativas²³⁵.

Desta forma, se o PL fosse aprovado tal como proposto, não haveria prejuízo às conclusões do presente estudo, visto que o legislador concederia liberdade de conformação ou de configuração para autorregulação pública setorizada a ser levada a efeito pelos órgãos e entidades da Administração Pública revestidos de competência normativa para regulação *interna corporis* em questões relacionadas à moralidade administrativa²³⁶. Evidentemente, dever-se-á observar a heterorregulação prevista no PL, tais como direitos aplicáveis em todos os casos em que há interação com IA, bem como o atendimento de deveres e responsabilidades específicas em função do grau de risco identificado na avaliação de cada caso.

²³² Ibid, art. 30 do PL.

²³³ Ibid, alínea “a” do § 2º do art. 30 do PL.

²³⁴ Ibid, alínea “c” do § 2º do art. 30 do PL.

²³⁵ Ibid, § 3º do art. 30 do PL.

²³⁶ Esse entendimento também é válido em relação ao PL n.º 21/2020 que adotou um modelo de heterorregulação pública principiológica, havendo, assim, espaço para a autorregulação interna e setorizada na Administração Pública.

CAPÍTULO IV - PROPOSTA DE DIRETRIZES PARA CONSTRUÇÃO DE NORMAS ÉTICO-JURÍDICAS RELATIVAS AO DESIGN E USO DE IA

4.1 Referências de parâmetros éticos para IA

Para definição de diretrizes derivadas da moralidade administrativa, que deverão servir como norte para construção de normativos específicos de órgãos e entidades da Administração Pública, é mister, antes, ter uma noção sobre princípios e parâmetros éticos que geralmente são identificados na relação com máquinas inteligentes.

A seguir serão apresentadas algumas das fontes mais difundidas na atualidade e que geralmente são utilizadas como referência de países e organismos internacionais que pretendem construir um arcabouço normativo relativo à IA, a saber: i) as Três Leis da Robótica de Isaac Asimov; ii) Os Princípios sobre IA da OCDE; iii) a Declaração do G20 sobre Princípios para IA Centrada nos Humanos; e iv) As Diretrizes de Ética para Inteligência Artificial Confiável da União Europeia.

4.1.1 As Três Leis da Robótica de Isaac Asimov

A ideia de se definir diretrizes éticas para nortear o design e uso de IA fora difundida especialmente pelo escritor de ficção científica Isaac Asimov, que cunhou as Três Leis da Robótica, apresentadas em conjunto pela primeira vez em *Runaround*²³⁷:

[...]

–Temos o seguinte. A primeira: um robô não pode ferir um ser humano ou, por inação, permitir que um ser humano venha a ser ferido.

– Certo! – A segunda –continuou Powell – um robô deve obedecer às ordens dadas por seres humanos, exceto nos casos em que tais ordens entrem em conflito com a Primeira Lei.

– Certo!

–E a terceira: um robô deve proteger sua própria existência, desde que tal proteção não entre em conflito com a Primeira ou com a Segunda Lei.

Asimov tinha por objetivo definir regras que expressassem a centralidade do ser humano, por um lado, e a subordinação das máquinas, de outro, tornando viável a coexistência

²³⁷ *Apud* SEIFFERT, A. S. Os robôs de Asimov e o futuro da humanidade. *Revista Eletrônica Da ANPHLAC*, v. 24, p. 374–393, 2018. Disponível em: <https://revista.anphlac.org.br/anphlac/article/view/2877>. Acesso em: 07 dez. 2021. p. 387.

pacífica entre pessoas e robôs inteligentes, de sorte que a IA não se rebelasse ou até chegasse a subjugar os homens. As referidas leis inspiraram inúmeros filmes e romances de ficção científica a respeito²³⁸.

As Três Leis da Robótica vem sendo objeto de discussão por cientistas e filósofos, especialmente em relação à aptidão para resolução de todos os possíveis problemas na relação homem-robô, visto que a IA não é dotada de uma capacidade (inerentemente humana) de discernimento entre o bem e o mal. Por conseguinte, um sistema computacional pode incorrer em violações às Leis da Robótica sem que sequer perceba. Além disso, as referidas leis não levam em consideração eventuais falhas de hardware ou software, ou mesmo limitações intrínsecas à tecnologia. Por fim, o conflito entre cada uma das leis também é fruto de problemas éticos complexos²³⁹.

Sem embargo é inegável que as Três Leis da Robótica tem o mérito de erigir a centralidade do ser humano como princípio fundamental orientador do comportamento e funcionamento de máquinas pensantes, sendo considerada uma das mais importantes referências na discussão de questões éticas relacionadas à IA.

4.1.2 Os Princípios sobre IA da OCDE e a Declaração do G20 sobre Princípios para IA Centrada nos Humanos

No plano internacional, os principais debates sobre regulação de IA giram em torno do estabelecimento de diretrizes e boas práticas para nortear governos e entidades privadas, de forma que os sistemas de IA sejam centrados no ser humano (*human centric AI*) e confiáveis (*trustworthy AI*). Dentre as inúmeras iniciativas nesse sentido, destacam-se os princípios da OCDE sobre IA e a Declaração do G20 sobre Princípios para IA Centrada nos Humanos.

Ainda que não seja juridicamente vinculante, a OCDE pretendeu definir um padrão internacional para auxiliar os governos na elaboração da legislação nacional a respeito. Em maio de 2019, estabeleceu cinco princípios para promover uma IA que seja inovadora, confiável, que respeite os direitos humanos e os valores democráticos. Além dos membros da OCDE, outros países, incluindo Argentina, Brasil, Costa Rica, Malta, Peru, Romênia e Ucrânia

²³⁸ O próprio Asimov escreveu o livro *I, Robot* em 1950, obra em que se baseou o filme homônimo de 2004, dirigido por Alex Proyas.

²³⁹ Cf. por exemplo, BOSTROM, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014.

já aderiram aos Princípios da AI²⁴⁰.

A OCDE recomenda, assim, a adoção dos seguintes princípios:

- 1 - A IA deve beneficiar as pessoas e o planeta ao impulsionar o crescimento inclusivo, o desenvolvimento sustentável e o bem-estar.
- 2 - Os sistemas de IA devem ser concebidos de forma a respeitar o Estado de Direito, os direitos humanos, os valores democráticos e a diversidade, e devem incluir salvaguardas adequadas - por exemplo, permitindo a intervenção humana quando necessário - para garantir uma sociedade igualitária e justa.
- 3 - 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.
- 4 - Os sistemas de IA devem funcionar de maneira robusta, segura e protegida ao longo de seus ciclos de vida e os riscos potenciais devem ser avaliados e gerenciados continuamente.
- 5 - Organizações e indivíduos desenvolvendo, implantando ou operando sistemas de IA devem ser responsabilizados por seu funcionamento adequado, de acordo com os princípios acima.

Para os governos, baseando-se nos princípios acima, a OCDE também fornece cinco recomendações²⁴¹:

- 1 - Facilitar o investimento público e privado em pesquisa e desenvolvimento para estimular a inovação em IA confiável.
- 2 - Promova ecossistemas de IA acessíveis com infraestrutura digital e tecnologias e mecanismos para compartilhar dados e conhecimento.
- 3 - Garanta um ambiente de política que abrirá o caminho para a implantação de sistemas de IA confiáveis.
- 4 - Capacite as pessoas com as habilidades para IA e apoie os trabalhadores para uma transição justa.
- 5 - Cooperar além das fronteiras e setores para progredir na gestão responsável de IA confiável.

A Declaração do G20 sobre Princípios para IA Centrada nos Humanos, por sua vez, consignou o seguinte²⁴²:

3. Inteligência Artificial Centrada no Homem (IA)

17. Reconhecendo os esforços realizados até agora por todas as partes interessadas em suas respectivas funções, incluindo governos, organizações internacionais, academia, sociedade civil e setor privado, e ciente de como a tecnologia impacta a sociedade, o G20 se esforça para fornecer um ambiente propício para IA centrada no ser humano que promova inovação e investimento, com foco particular em empreendedorismo digital, pesquisa e desenvolvimento, ampliação de startups nesta área e adoção de IA pelas micro, pequeno e médias empresas que enfrentam custos desproporcionalmente maiores para adotar a IA.

18. Reconhecemos que as tecnologias de IA podem ajudar a promover o crescimento econômico inclusivo, trazer grandes benefícios para a sociedade e capacitar os

²⁴⁰ OCDE. OECD Principles on AI. [S. l.], 2019. Disponível em: <https://www.oecd.org/going-digital/ai/principles/>. Acesso em: 9 dez. 2021.

²⁴¹ Ibid.

²⁴² G20. G20 Ministerial Statement on Trade and Digital Economy. 2019. Disponível em: <https://www.mofa.go.jp/files/000486596.pdf>. Acesso em: 27 dez. 2020.

indivíduos. O desenvolvimento responsável e o uso de IA podem ser uma força motriz para ajudar a avançar as metas de desenvolvimento sustentável e realizar uma sociedade sustentável e inclusiva, mitigar riscos para valores sociais mais amplos. Os benefícios trazidos pelo uso responsável da IA podem melhorar o ambiente de trabalho e a qualidade de vida e criar potencial para a realização de uma sociedade futura centrada no ser humano com oportunidades para todos, incluindo mulheres e meninas, bem como Grupos vulneráveis.

19. Ao mesmo tempo, também reconhecemos que a IA, como outras tecnologias emergentes, pode apresentar desafios da sociedade, incluindo as transições no mercado de trabalho, privacidade, segurança, ética questões, novas brechas digitais e a necessidade de capacitação em IA. Para fomentar a confiança pública e confiança nas tecnologias de IA e perceber totalmente seu potencial, estamos comprometidos com uma abordagem de IA centrada no ser humano, guiada pelos Princípios do G20 sobre IA, extraídos da Recomendação da OCDE sobre IA, que estão no Anexo e não são vinculativos. Este anexo inclui os seguintes princípios de “crescimento inclusivo, desenvolvimento sustentável e bem-estar”, “valores centrados no ser humano e justiça”, “transparência e explicabilidade”, “robustez, segurança e proteção” e “responsabilidade”. O Anexo também oferece orientação para consideração pelos formuladores de políticas com o objetivo de maximizar e compartilhar os benefícios da IA, enquanto minimiza os riscos e preocupações, com atenção especial à cooperação internacional e inclusão de países em desenvolvimento e populações sub-representadas.

20. Ao buscar a IA centrada no ser humano, os membros do G20 reconhecem a necessidade de continuar a promover a proteção da privacidade e dos dados pessoais de acordo com as estruturas aplicáveis. O G20 também reconhece a necessidade de promover a capacitação e o desenvolvimento de habilidades em IA. Cada um de nós continuará a lutar pela cooperação internacional e por trabalhar em conjunto com fóruns apropriados em áreas como pesquisa e desenvolvimento, desenvolvimento de políticas e compartilhamento de informações por meio do Repositório de Políticas Digitais do G20 e outros esforços abertos e colaborativos.

Nota-se, assim, que tanto os princípios da OCDE quanto a Declaração do G20 tem como enfoque principal orientar o desenvolvimento e uso de sistemas de IA de forma ética, confiável, responsável e centrada nos seres humanos. Os referidos princípios dão destaque para a promoção de uma governança adequada para garantir a transparência, explicabilidade, privacidade, direitos individuais, sociais e a proteção do meio ambiente, de sorte a beneficiar a sociedade como um todo.

Embora sejam considerados importantes por resultarem de um esforço e colaboração internacional voltados à definição de orientações globais sobre IA, o fato de os princípios da OCDE e a Declaração do G20 serem meras declarações de intenções, por demais amplas e vagas, sem incluir ações e medidas definidas, acaba prejudicando a concretização por parte de entidades públicas e privadas sobre como, na prática, desenvolver e utilizar a IA de maneira ética e responsável.

4.1.3 Diretrizes de Ética para IA Confiável da União Europeia

Embora a União Europeia não possua uma legislação específica sobre IA, em abril de 2021 a Comissão Europeia apresentou uma proposta de regulamento, trazendo disposições, dentre outras, sobre ética, responsabilidade, transparência e privacidade²⁴³.

Sem embargo, a Comissão Europeia estabeleceu uma estratégia para a IA em abril de 2018, definindo a abordagem do bloco europeu para o seu desenvolvimento e uso, incluindo recomendações sobre questões éticas e regulamentação²⁴⁴. Em abril de 2019, a Comissão emitiu uma orientação sobre Diretrizes de Ética para Inteligência Artificial Confiável²⁴⁵, estabelecendo os seguintes princípios basilares: iniciativa e controlo por humanos, robustez e segurança, privacidade e governação dos dados, transparência, diversidade, não discriminação e equidade, bem-estar societal e ambiental e responsabilização:

As orientações determinam que, para se alcançar uma “IA de confiança”, são necessárias três componentes: 1) Conformidade com a legislação, 2) Respeito dos princípios éticos e 3) Robustez.

Com base nestas três componentes e nos valores europeus definidos na secção 2, as orientações identificam sete requisitos essenciais que as aplicações de IA devem respeitar para serem consideradas de confiança. As orientações incluem igualmente uma lista de avaliação para ajudar a verificar se estes requisitos são cumpridos.

[...]

I. Iniciativa e controlo por humanos: Os sistemas de IA devem ajudar os indivíduos a fazerem escolhas melhores e mais informadas, em conformidade com os seus objetivos. Devem agir como facilitadores de uma sociedade próspera e equitativa, apoiando a atividade humana e promovendo os direitos fundamentais, e não reduzindo, limitando ou guiando a autonomia humana. O bem-estar geral do utilizador deve estar no centro da funcionalidade do sistema. O controlo humano contribui para garantir que um sistema de IA não prejudica a autonomia humana nem causa outros efeitos adversos. Em função do sistema específico baseado em inteligência artificial e da sua área de aplicação, devem ser assegurados os graus adequados de medidas de controlo, incluindo a capacidade de adaptação, a exatidão e a explicação dos sistemas baseados em inteligência artificial. A supervisão pode ser efetuada através de mecanismos de governação, como a garantia de uma abordagem que respeite os princípios da intervenção humana, da supervisão humana e da detenção do controlo por humanos («human-in-the-loop», «human-on-the-loop» ou «human-in-command»). Deve garantir-se que as autoridades públicas têm capacidade para exercer os seus poderes de supervisão em conformidade com os seus mandatos. Dado que todos os outros fatores são iguais, quanto menor for o grau de supervisão de um ser humano sobre um sistema de IA, maior é a necessidade de ensaios mais aprofundados e de uma governação mais rigorosa.

II. Robustez e segurança: Uma IA de confiança exige que os algoritmos sejam seguros, fiáveis e suficientemente robustos para lidar com erros ou incoerências durante todas as fases do ciclo de vida do sistema de IA e para lidar adequadamente

²⁴³ União Europeia. Proposta de Regulamento nº 2021/0106(COD), de 21 de abril de 2021. Inteligência Artificial. [S. l.], 21 abr. 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>. Acesso em: 12 fev. 2023.

²⁴⁴ UNIÃO EUROPEIA. Inteligência artificial para a Europa, COM/2018/237. Bruxelas, 25 abr. 2018. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52018DC0237>. Acesso em: 12 fev. 2023.

²⁴⁵ UNIÃO EUROPEIA. Aumentar a confiança numa inteligência artificial centrada no ser humano, COM (2019) 168 final. Bruxelas, 8 abr. 2019. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52019DC0168&from=EN>. Acesso em: 12 fev. 2023.

com resultados errados. Os sistemas de IA têm de ser fiáveis, suficientemente resilientes contra os ataques não dissimulados e tentativas mais subtis de manipular dados ou os próprios algoritmos, e devem assegurar um plano de recurso em caso de problemas. As suas decisões devem ser exatas, ou, pelo menos, refletir corretamente o seu nível de precisão, e os seus resultados devem ser reproduzíveis. Além disso, os sistemas de IA devem integrar mecanismos de proteção e de segurança desde a conceção para garantir que são comprovadamente seguros em todas as fases, tendo em conta a segurança física e mental de todas as partes envolvidas. Tal inclui a minimização e, sempre que possível, a reversibilidade de consequências ou de erros não intencionais na operação do sistema. Devem ser criados processos destinados a clarificar e avaliar os riscos potenciais associados à utilização de sistemas de IA em diferentes áreas de aplicação.

III. Privacidade e governação dos dados: A privacidade e a proteção de dados devem ser garantidas em todas as fases do ciclo de vida do sistema de IA. Os registos digitais de comportamento humano podem permitir aos sistemas de IA inferir não só as preferências dos indivíduos, a sua idade e o seu género, como também a sua orientação sexual, religião ou opinião política. A fim de permitir que as pessoas confiem no tratamento de dados, há que assegurar-lhes o pleno controlo sobre os seus próprios dados e a não utilização dos dados que lhes dizem respeito para as prejudicar ou discriminar. Além da proteção da privacidade e dos dados pessoais, devem ser cumpridos requisitos para garantir sistemas de IA de elevada qualidade. A qualidade dos conjuntos de dados utilizados é fundamental para o desempenho dos sistemas de IA. Os dados recolhidos estes podem refletir enviesamentos socialmente construídos ou conter inexatidões, erros e lapsos. Esta questão deve ser abordada antes da formação de um sistema de IA com qualquer conjunto de dados. Além disso, deve ser assegurada a integridade dos dados. Os processos e conjuntos de dados utilizados devem ser testados e documentados em cada fase, como as do planeamento, da formação, dos ensaios e da implantação. Este requisito deve aplicar-se igualmente aos sistemas de IA que não sejam desenvolvidos internamente, mas adquiridos noutros locais. Por último, o acesso aos dados deve ser adequadamente regulado e controlado.

IV. Transparência: Deve ser assegurada a rastreabilidade dos sistemas de IA; é importante registar e documentar as decisões tomadas pelos sistemas, bem como todo o processo (incluindo uma descrição da recolha e rotulagem de dados e uma descrição do algoritmo utilizado) que deu origem às decisões. Por este motivo, deve ser prestada, na medida do possível, uma explicação do processo de tomada de decisões algorítmicas, adaptado às pessoas envolvidas. Devem ser prosseguidos os trabalhos de investigação em curso com vista ao desenvolvimento de mecanismos de explicação. Além disso, devem estar disponíveis explicações sobre o grau de influência e orientação de um sistema de IA no processo de tomada de decisão organizacional, as escolhas de conceção do sistema, bem como a justificação da sua implantação (garantindo, assim, não só a transparência dos dados e do sistema, mas também a transparência do modelo empresarial). Por último, é importante comunicar adequadamente as capacidades e limitações do sistema de IA às diferentes partes interessadas envolvidas, de uma forma adequada ao caso em concreto. Além disso, os sistemas de IA devem ser identificáveis como tal, assegurando que os utilizadores sabem que estão a interagir com um sistema de IA e que pessoas são responsáveis por este sistema.

V. Diversidade, não discriminação e equidade: Os conjuntos de dados utilizados pelos sistemas de IA (tanto para a aprendizagem como para o funcionamento) podem ser afetados pela inclusão inadvertida de modelos de má conduta, incompletos e de má governação. A subsistirem, tais enviesamentos poderiam conduzir à discriminação (in)direta. Podem resultar danos também da exploração intencional de preconceitos (dos consumidores) ou da concorrência desleal. Além disso, o modo de desenvolvimento dos sistemas de IA (por exemplo, a forma de escrita do código de programação de um algoritmo) também pode ser tendencioso. Estas preocupações devem ser abordadas desde o início do desenvolvimento do sistema. A constituição de equipas de conceção diversas e a criação de mecanismos que garantam a participação, em especial dos cidadãos, no desenvolvimento da IA também podem ajudar a dar resposta a estas preocupações. É aconselhável consultar as partes interessadas que possam ser direta ou indiretamente afetadas pelo sistema ao longo de

todo o seu ciclo de vida. Os sistemas de IA devem ter em conta toda a gama de capacidades, competências e requisitos humanos, bem como garantir a acessibilidade através de uma abordagem de conceção universal, a fim de assegurar a igualdade de acesso das pessoas com deficiência.

VI. Bem-estar societal e ambiental: Para que a IA seja de confiança, há que ter em conta o seu impacto sobre o ambiente e outros seres sencientes. Idealmente, todos os seres humanos, incluindo as gerações futuras, devem beneficiar da biodiversidade e de um ambiente habitável. Devem, por conseguinte, ser incentivadas a sustentabilidade e a responsabilidade ecológica dos sistemas de IA. O mesmo se aplica às soluções de inteligência artificial que abordam domínios de interesse global, como, por exemplo, os Objetivos de Desenvolvimento Sustentável das Nações Unidas. Além disso, o impacto dos sistemas de IA deve ser considerado não apenas numa perspetiva individual, mas também na perspetiva de toda a sociedade. A utilização de sistemas de IA deve merecer uma atenção especial em situações relacionadas com o processo democrático, incluindo a tomada de decisões políticas e os contextos eleitorais. Além disso, há que ter em conta o impacto social da IA. Embora os sistemas de IA possam ser utilizados para reforçar as competências sociais, podem contribuir igualmente para a sua deterioração.

VII. Responsabilização: Devem ser criados mecanismos para assegurar a responsabilidade e a responsabilização dos sistemas de IA, bem como dos seus resultados, tanto antes como depois da sua aplicação. A auditabilidade dos sistemas de IA é fundamental neste contexto, uma vez que a avaliação dos sistemas de IA por auditores internos e externos, bem como a disponibilidade de tais relatórios de avaliação, contribuem fortemente para a fiabilidade da tecnologia. A auditabilidade externa deve ser assegurada, em especial, nas aplicações que afetam os direitos fundamentais, incluindo as aplicações críticas para a segurança. Os potenciais impactos negativos dos sistemas de IA devem ser identificados, avaliados, documentados e minimizados. A utilização de avaliações de impacto facilita este processo. Estas avaliações devem ser proporcionais à dimensão dos riscos que os sistemas de IA representam. Os compromissos entre os requisitos, muitas vezes inevitáveis, devem ser abordados de uma forma racional e metodológica, devendo ser tidos em conta. Por último, prever-se mecanismos acessíveis, que garantam uma reparação adequada, para as ocorrências de impactos adversos e injustos.

Portanto, para a UE, a confiança é um pré-requisito para garantir uma abordagem da IA centrada no ser humano, visto que a tecnologia não é um fim em si mesmo, mas um instrumento a serviço da pessoa humana com o objetivo de aumentar o seu bem-estar, fundamentando-se nos valores do respeito pela dignidade humana, liberdade, democracia, igualdade, Estado de Direito e respeito pelos direitos humanos²⁴⁶.

Sem embargo, tal como no caso das princípios da OCDE e a Declaração do G20, as Diretrizes de Ética para Inteligência Artificial Confiável carecem de força vinculante, o que prejudica a sua implementação na prática, visto que os desenvolvedores e usuários de sistemas de IA não estão obrigados a cumpri-las, além de serem por demais gerais e teóricas.

4.1.4 Referências nacionais de princípios reitores da IA

²⁴⁶ UNIÃO EUROPEIA. Aumentar a confiança numa inteligência artificial centrada no ser humano, COM (2019) 168 final. Bruxelas, 8 abr. 2019. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52019DC0168&from=EN>. Acesso em: 12 fev. 2023.

Em relação às referências de princípios que regem a IA no Brasil, conforme fora visto, ainda não há um marco legal sobre sistemas inteligentes. Sem embargo a EBIA, alinhada às diretrizes da OCDE, além de nortear as ações do estado brasileiro no sentido de estimular a pesquisa, a inovação e o desenvolvimento de soluções em IA, também prevê como ação estratégica “*estabelecer valores éticos para uso da IA na Administração Pública Federal*”²⁴⁷.

O relatório final da comissão de Juristas instituída para criação de um Marco Legal de IA, pode ser considerado uma importante e atual referência sobre princípios ético-jurídicos relacionado à inteligência computacional, visto que a comissão realizou um vasto trabalho, com a participação de juristas e especialista de renome nacional e internacional, além de audiências públicas e seminários²⁴⁸, conforme já fora mencionado. No PL apresentado pela Comissão constam os seguintes princípios²⁴⁹:

Art. 3º O desenvolvimento, implementação e uso de sistemas de inteligência artificial observarão a boa-fé e os seguintes princípios:

- I – crescimento inclusivo, desenvolvimento sustentável e bem-estar;
- II – autodeterminação e liberdade de decisão e de escolha;
- III – participação humana no ciclo da inteligência artificial e supervisão humana efetiva;
- IV – não discriminação;
- V – justiça, equidade e inclusão;
- VI – transparência, explicabilidade, inteligibilidade e auditabilidade;
- VII – confiabilidade e robustez dos sistemas de inteligência artificial e segurança da informação;
- VIII – devido processo legal, contestabilidade e contraditório;
- IX – rastreabilidade das decisões durante o ciclo de vida de sistemas de inteligência artificial como meio de prestação de contas e atribuição de responsabilidades a uma pessoa natural ou jurídica;
- X – prestação de contas, responsabilização e reparação integral de danos
- XI – prevenção, precaução e mitigação de riscos sistêmicos derivados de usos intencionais ou não intencionais e efeitos não previstos de sistemas de inteligência artificial; e
- XII – não maleficência e proporcionalidade entre os métodos empregados e as finalidades determinadas e legítimas dos sistemas de inteligência artificial.

Em que pese a extensa enumeração acima mencionada, questiona-se até que ponto convém realizar uma relação exaustiva de princípios aplicáveis ao contexto de IA. Além disso,

²⁴⁷ BRASIL. Ministério da Ciência, Tecnologia e Inovações/Gabinete do Ministro. Portaria GM Nº 4.617, DE 6 de Abril de 2021. Institui a Estratégia Brasileira de Inteligência Artificial e seus eixos temáticos. [S. l.], 12 abr. 2021. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-gm-n-4.617-de-6-de-abril-de-2021-*-313212172. Acesso em: 22 nov. 2021.

²⁴⁸ COMISSÃO DE JURISTAS RESPONSÁVEL POR SUBSIDIAR ELABORAÇÃO DE SUBSTITUTIVO SOBRE INTELIGÊNCIA ARTIFICIAL NO BRASIL. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 2002. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04>. Acesso em: 8 mar. 2023.

²⁴⁹ Ibid.

alguns dos princípios citados já se encontram contemplados em texto legal ou constitucional, tal como “VII - devido processo legal, contestabilidade e contraditório” previsto no art. 5º, inc. LV da CRFB.

Por outro lado, o marco legal de IA incorre no risco de omitir alguns princípios considerados basilares pela doutrina e jurisprudência, bem como de não se fazer os devidos temperamentos em relação à sua aplicabilidade à iniciativa privada ou ao Estado. Cite-se, por exemplo, a diretriz de disponibilidade de dados públicos, que é de observância obrigatória pelo Estado por força da LAI, excluindo-se, assim, a iniciativa privada, diretriz que será melhor trabalhada adiante.

Diante disso, os princípios previsto no PL de IA podem ser considerados de natureza meramente enumerativa, não prejudicando a esforço doutrinário e jurisprudencial no sentido de visualizar outros princípios e/ou diretrizes aplicáveis que não estejam expressamente contemplados em lei, tal como se proporá adiante na presente dissertação.

4.2 Proposta de diretrizes para regulação da IA na Administração Pública

À luz de todo o exposto, passamos a tecer algumas considerações sobre diretrizes que, na presente dissertação, são consideradas basilares e que servirão como parâmetro para órgãos e entidades da Administração Pública na elaboração da regulação setorial e *interna corporis* para o design e uso ético de IA.

Em que pese, conforme será visto, alguns desses parâmetros possam ser entendidos como autônomos e independentes do princípio da moralidade administrativa, é preciso recordar as lições de Wallace Paiva Martins Júnior já citado anteriormente, segundo o qual o princípio constitucional da moralidade administrativa é “um superprincípio, ou um princípio dos princípios”, devendo servir como alicerce chave hermenêutico-constitucional das diretrizes em relação à sua aplicação, em concreto, para casos relativos à adoção de sistemas inteligentes pela Administração Pública.

Também deve se levar em consideração que existem outras possibilidades de classificação ou mesmo de aglutinação dessas diretrizes, não havendo uma fórmula fixa ou universal, bastando para constatar isso as diferentes perspectivas e abordagens nacionais e internacionais antes mencionadas. Assim, a definição e organização dos parâmetros ora propostos pareceu a forma mais didática para os objetivos pretendidos na presente dissertação.

4.2.1 Excelência mediante o uso de IA

A adoção de uma perspectiva aristotélica de ética leva à adoção racional dos meios em vista à realização de um *telos* como um vetor fundamental do agir humano²⁵⁰. Trata-se de desenvolver hábitos operativos bons: virtudes ou *areté*, no sentido grego²⁵¹. A Administração Pública deve estimular o uso da IA como instrumento para atingimento da sua finalidade precípua de forma excelente: o atendimento ao interesse público, ao bem comum e à efetivação de direitos e garantias fundamentais.

Essa visão se contrapõe a uma perspectiva deontológica de regulação da IA, onde geralmente dá-se um peso excessivo à definição de limitadores e codificações, no dever e na obrigação, na definição do que “pode” ou “não pode” ser feito. O problema de uma perspectiva de “moral de obrigações” é que acaba perdendo-se de vista que o uso ético da tecnologia tem como consequência necessária extrair dela todas as suas potencialidades e benefícios visando à excelência no atingimento da finalidade da Administração.

Dentro de uma visão ética das virtudes, a adoção de IA pelo Estado é vista com bons olhos, até porque o Poder Público desempenha um papel fundamental no sentido de prover a infraestrutura necessária para o funcionamento e crescimento das inovações disruptivas no Brasil²⁵². Os investimentos em sistemas inteligentes devem ser direcionados especialmente naquelas atividades em que a inteligência computacional supera substancialmente o ser humano. Para ilustrar o argumento, apresentamos a seguir a resposta dada pelo sistema de IA *ChatGPT* ao ser questionado com a seguinte pergunta: “Quais são as atividades em que a Inteligência Artificial é comprovadamente superior à inteligência humana?”:

²⁵⁰ Cf. Capítulo III desta dissertação, sobre a relação entre direito e ética.

²⁵¹ “Areté era uma palavra utilizada no idioma grego que, entre outras coisas, trazia à tona certos traços da personalidade do homem. Sua correspondente mais usada na língua portuguesa, virtude, normalmente é definida como: “1. Disposição firme e habitual para a prática do bem. 2. Boa qualidade moral; força moral. 3. Ato virtuoso. 4. Qualidade própria para a produção de certos efeitos; eficácia” (VIRTUDE, 1996). Os ideais de virtude relacionados ao homem do campo em Hesíodo, ao homem político com os Sofistas e ao hábito orientado para o bem com Aristóteles teriam, todos, como precursores as sagas de Homero. Para compreendermos a ideia de areté, diferenciando-a das afirmações convencionais sobre virtude, é preciso nos situarmos no contexto da Grécia Antiga, por volta do século VIII a.C. Nessa época, segundo a tradição, teria vivido Homero, poeta grego a quem são atribuídas as epopeias *Ilíada* e *Odisseia*, tramas guerreiras repletas de ação, fruto da tradição oral compilada e remontada ao longo dos séculos. Em ambas se destaca a figura de um herói, sempre protagonista de grandes feitos corajosos e situado como intermédio entre os homens e os deuses, uma vez que, na maioria das vezes, é fruto do encontro amoroso entre as duas dimensões”. MARÇAL, Euclides Evangelista. ARETÉ E PAIDEIA: a excelência e a educação na era de Homero. [Pensamentoextemporaneo.com.br](https://pensamentoextemporaneo.com.br), [s. l.], 20 jan. 2013. Disponível em: <https://pensamentoextemporaneo.com.br/?p=2498>. Acesso em: 14 out. 2021.

²⁵² SADDY, André et al. *Aspectos Jurídicos das Novas Tecnologias (inovações) Disruptivas*. 2. ed. Rio de Janeiro: CEEJ, 2020, p. 26.

A inteligência artificial (IA) é particularmente boa em tarefas que envolvem grandes quantidades de dados, processamento repetitivo e velocidade. Algumas das áreas em que a IA é comprovadamente superior à inteligência humana incluem:

Cálculo e análise de grandes quantidades de dados: A IA pode processar grandes quantidades de dados muito mais rapidamente e de forma mais eficiente do que os seres humanos. Por exemplo, algoritmos de aprendizado de máquina podem analisar grandes quantidades de dados para identificar padrões e insights úteis em questão de minutos.

Jogos estratégicos: A IA pode superar os melhores jogadores humanos em jogos estratégicos complexos, como xadrez e Go. Os programas de IA usam algoritmos que analisam todas as possibilidades de movimentos futuros, permitindo que eles tomem as melhores decisões possíveis.

Reconhecimento de imagens: A IA pode identificar imagens e objetos com grande precisão e rapidez. Isso é especialmente útil em aplicações de vigilância e segurança, como câmeras de segurança que podem detectar pessoas e objetos suspeitos.

Processamento de linguagem natural: A IA pode processar grandes quantidades de texto e identificar padrões linguísticos para realizar tarefas como tradução automática, classificação de sentimentos e respostas a perguntas.

Previsão do tempo: A IA é frequentemente usada para prever o clima, com modelos de IA sendo capazes de analisar dados de sensores meteorológicos em tempo real para fazer previsões precisas.

Controle de tráfego aéreo: A IA é frequentemente usada para gerenciar o tráfego aéreo, com algoritmos de IA capazes de identificar e rastrear aeronaves em tempo real e ajudar a evitar colisões.

É importante notar que, embora a IA possa ser superior à inteligência humana em certas tarefas específicas, ela ainda não pode replicar a inteligência geral e a flexibilidade do pensamento humano. A inteligência artificial é uma ferramenta poderosa, mas é apenas uma das muitas ferramentas que podemos usar para melhorar a vida humana²⁵³.

No caso das Cidades Inteligentes (*Smart Cities*), vislumbram-se inúmeros benefícios decorrentes da adoção de IA, IoT, e tecnologia 5G, dentre outras soluções inovadoras, de forma a proporcionar o aperfeiçoamento dos serviços públicos existentes e soluções inovadoras, com ganhos em eficiência energética, mobilidade urbana, qualidade do ar e segurança pública, por exemplo. Proporciona-se, destarte, uma gestão urbana eficiente, sustentável, proativa e em tempo real; conectividade entre cidadãos e governo; e melhoria da qualidade de vida da população²⁵⁴.

Uma Cidade Inteligente caracteriza-se pela utilização de sensores inteligentes, com tecnologia IoT, instalados por toda a cidade, para coletar dados em tempo real nas mais diversas searas de interesse público, como tráfego urbano, qualidade do ar, consumo de energia, segurança pública, abastecimento de água, saneamento básico, bem como informações providas por cidadãos. As informações coletadas formam um *big data* que é processado e analisado por sistemas de IA, que detectam problemas, padrões e tendências e fazem recomendações,

²⁵³ TEIXEIRA, Raphael Lobato Collet Janny. Exemplos de Superioridade da IA. 18/02/2023. Conversa realizada com o assistente virtual ChatGPT. Disponível em: <https://chat.openai.com/chat/8ec0a4fc-607b-4cb4-9310-ce4ece0cab9>. Acesso em: 18 fev. 2023.

²⁵⁴ Cf. RATTI, Carlo; CLAUDEL, Matthew. *The City of Tomorrow: Sensors, Networks, Hackers, and the Future of Urban Life*. New Haven: Yale University Press, 2016.

proporcionando aos gestores públicos uma análise mais apurada para auxílio na tomada de decisão e implementando soluções baseadas em dados. No caso do controle de tráfego, por exemplo, sensores de trânsito coletam informações que ajudam a regular os temporizadores de semáforos, bem como orientar os motoristas em tempo real, melhorando o trânsito e reduzindo o congestionamento.

No caso do Brasil, deve-se fazer uma combinação entre EBIA e Plano Nacional de Internet das Coisas previsto no Decreto nº 9.854/2019²⁵⁵, de sorte a estimular, disseminar e orientar a formação de Cidades Inteligentes, proporcionando cidades eficientes, conectadas e sustentáveis, sem prejuízo de se buscar outras searas do serviço público em que IA e IoT interagem. O Plano Nacional de Internet das Coisas estabelece que saúde, cidades, indústria e rural são campos que deverão ser priorizados na implementação da política²⁵⁶.

Enfim, a diretriz de excelência motiva a adoção de mecanismos de estímulo ao uso de inteligência sintética pelo Poder Público, nos mais diversos campos de atuação da Administração, como: investimentos em pesquisa e desenvolvimento, promoção de ecossistemas e de políticas de estímulo à implantação de sistemas de IA, melhoria da eficiência e inovação na prestação de serviços públicos, economicidade e controle de gastos, transparência e combate à corrupção.

4.2.2 Centralidade do ser humano

A centralidade do ser humano (*human-centric AI*) é decorrência necessária da dignidade da pessoa humana, sendo considerado em princípios e declarações internacionais como o principal vetor da utilização ética de IA, conforme pudemos observar nos Princípios da OCDE sobre IA e na Declaração do G20 sobre Princípios para IA Centrada nos Humanos. Nas palavras do saudoso Danilo Doneda²⁵⁷, proferidas quando da instalação da Comissão de

²⁵⁵ BRASIL. Decreto nº 9854, de 25 de junho de 2019. Institui o Plano Nacional de Internet das Coisas e dispõe sobre a Câmara de Gestão e Acompanhamento do Desenvolvimento de Sistemas de Comunicação Máquina a Máquina e Internet das Coisas. [S. l.], 26 jun. 2019. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/d9854.htm. Acesso em: 4 mar. 2023.

²⁵⁶ Ibid, art. 4º

²⁵⁷ DONEDA, Danilo Cesar Maganhoto. Apud. COMISSÃO DE JURISTAS RESPONSÁVEL POR SUBSIDIAR ELABORAÇÃO DE SUBSTITUTIVO SOBRE INTELIGÊNCIA ARTIFICIAL NO BRASIL. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil.. Brasília, 2002. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04>. Acesso em: 8 mar. 2023.

Juristas para elaboração do PL de IA:

“[...] a centralidade do elemento humano [que] deve ser enfatizada em todos os seus aspectos, em todos os pontos necessários. Nossa missão certamente será propor e pesquisar instrumentos e soluções regulatórias que não contradigam e não impeçam a aplicação da inteligência artificial, que pode ser, inclusive, necessária, mas que facilitem e incentivem a sua adoção, diminuindo riscos e garantindo a segurança jurídica [...]”.

Trata-se de uma diretriz fundamental, que acaba influenciando ou produzindo derivações em outras diretrizes, conforme será visto de forma subsequente. A centralidade do ser humano materializa-se, em apertada síntese, na primeira lei de Asimov antes citada: “um robô não pode ferir um ser humano ou, por inação, permitir que um ser humano venha a ser ferido”.

Significa dizer que a IA deve ser direcionada ao bem-estar social e individual da pessoa humana. Deve não só respeitar, mas ser um instrumento de efetivação de direitos e garantias fundamentais. Não pode, por ação ou omissão, causar danos aos administrados. Deve ser implantada com responsabilidade social privilegiar o *Human Augmentation*, ao mesmo tempo em que cria novas oportunidades de trabalho.

Trata-se, assim, de compreender os sistemas inteligentes não como um fim em si mesmo, ou simplesmente como uma ferramenta para aumento de eficiência produtiva e redução de custos: os sistemas de IA devem estar a serviço do bem integral da pessoa singular, da sociedade humana e do meio ambiente, ou, tal como consta nas Diretrizes de Ética para Inteligência Artificial Confiável da União Europeia acima mencionados, “devem agir como facilitadores de uma sociedade próspera e equitativa, apoiando a atividade humana e promovendo os direitos fundamentais, e não reduzindo, limitando ou guiando a autonomia humana”.

A implantação de inteligência computacional com responsabilidade social encontra, outrossim, guarida constitucional entre os direitos sociais previstos no artigo 7º da CRFB, que garante em seu inciso XXVII a proteção em face da automação, na forma da lei, como um direito dos trabalhadores urbanos e rurais²⁵⁸. Conseqüentemente, um servidor público não poderá ser simplesmente substituído por um robô e descartado. Os gestores da Administração Pública devem proporcionar a adequada capacitação dos seus servidores na utilização de sistemas inteligentes. Por outro lado, os servidores poderão ser realocados para realizar atividades estratégicas e que exijam atributos inerentes à pessoa humana e que a IA possui um desempenho menor ou até inexistente, como intuição, empatia, criatividade, flexibilidade,

²⁵⁸ Constituição da República Federativa do Brasil de 1988. Art. 7º, XXVII. [S. l.: S. n.], 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 7 dez. 2021.

sopesamento de valores, pensamento crítico e análise casuística. Em todo caso, os gestores da Administração devem assegurar uma transição justa, segura e adequada.

Deve-se levar em consideração, ainda, que a simbiose entre inteligência humana e artificial (*Human Augmentation*) pode ser a melhor opção para obtenção de um resultado excelente. Garry Kasparov²⁵⁹ ao ser derrotado pela IA da IBM, desenvolve reflexão neste sentido, ressaltando, ao comentar a lenda de John Henry, que não devemos temer as máquinas inteligentes, mas trabalhar com elas. A inteligência computacional funcionaria como uma espécie de “virtude” ou “segunda natureza” agindo de forma a complementar, amplificar e aperfeiçoar as capacidades humanas. No caso da aplicação de Inteligência Operacional, por exemplo, a IA relaciona dados de milhares de fontes diferentes e apresenta as informações geradas de maneira visual e simplificada por meio de painel interativo, identificando comportamentos anômalos com alertas inteligentes e prevendo possíveis falhas. Os operadores, assim, identificam os alertas e tomam decisões de maneira assertiva. Com a utilização de Inteligência Operacional, empresas têm obtido resultados impressionantes de aumento na produtividade e *compliance* em relação ao cumprimento de leis²⁶⁰.

No direito europeu, relativamente ao tratamento automatizado de dados, o GDPR estipula no Considerando 71 o seguinte:

(71) O titular dos dados deverá ter o direito de não ficar sujeito a uma decisão, que poderá incluir uma medida, que avalie aspetos pessoais que lhe digam respeito, que se baseie exclusivamente no tratamento automatizado e que produza efeitos jurídicos que lhe digam respeito ou o afetem significativamente de modo similar, como a recusa automática de um pedido de crédito por via eletrónica ou práticas de recrutamento eletrónico sem qualquer intervenção humana. Esse tratamento inclui a definição de perfis mediante qualquer forma de tratamento automatizado de dados pessoais para avaliar aspetos pessoais relativos a uma pessoa singular, em especial a análise e previsão de aspetos relacionados com o desempenho profissional, a situação económica, saúde, preferências ou interesses pessoais, fiabilidade ou comportamento, localização ou deslocações do titular dos dados, quando produza efeitos jurídicos que lhe digam respeito ou a afetem significativamente de forma similar. No entanto, a tomada de decisões com base nesse tratamento, incluindo a definição de perfis, deverá ser permitida se expressamente autorizada pelo direito da União ou dos Estados-Membros aplicável ao responsável pelo tratamento, incluindo para efeitos de controlo e prevenção de fraudes e da evasão fiscal, conduzida nos termos dos regulamentos, normas e recomendações das instituições da União ou das entidades nacionais de controlo, e para garantir a segurança e a fiabilidade do serviço prestado pelo responsável pelo tratamento, ou se for necessária para a celebração ou execução de um contrato entre o titular dos dados e o responsável pelo tratamento, ou mediante o consentimento explícito do titular. Em qualquer dos casos, tal tratamento deverá ser acompanhado das garantias adequadas, que deverão incluir a informação específica

²⁵⁹ KASPAROV, Garry. Palestra proferida no TED Taks, Vancouver Bc, abr. 2017. Disponível em: https://www.youtube.com/watch?v=NP8xt8o4_5Q. Acesso em: 30 out. 2021.

²⁶⁰ BORELLA, Augusto. Inteligência Operacional auxilia empresas no cumprimento de tarefas e regulações. Inforchannel, página web, 11 nov. 2021. Disponível em: <https://inforchannel.com.br/2021/11/11/inteligencia-operacional-auxilia-empresas-no-cumprimento-de-tarefas-e-regulacoes/>. Acesso em: 25 nov. 2021.

ao titular dos dados e o direito de obter a intervenção humana, de manifestar o seu ponto de vista, de obter uma explicação sobre a decisão tomada na sequência dessa avaliação e de contestar a decisão. Essa medida não deverá dizer respeito a uma criança.

Dando concretude ao referido considerando, o GDPR estipula que as pessoas singulares têm o direito de não ficar sujeitas a nenhuma decisão tomada exclusivamente com base em sistemas automatizados, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar, com algumas ressalvas²⁶¹.

4.2.3 Segurança e Precaução no uso de sistemas inteligentes

Com relação à segurança dos sistemas de IA, os órgãos e entidades da Administração devem adotar mecanismos que comprovadamente assegurem a sua confiança e robustez, *by design*, garantindo o seu correto funcionamento durante todas as fases do ciclo de vida da tecnologia. As medidas de segurança a serem adotadas devem levar em consideração a integridade física e mental da pessoa humana.

Conforme tratado nos aspectos técnicos da IA, pesquisas afirmam que a instabilidade é o calcanhar de Aquiles da IA moderna, visto que uma rede neural inerentemente confiável não pode ser construída e, por conseguinte, sistemas inteligentes seriam confiáveis apenas em áreas específicas e usando métodos específicos²⁶². O problema de instabilidade do sistema pode se tornar um grande perigo quando utilizado em situações de alto risco de danos, como diagnóstico

²⁶¹ “art. 22.1. 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.

2. O n. 1 não se aplica se a decisão:

a) For necessária para a celebração ou a execução de um contrato entre o titular dos dados e um responsável pelo tratamento;

b) For autorizada pelo direito da União ou do Estado-Membro a que o responsável pelo tratamento estiver sujeito, e na qual estejam igualmente previstas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular dos dados; ou

c) For baseada no consentimento explícito do titular dos dados”.

[...]

4. As decisões a que se refere o n. 2 não se baseiam nas categorias especiais de dados pessoais a que se refere o artigo 9.o, n. 1, a não ser que o n. 2, alínea a) ou g), do mesmo artigo sejam aplicáveis e sejam aplicadas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular”.

In: União Europeia. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho e do Conselho. Relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679&qid=1677953718324&from=PT>. Acesso em: 04 mar. 2022.

²⁶² Paradoxo matemático expõe limites da Inteligência Artificial. *Inovação Tecnológica*, 28 mar. 2022. Disponível em: <https://www.inovacaotecnologica.com.br/noticias/noticia.php?artigo=paradoxo-matematico-expoe-limites-inteligencia-artificial&id=010150220328#.Y18AMXbMKU1>. Acesso em: 30 out. 2022.

de doenças e veículos autônomos. Segundo explicou Matthew Colbrook²⁶³, um dos autores do artigo:

O problema está nas áreas em que você precisa de garantia, porque muitos sistemas de IA são uma caixa preta," [...] Em algumas situações, está tudo bem que uma IA cometa erros, mas ela precisa ser honesta sobre isso. E não é isso que estamos vendo para muitos sistemas - não há como saber quando eles estão mais confiantes ou menos confiantes sobre uma decisão.

Deve-se, assim, adotar cuidados de sorte a evitar falhas de percepção da realidade com potencial de grandes danos, tal como ocorrera nos acidentes do Boeing 737 MAX que, conforme já fora mencionado, teve a sua autorização de voo suspensa em todo o mundo entre março de 2019 a novembro de 2020, depois que se identificou que um defeito no sistema de controle de voo causou a queda de duas aeronaves. Por esta razão, é recomendado adicionalmente que os sistemas possuam um plano de contingência para o caso de falhas. Como medida de precaução, recomenda-se realização prévia de estudo de impacto de sistemas de IA, conforme segue.

A precaução é um relevante parâmetro para regulação de IA, especialmente a partir da definição do grau de riscos de danos ou do impacto a direitos fundamentais causados por potenciais malefícios, de sorte a se adotar medidas *ex ante* de proteção. A precaução no uso de sistemas inteligentes enseja a necessidade de prévia avaliação para identificação de potenciais riscos, bem como adoção de medidas para sua eliminação ou minimização. Assegura-se, assim, que um sistema de IA esteja eticamente adequado desde o seu desenvolvimento (*ethics by design*).

A origem deste instituto remonta ao direito ambiental. Uma das suas primeiras formulações ocorreu na Declaração do Rio sobre Meio Ambiente e Desenvolvimento de 1992 (Rio 92), onde se declarou que uma abordagem precaucionária deveria ser adotada pelos países para a proteção do meio ambiente²⁶⁴. Na seara ambiental, a legislação brasileira prevê instrumentos de natureza precaucionária, como o Estudo de Impacto Ambiental (EIA) e o Relatório de Impacto Ambiental (RIMA) para avaliação dos impactos ambientais de empreendimentos e atividades potencialmente poluidoras.²⁶⁵

²⁶³ Ibid.

²⁶⁴ BIONI, Bruno Ricardo; LUCIANO, Maria. O princípio da precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada. In: FRAZÃO, Ana (coord.); MULHOLLAND, Caitlin (coord.). Revista de Direito da Inteligência Artificial, Ética e Regulação. São Paulo: Revista dos Tribunais, 2019. p. 207-228.

²⁶⁵ Cf. A Lei n.º 6.938/1981, que instituiu a Política Nacional do Meio Ambiente e Lei n.º 9.605/1998, Lei de Crimes Ambientais.

Sobre a aplicação do princípio da precaução na regulação de IA, Bruno Ricardo Bioni e Maria Luciano ressaltam o seguinte²⁶⁶:

O princípio da precaução fornece um substrato importante para se pensar medidas e estratégias de regulação de IA, notadamente como lidar com situações de riscos de danos ou de desconhecimento dos potenciais malefícios e benefícios desse tipo de tecnologia. A automatização de processos de tomadas de decisão, a partir do emprego de IA, não deve se constituir como um argumento ingênuo em defesa de sua objetividade e neutralidade. Tais circuitos decisórios carregam escolhas das entidades e pessoas envolvidas na sua construção, sendo modulado pela agenda política e aspectos socioeconômico, de forma implícita ou explícita, que lhes são subjacentes (Data & Society, 2018).

Diante disso, o princípio da precaução apresenta dois vetores de regulação que merecem atenção: a) a abertura do debate regulatório a todos os atores envolvidos na implementação dessa tecnologia (e nas escolhas que ela impõe), de desenvolvedores àqueles que sofrerão seus possíveis efeitos, o que é um requisito obrigatório de um sistema democrático com históricas dinâmicas de assimetria de poder e informação; b) a atribuição de obrigações que reduzam as incertezas quantos aos benefícios e riscos em questão, de sorte a determinar a adoção ou não de IA.

Conforme fora visto no tópico referente à análise do Relatório da Comissão de Juristas, no PL de IA optou-se por adotar mecanismos de avaliação prévia de riscos, com uma carga regulatória e obrigacional proporcional aos impactos que um sistema de IA possa causar a direitos e garantias fundamentais em função do contexto de aplicação da tecnologia, de sorte a mitigar ou eliminar os riscos identificados.

Assim, para se dar efetividade ao princípio da precaução, dentro de uma governança e da regulação setorial de IA, propõe-se que os órgãos e entidades da Administração Pública façam a utilização combinada de três ferramentas, a saber: a) elaboração de um relatório de impacto de IA; b) um *framework* de avaliação de riscos a direitos; e, a partir dos documentos gerados em “a” e “b”, c) definir as medidas mitigadoras e/ou eliminadoras dos riscos e potenciais danos identificados.

Em relação ao relatório de impacto de IA, sugere-se recorrer às disposições da legislação de proteção de dados e privacidade, especialmente o Relatório de Impacto de Proteção de Dados, documento que será analisado mais detalhadamente adiante.

Sobre o *framework* de avaliação de riscos a direitos, há diversos modelos específicos propostos por entidades com o objetivo de assegurar que um sistema seja utilizado de forma ética e confiável. Só para citar alguns exemplos, cita-se o “*Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities*” da U.S. Government

²⁶⁶ BIONI, Bruno Ricardo; LUCIANO, Maria. O princípio da precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada. In: FRAZÃO, Ana (coord.); MULHOLLAND, Caitlin (coord.). *Revista de Direito da Inteligência Artificial, Ética e Regulação*. São Paulo: Revista dos Tribunais, 2019. p. 207-228.

Accountability Office - GAO²⁶⁷, para processos de negócio do governo; “*Guidance on the AI auditing framework*” do Information Commissioner's Office (ICO), autoridade nacional de proteção de dados do Reino Unido, que fornece um toolkit de suporte para auditoria²⁶⁸; “*Assessment List for Trustworthy Artificial Intelligence – (ALTAI)*” da Comissão Europeia, para avaliar se o sistema de IA atende às Diretrizes de Ética para Inteligência Artificial Confiável da União Europeia²⁶⁹; e “*Estrutura de avaliação de riscos a direitos e de transparência*”, da Transparência Brasil, voltada para garantir o controle social no desenvolvimento, aquisição e implementação de IA pelo Estado²⁷⁰.

4.2.4 Controle e Supervisão por seres humanos

A ideia de controle humano encontra-se bem expressada na segunda lei de Asimov antes citada: “um robô deve obedecer às ordens dadas por seres humanos, exceto nos casos em que tais ordens entrem em conflito com a Primeira Lei”. Pela referida diretriz, adotam-se medidas para garantir que um sistema de IA encontra-se subordinado aos seres humanos, que não atuem de forma autônoma e prejudicial ou que não se “rebelem”, assegurando-se a sua segurança e confiabilidade. Evitam-se, assim, situações como a ilustrada no filme de ficção científica “*2001: Uma Odisséia no Espaço*”²⁷¹, onde HAL 9000, um supercomputador que controla a nave espacial Discovery One, rebela-se contra a sua tripulação, matando vários astronautas.

A supervisão, por sua vez, é especialmente relevante nos casos de sistemas opacos de IA, provendo meios, condições e informações para a sua sindicabilidade, ou seja, que sejam passíveis de avaliação por auditores internos e externos, mormente em relação ao parâmetros ou variáveis utilizadas pelos desenvolvedores, bem como as bases de dados utilizadas.

A supervisão garante, ainda, que o sistema inteligente não seja utilizado para prejudicar terceiros ou praticar atos ilícitos, como no caso dos *deepfakes* ou *synthetic reality*, onde se

²⁶⁷ Cf. GOVERNMENT ACCOUNTABILITY OFFICE - GAO. Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities. [S. l.], 2021. Disponível em: <https://www.gao.gov/assets/gao-21-519sp.pdf>. Acesso em: 09 mar. 2023.

²⁶⁸ Cf. INFORMATION COMMISSIONER'S OFFICE (ICO). Guidance on the AI auditing framework: Draft guidance for consultation. [S. l.], 2020. Disponível em: <https://ico.org.uk/media/2617219/guidance-on-the-ai-auditing-framework-draft-for-consultation.pdf>. Acesso em: 9 mar. 2023.

²⁶⁹ Cf. HIGH-LEVEL EXPERT GROUP ON AI. Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment. [S. l.], 2020. Disponível em: <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>. Acesso em: 9 mar. 2023.

²⁷⁰ Cf. TRANSPARÊNCIA BRASIL. Estrutura de Avaliação de Riscos, 2020. Disponível em: https://www.transparencia.org.br/downloads/publicacoes/Estrutura_Avaliacao_Risco.pdf. Acesso em: 9 mar. 2023.

²⁷¹ KUBRICK, Stanley. 2001: Uma Odisséia no Espaço. Produção da Metro-Goldwyn-Mayer (MGM), 1968. 1 filme (148 min), son., color.

produzem vídeos ou imagens com o objetivo de disseminar *fakenews*. Nessas situações, é importante que haja uma supervisão adequada, especialmente por autoridades públicas.

O Controle e supervisão por seres humanos devem ser diretamente proporcionais ao grau de potenciais prejuízos e violações aos direitos e interesses da pessoa singular, a partir dos quais a Administração Pública deverá definir os casos em que uma IA poderá funcionar: de i) forma autônoma, ii) com intervenção humana e iii) nos casos em que o seu uso será proibido, se representar um malefício insanável e inaceitável.

Com relação ao uso de IA de forma autônoma e sem intervenção humana, deverá ser restrita às atividades em que o risco de prejuízo aos direitos dos administrados seja inexistente ou muito reduzido. Aplicar-se-ia aos atos puramente burocráticos e organizativos, análise de dados anonimizados para fins estatísticos etc. Nestes casos poder-se-á utilizar algoritmos opacos, em que a explicação sobre o porquê da geração de um resultado nem sempre é possível, sem prejuízo de, em conformidade com o princípio da transparência que será adiante tratado, a Administração Pública fornecer informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada.

Nos casos em que seja identificado risco de prejuízo relevante a direitos e interesses dos administrados, recomenda-se que haja intervenção humana. A depender do grau de risco identificado, a intervenção poderá ocorrer sob a forma de revisão necessária da decisão automatizada ou pela utilização de IA como uma ferramenta auxiliar na tomada de decisão do servidor público (*Human Augmentation*). Conforme fora ressaltado no tópico referente à margem de livre apreciação do princípio da moralidade administrativa como conceito jurídico indeterminado, nos casos de zona de incerteza hermenêutica, conflito entre igualdade material e formal e nos juízos de equidade, é recomendável a adoção de mecanismos de revisão humana em decisões automatizadas.

Por fim, em casos extremos, poder-se-á cogitar, inclusive, da proibição da utilização de inteligência computacional pela Administração Pública, quando representarem um risco insanável para direitos fundamentais. Por exemplo, algumas entidades defendem que deveria haver uma proibição total do uso de IA para sistemas de pontuação social, categorização biométrica discriminatória e sistemas usados para prever atividades criminosas futuras²⁷².

²⁷² Neste sentido, os Direitos Digitais Europeus (EDRi) e 119 organizações da sociedade civil lançaram uma declaração coletiva para pedir uma Lei de Inteligência Artificial (AIA) que priorize os direitos fundamentais. In: EUROPEAN DIGITAL RIGHTS. An EU Artificial Intelligence Act for Fundamental Rights: A Civil Society Statement. Internet, 30 nov. 2021. Disponível em: <https://edri.org/wp-content/uploads/2021/11/Political-statement-on-AI-Act.pdf>. Acesso em: 5 dez. 2021.

4.2.5 Explicabilidade e Revisibilidade das decisões automatizadas

Um dos problemas éticos apontados na presente dissertação refere-se aos algoritmos opacos de Inteligência Artificial (*black box*), onde se questiona a obscuridade e falta de fundamentação de decisões automatizadas da Administração Pública que afetem direitos e liberdades fundamentais.

A explicabilidade tem como fundamento a necessidade de motivação das decisões administrativas. Sobre o tema, comenta André Saddy²⁷³:

“Hodiernamente, todo e qualquer ato administrativo, seja ele decorrente de uma manifestação de vontade sem ou com liberdade, isto é, vinculado ou com algum tipo de subjetividade ou autonomia pública, deve ser motivado, excetuando algumas situações. Tratar-se-á da divergência quanto à necessidade de motivação dos atos “discricionários” no Capítulo XXII - Atos Administrativos Unilaterais, mas, desde já, posiciona-se no sentido supra e, inclusive, concorda-se com aqueles que afirmam que a obrigação de motivar é diretamente proporcional à margem de atuação que a norma concede à Administração ou, o que é o mesmo, quanto maior a margem de liberdade, a intensidade ou profundidade da motivação do ato se incrementa.

Por conseguinte, as decisões administrativas que afetem direitos e interesses dos administrados devem ser fundamentadas de forma humanamente inteligível, ou seja, devem ser passíveis de interpretação e compreensão humana. Na utilização de um sistema automatizado, a obrigação de motivar deve ser proporcional não só à margem de liberdade concedida à Administração, mas levando-se em consideração, especialmente, os impactos a direitos e garantias fundamentais.

Ademais, a explicabilidade é condição *sine qua non* para a contestabilidade e revisibilidade de decisões, garantindo o exercício do contraditório e ampla defesa, direitos constitucionalmente assegurados nos processos administrativos²⁷⁴, permitindo, assim, que os administrados compreendam e eventualmente contestem as referidas decisões nas instâncias competentes.

A LGPD prevê o direito de revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem os interesses do titular da informação, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade²⁷⁵.

²⁷³ Cf. SADDY, André. Curso de direito administrativo brasileiro. 2. ed. Rio de Janeiro: CEEJ, 2023, v.1, p. 384.

²⁷⁴ art. 5º, inc. LV: “Aos litigantes, em processo judicial ou administrativo, e aos acusados em geral são assegurados o contraditório e ampla defesa, com os meios e recursos a ela inerentes”. BRASIL. [Constituição (1988)]. Constituição da República Federativa do Brasil de 1988. [S. l.: s. n.], 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 7 dez. 2021.

²⁷⁵ Cf. Ibid, art. 20.

Em sentido semelhante, o GDPR determina que, em casos em que se admite o tratamento automatizado de dados pessoais, o responsável pelo tratamento aplique medidas adequadas para salvaguardar os direitos, liberdades e legítimos interesses do titular dos dados, mormente o direito de solicitar uma revisão humana, manifestar o seu ponto de vista e contestar a decisão²⁷⁶.

A explicabilidade é relevante, também, para que o indivíduo possa exercer o direito de correção, atualização, bem como eliminação de dados desnecessários, excessivos ou tratados em desconformidade com o disposto na legislação de privacidade²⁷⁷.

Por fim, João Sergio dos Santos Soares Pereira²⁷⁸ entende que:

É preciso considerar que predições não envolvem conhecimento jurídico e servem para auxiliar. A correlação de palavras e o encontro de eventos não se iguala a devida fundamentação normativa contemporânea que se espera de atos decisórios que envolvem pessoas, sensibilidades e não apenas processos ou atos materiais da vida que esperam a solução advinda do Estado. As respostas advindas do Poder Público interferem na vida dos cidadãos que serão afetadas no emprego de quaisquer das etapas de implementação tecnológica no Poder Executivo.

Portanto, nos casos em que a opacidade é inerente à tecnologia de IA adotada, não é recomendável que as decisões administrativas sejam completamente substituídas por sistemas inteligentes.

4.2.6 Correção de vieses algorítmicos

Outro aspecto relevante a ser considerado como um corolário da centralidade e dignidade da pessoa humana relaciona-se à necessidade de constante verificação e correção de vieses algorítmicos (*bias*). Ao tratar da problemática relativa ao tema, fora apontado que o *bias* pode assumir diversas formas, como vieses estatísticos, sociais, técnicos e emergentes.

O algoritmo de IA não deve produzir resultados que sejam injustamente desiguais em

²⁷⁶ “Art. 22.3. Nos casos a que se referem o n. 2, alíneas a) e c), o responsável pelo tratamento aplica medidas adequadas para salvaguardar os direitos e liberdades e legítimos interesses do titular dos dados, designadamente o direito de, pelo menos, obter intervenção humana por parte do responsável, manifestar o seu ponto de vista e contestar a decisão”. In: União Europeia. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho. Relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679&qid=1677953718324&from=PT>. Acesso em: 04 mar. 2022.

²⁷⁷ Cf. art. 18, III e IV da LGPD.

²⁷⁸ PEREIRA, João Sergio dos Santos Soares. As decisões administrativas robóticas: das possibilidades aos limites. In: SADDY, André et al. Inteligência Artificial e Direito Administrativo. Rio de Janeiro: Centro para Estudos Empírico-Jurídicos (CEEJ), 2022.

razão de refletir preconceitos existentes na sociedade. Por esta razão, via de regra, as pessoas não podem ser desequiparadas com fundamento em dados sensíveis, ou seja, qualquer dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico²⁷⁹, bem como outras informações que podem gerar algum tipo de discriminação, tais como nacionalidade ou renda.

Ressalte-se, sem embargo, que o simples fato de se dar um tratamento desigual não reflete, necessariamente, uma injustiça ou preconceito. Na clássica frase de Rui Barbosa²⁸⁰ “A regra da igualdade não consiste senão em quinhoeir desigualmente aos desiguais, na medida em que se desigualem”. Celso Antônio Bandeira de Mello ressalta²⁸¹:

Há ofensa ao preceito constitucional da isonomia quando:

I — A norma singulariza atual e definitivamente um destinatário determinado, ao invés de abranger uma categoria de pessoas, ou uma pessoa futura e indeterminada.

II — A norma adota como critério discriminador, para fins de diferenciação de regimes, elemento não residente nos fatos, situações ou pessoas por tal modo desequiparadas. É o que ocorre quando pretende tomar o fator “tempo” — que não descansa no objeto — como critério diferencial.

III — A norma atribui tratamentos jurídicos diferentes em atenção a fator de discrimen adotado que, entretanto, não guarda relação de pertinência lógica com a disparidade de regimes outorgados.

IV — A norma supõe relação de pertinência lógica existente em abstrato, mas o discrimen estabelecido conduz a efeitos contrapostos ou de qualquer modo dissonantes dos interesses prestigiados constitucionalmente.

V — A interpretação da norma extrai dela distinções, discriminações, desequiparações que não foram professadamente assumidos por ela de modo claro, ainda que por via implícita.

Para identificação de um viés algorítmico ou desvio de poder praticado por inteligência computacional, é mister que haja, nas palavras de Vanice Regina Lírio do Valle²⁸², uma “avaliação do alinhamento entre a finalidade posta à Administração Pública pela regra de competência, e aquela efetivamente desenhada ou servida (ainda que hipoteticamente) pela aplicação de IA”.

Algumas boas práticas no sentido de identificar e corrigir vieses algorítmicos são a

²⁷⁹ Cf. art. 5º, II da LGPD.

²⁸⁰ BARBOSA, Rui. Oração aos Moços. Brasília: Senado Federal, 2019. Disponível em: https://www2.senado.leg.br/bdsf/bitstream/handle/id/564558/Oracao_aos_mocos_Rui_Barbosa.pdf. Acesso em: 27 fev. 2023. p. 36.

²⁸¹ MELLO, Celso Antônio Bandeira de. O conteúdo jurídico do princípio da igualdade. 3. ed. São Paulo: Malheiros editores, 2007. Disponível em: https://edisciplinas.usp.br/pluginfile.php/5367569/mod_resource/content/2/MELLO_Princ%C3%ADpio%20da%20Iguerdade.pdf. Acesso em: 27 fev. 2023. p. 47.

²⁸² VALLE, Vanice Regina Lírio do. Administração Pública, viés algorítmico e desvio de finalidade: existe conciliação possível? In: ZOCKUN, Maurício; GABARDO, Emerson. Direito Administrativo e Inovação: crises e soluções. Curitiba: Íthala, 2022. p. 503-515.

depuração da qualidade dos dados utilizados, a criação de sistemas de monitoramento e verificação de vieses, a transparência em relação ao *modus operandi* e dos dados utilizados para treinamento da IA, bem como disponibilização dos códigos-fonte do sistema para avaliação pela sociedade.

4.2.7 Transparência e Disponibilidade de dados públicos

Em relação à transparência sobre o uso de IA pela Administração Pública, deve-se dar publicidade e informar previamente sobre a adoção de sistemas inteligentes na interação com os administrados, bem como sobre decisões administrativas robóticas que afetem direitos. A transparência é especialmente relevante nos casos de decisões tomadas por algoritmos opacos, materializando-se na obrigação de dar publicidade sobre os parâmetros definidos no design e as bases de dados utilizadas para treinamento e alimentação do sistema.

No direito europeu, o GPDR assegura o direito de os indivíduos serem informados sobre a existência de decisões automatizadas, incluindo a definição de perfis, bem como informações úteis relativas à lógica subjacente, a importância e as consequências previstas de tal tratamento para o titular dos dados²⁸³.

Conforme preceituado na LGPD, a Administração Pública deve fornecer informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, salvo nos casos de segredos comercial e industrial, hipótese em que a Autoridade Nacional de Proteção de Dados poderá realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais.²⁸⁴ As referidas informações abarcam tanto a documentação do projeto de design da IA, com as fontes de dados que alimentam esses sistemas. Da mesma forma, a publicidade enseja a obrigação de disponibilidade de relatórios de avaliação de sistemas por auditorias internas e externas para verificação de ocorrência de vieses, por exemplo.

Por outro lado, tendo em vista que a acurácia do funcionamento de IA é diretamente proporcional à entrada de dados e informações, a Administração Pública deve possuir uma política de dados abertos, de forma a garantir a disponibilidade e acesso aos dados públicos,

²⁸³ Cf. União Europeia. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho. Relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679&qid=1677953718324&from=PT>. Acesso em: 04 mar. 2022.

²⁸⁴ Cf. parágrafos 1º e 2º do art. 20 da LGPD.

nos termos da LAI, em formatos que facilitem a utilização de computadores inteligentes, respeitado o disposto na LGPD e o sigilo das informações empresariais.

A EBIA lembra que a Política de Dados Abertos do Poder Executivo Federal encontra-se prevista no Decreto nº 8.771/2016, podendo servir para a alimentação de sistemas de Inteligência Artificial. Ressalta, por isso, a importância de diretrizes sobre o uso ético de dados abertos²⁸⁵. Acrescenta, ainda, o seguinte²⁸⁶:

[...]

A OCDE publicou em 2018 o relatório "Open Government Data Report", o qual destaca que a melhora do acesso a dados governamentais propicia oportunidades para inovadores governamentais e não-governamentais criarem novas formas de solucionar problemas de nossa sociedade. Isso implica o engajamento e participação de partes interessadas não institucionais, como do setor privado, academia, setor sem fins lucrativos, no processo de política de dados abertos. Dados abertos podem se tornar a "plataforma" que alimenta o desenvolvimento de aplicações e soluções úteis, ou seja, fica cada vez mais evidente o valor dos dados abertos em produzir benefícios econômicos ao setor público e à economia como um todo, ao facilitar novas oportunidades de negócios e ajudar indivíduos, empresas e o governo a tomar melhores decisões com base em mais informações disponíveis. O Governo Brasileiro segue em linha com tais direcionamentos. Não há dúvidas de que os benefícios advindos do uso de dados governamentais abertos, além de significantes, são transversais, na medida em que atingem áreas inicialmente não previstas. Promover o intercâmbio de dados abertos entre entidades da Administração Pública e entre estas e o setor privado, sempre com respeito ao direito à proteção de dados pessoais e ao segredo comercial.

Vemos, assim, que a maior disponibilidade de dados governamentais combinada com a utilização de IA abrem novos horizontes sobre otimização e precisão em relação à destinação de recursos públicos, bem como na identificação de deficiências e falhas na prestação de serviços. Ademais, trata-se de uma grande oportunidade para um controle mais eficaz da Administração Pública, bem como para o desenvolvimento de soluções inovadoras no interesse da sociedade, inclusive com maior participação de empresas, universidades e instituições não governamentais.

4.2.8 Privacidade e Segurança da Informação

A Administração Pública deve assegurar que os seus sistemas de IA estejam em

²⁸⁵ BRASIL. Ministério da Ciência, Tecnologia e Inovações/Gabinete do Ministro. Portaria GM nº 4.617, de 6 de abril de 2021. Institui a Estratégia Brasileira de Inteligência Artificial e seus eixos temáticos. [S. l.], 12 abr. 2021. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-gm-n-4.617-de-6-de-abril-de-2021-*-313212172. Acesso em: 22 nov. 2021.

²⁸⁶ Ibid.

conformidade com a legislação de privacidade em vigor desde a concepção (*privacy by design*). Também deve observar as normas e requisitos de segurança da informação, de forma a assegurar disponibilidade, integridade, confidencialidade e autenticidade de informações especialmente daquelas consideradas sigilosas por razões de segurança da sociedade e do Estado, protegidas por sigilo legal ou que devem ser mantidas em segredo por razões empresariais (*trade secret*, segredo de negócio ou tudo que possa representar um diferencial competitivo), como no caso de empresas estatais exploradoras de atividades econômicas²⁸⁷. Deve-se, outrossim, adotar mecanismos de proteção lógicos e físicos que assegurem a segurança e resiliência de sistemas contra-ataques e acesso de terceiros não autorizados.

Com o advento da Lei Geral de Proteção de Dados, a proteção de dados das pessoas naturais foi alçada a um novo patamar, consagrando o direito à privacidade e à autodeterminação informativa como regras motrizes a serem observadas por instituições públicas e privadas que manuseiam dados pessoais. A LGPD visa criar uma cultura de privacidade e proteção anterior à ocorrência de um evento nocivo, a partir da adoção de um modelo *ex ante* de proteção, estimulando a adoção de medidas preventivas e protetivas no acesso e uso de informações das pessoas naturais. Com relação ao tratamento de dados pessoais pelo Poder Público, estabelece que deverá ser restrito à execução das competências legais de interesse público ou cumprimento das atribuições legais do serviço público²⁸⁸.

Constata-se, assim, que a LGPD apresenta um ferramental que pode ser aplicado à IA²⁸⁹. Tal como fora abordado no tópico referente à diretriz de Segurança e Precaução no uso de sistemas inteligentes, nas situações que podem gerar riscos às liberdades civis e aos direitos fundamentais, recomenda-se a elaboração do Relatório de Impacto de Proteção de Dados (RIPD), documentação que contém a descrição dos processos de tratamento de dados pessoais, bem como medidas, salvaguardas e mecanismos de mitigação de risco.²⁹⁰

Embora não haja uma exigência legal expressa obrigando a feitura do RIPD, trata-se de uma boa prática, visto que propicia um mapeamento de riscos e possíveis inconformidades relacionadas ao tratamento de dados pessoais por IA definindo medidas concretas para

²⁸⁷ Cf. a LAI.

²⁸⁸ Cf. art. 23 da LGPD.

²⁸⁹ BIONI, Bruno Ricardo; LUCIANO, Maria. O princípio da precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada. In: FRAZÃO, Ana (coord.); MULHOLLAND, Caitlin (coord.). *Revista de Direito da Inteligência Artificial, Ética e Regulação*. São Paulo: Revista dos Tribunais, 2019. p. 207-228.

²⁹⁰ Cf. art. 5, XVII da LGPD

eliminação ou mitigação de riscos. Além disso, a Autoridade Nacional de Proteção de Dados (ANPD)²⁹¹ poderá exigir a sua elaboração pelo Poder Público²⁹².

Com relação ao compartilhamento de dados pessoais entre membros e entidades da Administração Pública, o STF declarou a constitucionalidade parcial do Decreto 10.046/2019, que dispõe sobre o tema. As ações questionando a constitucionalidade foram ajuizadas, respectivamente, pelo Conselho Federal da Ordem dos Advogados do Brasil e pelo Partido Socialista Brasileiro, que alegavam que o referido Decreto violava a legislação de privacidade, além de representar uma ameaça de controle e vigilância totalitária por parte do Estado.

Por maioria dos votos, o Supremo decidiu que órgãos e entidades da Administração Pública Federal podem compartilhar dados pessoais entre si, desde que com a observância de alguns critérios. O voto condutor do relator, ministro Gilmar Mendes²⁹³, ressaltou que o Estado deve cumprir integralmente as garantias e os procedimentos estabelecidos na LGPD compatíveis com o setor público. Entre eles, citou mecanismos rigorosos de controle de acesso ao Cadastro Base do Cidadão, publicidade do compartilhamento ou do acesso a banco de dados pessoais e fornecimento de informações claras e atualizadas sobre previsão legal, finalidade e práticas utilizadas.

Nos casos em que órgãos e entidades públicas utilizarem dados de forma contrária aos parâmetros legais e constitucionais, o Tribunal concluiu que o Estado poderá acionar servidores e agentes políticos responsáveis por atos ilícitos, visando ao ressarcimento de eventuais danos²⁹⁴. Por fim, o STF determinou a reestruturação do Comitê Central de Governança de Dados pelo prazo de 60 dias, a partir da publicação da ata de julgamento.

Analisando o referido aresto, entende-se que o STF poderia ter deixado de forma clara e expressa a necessidade de consentimento prévio e inequívoco do cidadão para autorizar compartilhamento de dados pessoais entre órgãos e entidades da Administração Pública nos casos em que não há uma base legal, nos termos da LGPD, reafirmando, assim, o princípio da autodeterminação informativa, visto que essa questão não foi devidamente ventilada no acórdão.

No que tange à segurança da informação, o Decreto nº 9.637/2018, que institui a Política Nacional de Segurança da Informação (PNSI), no âmbito da Administração Pública Federal

²⁹¹ A ANPD é uma autarquia de natureza especial, dotada de autonomia técnica e decisória e patrimônio próprio responsável por zelar, implementar e fiscalizar o cumprimento desta Lei em todo o território nacional. Cf. arts. 5º, XIX e 55-A da LGPD.

²⁹² Cf. art. 32 da LGPD

²⁹³ BRASIL. Supremo Tribunal Federal. ADI 6649 e ADPF 695, rel. min. Gilmar Mendes, Ata nº 26, de 14/09/2022. Diário da Justiça Eletrônico (DJE) nº 191, divulgado em 23/09/2022.

²⁹⁴ Ibid.

dispõe que compete ao Gabinete de Segurança Institucional da Presidência da República (GSI), estabelecer os requisitos mínimos de segurança para o uso dos produtos que incorporem recursos de segurança da informação, de modo a assegurar a disponibilidade, a integridade, a confidencialidade e a autenticidade da informação e garantir a interoperabilidade entre os sistemas de segurança da informação, ressalvadas as competências específicas de outros órgãos²⁹⁵. Desta forma, os sistemas de IA desenvolvidos ou utilizados pela Administração Pública Federal devem observar as normas e diretrizes de segurança da informação emitidas pelo GSI.

4.2.9 Responsabilidade pelo adequado funcionamento

Conforme fora visto acima, o quinto Princípio da AI definido pela OCDE estabelece a obrigação de *accountability* de entidades e indivíduos no desenvolvimento, implantação e operação de sistemas de IA, com a consequente responsabilização por seu funcionamento adequado, observando-se todos os princípios anteriormente enumerados. Esse parâmetro também é destacado na declaração de Diretrizes de Ética para Inteligência Artificial Confiável da União Europeia, estabelecendo a necessidade de se criar mecanismos asseguradores da responsabilidade e da responsabilização dos sistemas de IA.

A diretriz de Responsabilidade pelo adequado funcionamento enseja a necessidade de a Administração Pública, ao desenvolver e utilizar IA, verificar de forma apriorística (*ex ante*) o atendimento aos princípios, diretrizes e normas definidos para o atendimento da moralidade administrativa, bem como, após a implantação de sistemas de IA (*ex post*), fazer revisões de conformidade de forma periódica.

Ademais, a Administração Pública deve assumir a responsabilidade pelas consequências decorrentes do uso de IA, inclusive por perdas e danos em razão de eventuais falhas, mal funcionamento e inconformidades legais e normativas.

No PL de IA, em relação à responsabilidade civil, a Comissão de Juristas apresentou uma proposta que engloba a responsabilidade do fornecedor e do operador por dano patrimonial, moral, individual ou coletivo, a depender do grau de risco identificado, mas

²⁹⁵ BRASIL. Decreto nº 9.637, de 26 de dezembro de 2018. Institui a Política Nacional de Segurança da Informação, dispõe sobre a governança da segurança da informação, e altera o Decreto nº 2.295, de 4 de agosto de 1997, que regulamenta o disposto no art. 24, caput, inciso IX, da Lei n.º 8.666, de 21 de junho de 1993, e dispõe sobre a dispensa de licitação nos casos que possam comprometer a segurança nacional. [S. 1.], 27 dez. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/decreto/D9637.htm. Acesso em: 28 fev. 2023. Art. 12, IX.

independentemente do grau de autonomia do sistema. Quando se tratar de sistema e IA de alto risco ou de risco excessivo, o fornecedor ou operador respondem objetivamente pelos danos causados, na medida da participação de cada um no dano. Por outro lado, quando se tratar de sistema que não seja de alto risco, a culpa do agente causador do dano será presumida, aplicando-se a inversão do ônus da prova em favor da vítima²⁹⁶. Por fim, a responsabilidade por danos causados por sistemas no âmbito das relações de consumo permanece sujeita ao Código de Defesa do Consumidor²⁹⁷.

4.3 Considerações finais

Conforme fora visto no capítulo anterior, afora a adoção da autorregulação pública setorial, recomenda-se que haja uma heterorregulação centralizada que estabeleça um sistema de governança na Administração Pública como um todo, com a função de coordenação e supervisão geral entre setores, inclusive, aproveitando-se de estruturas e procedimentos de governança já criados por força do disposto na Lei Geral de Proteção de Dados, de forma a se evitar superposições e conflitos regulatórios.

Essa heterorregulação pública poderia ser instituída mediante Decreto dos chefes dos Poderes Executivos, com fundamento no princípio da moralidade administrativa, estipulando uma governança geral de IA, a ser observada pelos órgãos e entidades inferiores ou subordinados. O Decreto possui a vantagem de maior flexibilidade para promover de forma mais célere as alterações, modificações e atualizações tão necessárias para atender aos desafios, novidades e questionamentos éticos que surgem diuturnamente, não sujeito, assim, às vicissitudes naturais do processo legislativo, mais rígido, moroso e incerto.

²⁹⁶“Art. 27. O fornecedor ou operador de sistema de inteligência artificial que cause dano patrimonial, moral, individual ou coletivo é obrigado a repará-lo integralmente, independentemente do grau de autonomia do sistema. § 1º Quando se tratar de sistema de inteligência artificial de alto risco ou de risco excessivo, o fornecedor ou operador respondem objetivamente pelos danos causados, na medida de sua participação no dano.

§ 2º Quando não se tratar de sistema de inteligência artificial de alto risco, a culpa do agente causador do dano será presumida, aplicando-se a inversão do ônus da prova em favor da vítima”. In: Comissão de Juristas Responsável por Subsidiar Elaboração de Substitutivo sobre Inteligência Artificial no Brasil. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei n.º 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 2002. 114 p. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04>. Acesso em: 8 mar. 2023.

²⁹⁷ Ibid., art. 29.: “As hipóteses de responsabilização civil decorrentes de danos causados por sistemas de inteligência artificial no âmbito das relações de consumo permanecem sujeitas às regras previstas na Lei n.º 8.078, de 11 de setembro de 1990 (Código de Defesa do Consumidor), sem prejuízo da aplicação das demais normas desta Lei”.

Por outro lado, conforme fora visto anteriormente, se o PL da Comissão de Juristas ou o PL n.º 21/2020 fossem aprovados tal como propostos, não haveria prejuízo às conclusões do presente estudo, visto que o legislador concederia liberdade de conformação ou de configuração para autorregulação pública setorializada. De qualquer maneira, seja qual for o PL sobre IA que venha a ser aprovado pelo Congresso, recomenda-se que haja um capítulo específico tratando da regulamentação da IA no âmbito da Administração Pública. Para tanto, sugere-se a seguinte redação:

CAPÍTULO XXX – Regulamentação da IA no âmbito da Administração Pública

Art. x. Os chefes dos Poderes Executivos deverão emitir um Decreto regulamentando o disposto na presente Lei, com o objetivo de instituir a governança, organização, coordenação e supervisão geral entre os órgãos e entidades da Administração Pública a eles subordinados, observando-se, em especial, as seguintes diretrizes:

- I) Excelência mediante o uso de IA;
- II) Centralidade do ser humano;
- II) Segurança e Prevenção;
- IV) Controle e Supervisão por seres humanos;
- V) Explicabilidade e Revisibilidade das decisões automatizadas;
- VI) Correção de vieses algorítmicos;
- VII) Transparência e Disponibilidade de dados públicos;
- VIII) Privacidade e Segurança da Informação;
- IX) Responsabilidade pelo adequado funcionamento.

§1 A estruturação da governança deverá prever um comitê responsável por assegurar a conformidade e observância dos princípios, diretrizes e normas prevista nesta Lei no âmbito da Administração Pública; realizar treinamentos e medidas de conscientização; auditorias e certificações nos órgãos e entidades subordinados.

§2 As atribuições previstas no parágrafo §1 poderão ser delegadas aos diferentes órgãos e entidades da Administração, mediante a criação de subcomitês.

§3 O comitê e subcomitês previstos nos parágrafos anteriores poderão possuir assentos destinados a especialistas e representantes da sociedade civil que agreguem à discussão sobre parâmetros éticos com uma perspectiva exógena à Administração.

Art. y. Os órgãos e entidades da Administração Pública deverão emitir autorregulações que estipulem a governança, organização, bem como normas e orientações de conduta específicas e concretas para cada contexto de implementação no âmbito interno, observando-se, em especial, as diretrizes mencionadas no Art. x.

§ 1 Os órgãos e entidades da Administração Pública deverão possuir um canal de comunicação entre a sociedade e a Administração Pública para emitir esclarecimentos públicos; receber solicitações e denúncias.

Art. z. As minutas de normas relativas à governança, políticas, diretrizes, e códigos de conduta relativos à IA deverão ser submetidos a um escrutínio público, mediante a criação de espaços abertos para consultas públicas, audiências e debates que estimulem a discussão e participação da sociedade civil na concretização normativa dos princípios éticos a serem observados na pesquisa, desenvolvimento e uso da IA no âmbito da Administração Pública.

CONCLUSÕES

Vivemos em uma época de acelerada transformação digital, com repercussões diretas em organizações públicas e privadas, tornando imprescindível o desenvolvimento de competências digitais e a adoção de novas tecnologias. Em relação à Administração Pública, vislumbram-se inúmeros benefícios decorrentes do uso de IA, com ganhos em economicidade e eficiência, inovação na prestação de serviços públicos, melhorias em políticas públicas e aperfeiçoamento dos mecanismos de controle da atividade estatal em relação à moralidade, legalidade e combate à corrupção.

A IA pode ser definida, de forma resumida, como um sistema computacional dotado de autoaprendizagem, capaz de alterar as instruções iniciais e até criar novas devido a técnicas de *machine learning*, bem como propor respostas de forma automática, sequer imaginadas por seus programadores originais.

Em que pese todos os benefícios oriundos do uso de IA, exurgiram debates e questionamentos de natureza ética, tendo em vista a constatação de diversos riscos relacionados a esta tecnologia, tais como assimetria informacional entre entidades públicas e privadas, de um lado, e a pessoa natural, de outro, e utilização de IA para previsão, controle e exploração de hipossuficiências emocionais e intelecto-volitivas com o objetivo de manipular ou induzir o comportamento humano como produto. Em relação ao Estado, a formação de *Big Datas* de dados pessoais, muitas vezes coletados compulsoriamente e compartilhado entre os órgãos e entidades da Administração, aliado ao uso de IA, podem representar riscos e ameaças ao Estado Democrático de Direito, bem como a direitos e garantias fundamentais, controle e vigilantismo estatal. Além disso, questiona-se a falta de transparência e opacidade de decisões automatizadas (*black box*), injustiças e vieses algorítmicos (*bias*), substituição da subjetividade do gestor pela “subjetividade” do robô e impedimento ou dificuldade de acesso a dados públicos.

Deve-se, assim, buscar o adequado equilíbrio entre o uso justo e ético da IA, centrado no ser humano e respeitando valores democráticos, direitos e garantias fundamentais, por um lado, e o estímulo à inovação e ao desenvolvimento, por outro. Para o estabelecimento da excelência na prestação de serviços públicos, é imprescindível levar-se em consideração que se trata de uma tecnologia cujas possibilidades ainda não foram plenamente desvendadas ou compreendidas.

O direito pode receber da ética dados ou influxos para resolver situações que dependem de uma avaliação de ordem moral. Sendo realidades indissociáveis, as leis não são neutras em

relação à moral. No direito administrativo, adota-se o princípio da moralidade administrativa, que norteia os atos do servidor sob o prisma do agir ético. O princípio da moralidade administrativa, positivado e alcançado ao status constitucional na CRFB, é um superprincípio, ou um princípio dos princípios. Está umbilicalmente ligado ao conceito de excelência e boa administração, ao elemento ético, à honestidade, ao interesse público e à noção de bem comum. O ato do servidor público deve conformar-se à moralidade para ser revestido de plena legalidade. Não basta, assim, que um ato administrativo esteja subsumido à lei do ponto de vista formal: é preciso também que seja praticado segundo os ditames da ética.

Recorrer à ética é especialmente relevante nas situações em que ainda não possam ser efetivamente objeto de regulação ou de atuação direta de leis, mas que, haja vista a relevância dos valores envolvidos, exigem uma resposta célere e adequada. Na Administração Pública, o princípio da moralidade administrativa é o fundamento e fonte da força jurígena para solução de questões de natureza ético-jurídicas relacionadas à adoção de IA.

A jurisprudência do STF é pacífica no sentido de reconhecer a autonomia e aplicabilidade imediata, pelos agentes da Administração, dos princípios previstos no art. 37, *caput*, da CRFB, independente de lei dispor a respeito. Na esteira da doutrina de Peter Häberle, as esferas administrativas inferiores podem ser consideradas como participantes do processo constitucional de construção do direito material, na medida que poderiam exercer a hermenêutica do princípio da moralidade administrativa.

Há evidentes limitações relacionadas ao ordenamento jurídico atual, construído para um “mundo analógico”, devendo ser modificado e complementado para atendimento das exigências da transformação digital.

Uma heterorregulação pública que seja geral e detalhista incorreria no risco de cair em rápida obsolescência, visto que se trata de tecnologia em constante e acelerada evolução, com inovações, funcionalidades e soluções sequer imaginadas pelo legislador. Além disso, é impossível dar tratamento legal de forma casuística a todas as situações em que a Administração pode se deparar com a necessidade de adotar ou regulamentar mecanismos automatizados, tais como regulação de atividades de particulares, uso interno da Administração, prestação de serviços públicos ou execução de políticas públicas. Por outro lado, uma heterorregulação pública principiológica não daria o devido tratamento a casos concretos relacionados ao cotidiano de quem desenvolve e usa IA, com risco de se tornar inócua e pouco efetiva.

Por outro lado, a concessão de poder de hermenêutica constitucional para definição do agir, em consonância com a moralidade administrativa, a quaisquer servidores públicos responsáveis pela programação, desenvolvimento, treinamento ou aplicação de um algoritmo

de IA certamente ensejaria maior insegurança jurídica: a ausência de um parâmetro normativo concreto e direcionador dificultaria sobremaneira a tarefa do agente, na medida em que é difícil vislumbrar como se concretiza o agir ético em determinadas situações e circunstâncias, especialmente em vista da novidade e especificidade dessas novas tecnologias. Corre-se, ainda, o risco de haver decisões equivocadas e contraditórias entre diferentes órgãos e entes da Administração.

No atual estágio de desenvolvimento da IA, parece recomendável que haja uma maior desconcentração e descentralização da sua regulação. A regulamentação interna e setorial da IA entre os diferentes órgãos e entidades da Administração responderia à necessidade de dar um tratamento jurídico mais célere, concreto e adequado aos questionamentos éticos decorrentes da adoção de máquinas inteligentes, evitando-se uma regulamentação legal açodada.

Diante disso, a solução aqui proposta está na autorregulação pública setorial, ou seja, os órgãos e entidades da Administração Pública revestidos de competência normativa e dentro da sua esfera de atribuição deverão recorrer ao princípio da moralidade administrativa para regulamentar o design e uso dos sistemas inteligentes, conduzindo os seus agentes para que atuem de forma ética e constitucionalmente orientada, respeitando direitos e garantias fundamentais, não só para coibir o mau uso ou desvio para finalidades escusas, mas também para estimular a adoção da tecnologia de forma excelente e direcionada para o interesse público e bem comum.

Pelo exercício da liberdade conformadora ou configuradora, o órgão ou ente administrativo competente deverá expedir normativo que regulamentará o design e uso ético da IA no âmbito de sua atividade. A liberdade de conformação ou de configuração não significa autonomia absoluta para criação de regras. A Administração Pública deve observância ao princípio da juridicidade, ou seja, a atuação administrativa está submetida ao Direito como um todo, e não unicamente à estrita legalidade administrativa. Além disso, os órgãos e entidades da Administração Pública, em geral, possuem normas que regem o processo interno de gênese normativa, atribuindo competência para elaboração, aprovação e alteração de atos normativos sobre matérias de sua atribuição e obrigando ao seu cumprimento àqueles que lhe estão jurisdictionados, sob pena de responsabilidade.

A liberdade configuradora do normativo a ser elaborado, tem, ainda, a sua autonomia restrita à margem de apreciação da moralidade administrativa enquanto conceito jurídico indeterminado, observando-se os limites próprios das zonas de certeza positiva, incerteza e

certeza negativa em relação à definição ou aplicabilidade de uma diretriz, norma ou conduta como derivados da moralidade administrativa.

Na esteira de Konrad Hesse, e em concerto com os princípios da correção funcional e da máxima efetividade, a interpretação adequada é aquela que consegue concretizar, de forma excelente, o sentido (*Sinn*) da proposição normativa dentro das condições reais dominantes numa determinada situação. No caso objeto de estudo, diante de novos riscos e situações derivados do avanço tecnológico, o intérprete deve buscar a concretização da renovação da força normativa do princípio constitucional da moralidade administrativa, sem ultrapassar, entretanto, a fronteira hermenêutica para este conceito jurídico indeterminado, ou seja, o seu sentido teleológico, sob pena de se autoconferir poder constituinte de forma oblíqua.

Para definição de diretrizes derivadas da moralidade administrativa, que deverão servir como norte para construção de normativos específicos de órgãos e entidades da Administração Pública, recorreu-se às fontes mais difundidas na atualidade e que geralmente são utilizadas como referência de países e organismos internacionais que pretendem construir um arcabouço normativo relativo à IA, a saber: i) as Três Leis da Robótica de Isaac Asimov; ii) Os Princípios sobre IA da OCDE; iii) a Declaração do G20 sobre Princípios para IA Centrada nos Humanos; e iv) As Diretrizes de Ética para Inteligência Artificial Confiável da União Europeia.

A presente dissertação propõe nove diretrizes como parâmetro para órgãos e entidades da Administração Pública na elaboração da regulação setorial e *interna corporis* para o *design* e uso ético de IA, a saber: i) Excelência mediante o uso de IA; ii) Centralidade do ser humano; iii) Segurança e Precaução; iv) Controle e Supervisão por seres humanos; v) Explicabilidade e Revisibilidade das decisões automatizadas; vi) Correção de vieses algorítmicos; vii) Transparência e Disponibilidade de dados públicos; viii) Privacidade e Segurança da Informação; e ix) Responsabilidade pelo adequado funcionamento.

A Administração Pública deve estimular o uso da IA como instrumento para atingimento da sua finalidade precípua de forma excelente: o atendimento ao interesse público, ao bem comum e à efetivação de direitos e garantias fundamentais; extraíndo dela todas as suas potencialidades e benefícios, especialmente naquelas atividades em que a inteligência computacional supera substancialmente o ser humano. A diretriz de Excelência motiva a adoção de mecanismos de estímulo ao uso de inteligência sintética pelo Poder Público, nos mais diversos campos de atuação da Administração, como: investimentos em pesquisa e desenvolvimento, promoção de ecossistemas e de políticas de estímulo à implantação de sistemas de IA, melhoria da eficiência e inovação na prestação de serviços públicos, economicidade e controle de gastos, transparência e combate à corrupção. No caso das Cidades

Inteligentes (*Smart Cities*), vislumbram-se inúmeros benefícios decorrentes da adoção de IA, IoT, e tecnologia 5G, dentre outras soluções inovadoras, de forma a proporcionar o aperfeiçoamento dos serviços públicos existentes e soluções inovadoras, com ganhos em eficiência energética, mobilidade urbana, qualidade do ar e segurança pública, por exemplo. Proporciona-se, destarte, uma gestão urbana eficiente, sustentável, proativa e em tempo real; conectividade entre cidadãos e governo; e melhoria da qualidade de vida da população

A centralidade no ser humano decorre da dignidade da pessoa humana, trata-se de uma diretriz fundamental, que acaba influenciando ou produzindo derivações em outras diretrizes. Materializa-se, em síntese, na primeira lei de Asimov: “um robô não pode ferir um ser humano ou, por inação, permitir que um ser humano venha a ser ferido”. Significa dizer que a IA deve ser direcionada ao bem-estar social e individual da pessoa humana. Deve não só respeitar, mas ser um instrumento de efetivação de direitos e garantias fundamentais. Não pode, por ação ou omissão, causar danos aos administrados. Deve ser implantada com responsabilidade social privilegiar o *Human Augmentation*, ao mesmo tempo em que cria novas oportunidades de trabalho. Trata-se, assim, de compreender os sistemas inteligentes não como um fim em si mesmo, ou simplesmente como uma ferramenta para aumento de eficiência produtiva e redução de custos: os sistemas de IA devem estar a serviço do bem integral da pessoa singular, da sociedade humana e do meio ambiente

Com relação à Segurança dos sistemas de IA, os órgãos e entidades da Administração devem adotar mecanismos que comprovadamente assegurem a sua confiança e robustez, *by design*, garantindo o seu correto funcionamento durante todas as fases do ciclo de vida da tecnologia. As medidas de segurança a serem adotadas devem levar em consideração a integridade física e mental da pessoa humana. A Precaução no uso de sistemas inteligentes, por sua vez, enseja a necessidade de prévia avaliação para identificação de potenciais riscos, bem como adoção de medidas para sua eliminação ou minimização. Assegura-se, assim, que um sistema de IA esteja eticamente adequado desde o seu desenvolvimento (*ethics by design*). Para se dar efetividade à diretriz de Precaução, dentro de uma governança e da regulação setorial de IA, propõe-se que os órgãos e entidades da Administração Pública façam a utilização combinada de três ferramentas, a saber: a) elaboração de um relatório de impacto de IA; b) um *framework* de avaliação de riscos a direitos; e, a partir dos documentos gerados em “a” e “b”, c) definir as medidas mitigadoras e/ou eliminadoras dos riscos e potenciais danos identificados.

A ideia de Controle humano encontra-se bem expressada na segunda lei de Asimov: “um robô deve obedecer às ordens dadas por seres humanos, exceto nos casos em que tais

ordens entrem em conflito com a Primeira Lei”. Pela referida diretriz, adotam-se medidas para garantir que um sistema de IA encontra-se subordinado aos seres humanos, que não atuem de forma autônoma e prejudicial ou que não se "rebelem", assegurando-se a sua segurança e confiabilidade. A Supervisão, por sua vez, é especialmente relevante nos casos de sistemas opacos de IA, provendo meios, condições e informações para a sua sindicabilidade, ou seja, que sejam passíveis de avaliação por auditores internos e externos, mormente em relação ao parâmetros ou variáveis utilizadas pelos desenvolvedores, bem como as bases de dados utilizadas. Garante, ainda, que o sistema inteligente não seja utilizado para prejudicar terceiros ou praticar atos ilícitos, como no caso dos *deepfakes* ou *synthetic reality*. Nessas situações, é importante que haja uma supervisão adequada, especialmente por autoridades públicas. O Controle e Supervisão por seres humanos devem ser diretamente proporcionais ao grau de potenciais prejuízos e violações aos direitos e interesses da pessoa singular, a partir dos quais a Administração Pública deverá definir os casos em que uma IA poderá funcionar: de i) forma autônoma, ii) com intervenção humana e iii) nos casos em que o seu uso será proibido, se representar um malefício insanável e inaceitável.

As decisões administrativas que afetem direitos e interesses dos administrados devem ser fundamentadas de forma humanamente inteligível, ou seja, devem ser passíveis de interpretação e compreensão humana. Na utilização de um sistema automatizado, a obrigação de motivar deve ser proporcional não só à margem de liberdade concedida à Administração, mas levando-se em consideração, especialmente, os impactos a direitos e garantias fundamentais. Além disso, a explicabilidade é condição *sine qua non* para a contestabilidade e revisibilidade de decisões, garantindo o exercício do contraditório e ampla defesa, direitos constitucionalmente assegurados nos processos administrativos, permitindo, assim, que os administrados compreendam e eventualmente contestem as referidas decisões nas instâncias competentes.

O algoritmo de IA não deve produzir resultados que sejam injustamente desiguais em razão de refletir preconceitos existentes na sociedade. Por esta razão, via de regra, as pessoas não podem ser desequiparadas com fundamento em dados sensíveis, ou seja, qualquer dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, bem como outras informações que podem gerar algum tipo de discriminação, tais como nacionalidade ou renda. Sem embargo, o simples fato de se dar um tratamento desigual não reflete, necessariamente, uma injustiça ou preconceito, visto que o fator de *discrimen* pode guardar relação de pertinência lógica e consonância com os

interesses prestigiados constitucionalmente. Algumas boas práticas no sentido de identificar e corrigir vieses algorítmicos são a depuração da qualidade dos dados utilizados, a criação de sistemas de monitoramento e verificação de vieses, a transparência em relação ao *modus operandi* e dos dados utilizados para treinamento da IA, bem como disponibilização dos códigos-fonte do sistema para avaliação pela sociedade.

Em relação à Transparência sobre o uso de IA pela Administração Pública, deve-se dar publicidade e informar previamente sobre a adoção de sistemas inteligentes na interação com os administrados, bem como sobre decisões administrativas robóticas que afetem direitos. A Transparência é especialmente relevante nos casos de decisões tomadas por algoritmos opacos, materializando-se na obrigação de dar publicidade sobre os parâmetros definidos no design e as bases de dados utilizadas para treinamento e alimentação do sistema. Por outro lado, tendo em vista que a acurácia do funcionamento de IA é diretamente proporcional à entrada de dados e informações, a Administração Pública deve possuir uma política de dados abertos, de forma a garantir a Disponibilidade e acesso aos dados públicos, nos termos da LAI, em formatos que facilitem a utilização de computadores inteligentes, respeitado o disposto na LGPD e o sigilo das informações empresariais. A maior disponibilidade de dados governamentais combinada com a utilização de IA abrem novos horizontes sobre otimização e precisão em relação à destinação de recursos públicos, bem como na identificação de deficiências e falhas na prestação de serviços. Ademais, trata-se de uma grande oportunidade para um controle mais eficaz da Administração Pública, bem como para o desenvolvimento de soluções inovadoras no interesse da sociedade, inclusive com maior participação de empresas, universidades e instituições não governamentais.

A Administração Pública deve assegurar que os seus sistemas de IA estejam em conformidade com a legislação de privacidade em vigor desde a concepção (*privacy by design*). Também deve observar as normas e requisitos de segurança da informação, de forma a assegurar disponibilidade, integridade, confidencialidade e autenticidade de informações especialmente daquelas consideradas sigilosas por razões de segurança da sociedade e do Estado, protegidas por sigilo legal ou que devem ser mantidas em segredo por razões empresariais (*trade secret*, segredo de negócio ou tudo que possa representar um diferencial competitivo), como no caso de empresas estatais exploradoras de atividades econômicas. Deve-se, outrossim, adotar mecanismos de proteção lógicos e físicos que assegurem a segurança e resiliência de sistemas contra-ataques e acesso de terceiros não autorizados.

A diretriz de Responsabilidade pelo adequado funcionamento enseja a necessidade de a Administração Pública, ao desenvolver e utilizar IA, verificar de forma apriorística (*ex ante*)

o atendimento aos princípios, diretrizes e normas definidos para o atendimento da moralidade administrativa, bem como, após a implantação de sistemas de IA (*ex post*), fazer revisões de conformidade de forma periódica. Ademais, a Administração deve assumir a responsabilidade pelas consequências decorrentes do uso de IA, inclusive por perdas e danos em razão de eventuais falhas, mal funcionamento e inconformidades legais e normativas.

REFERÊNCIAS

AGÊNCIA SENADO. Para especialistas, PL sobre Inteligência Artificial precisa de mais debates. Senado Notícias, Brasília, 9 dez. 2021. Disponível em: <https://www12.senado.leg.br/noticias/materias/2021/07/16/para-especialistas-pl-sobre-inteligencia-artificial-precisa-de-mais-debates>. Acesso em: 9 dez. 2021.

_____. Brasil poderá ter marco regulatório para a inteligência artificial. Senado Notícias, Brasília, 30 out. 2022. Disponível em: <https://www12.senado.leg.br/noticias/materias/2022/03/30/brasil-podera-ter-marco-regulatorio-para-a-inteligencia-artificial>. Acesso em: 30 out. 2022.

_____. Comissão conclui texto sobre regulação da inteligência artificial no Brasil. Senado Notícias, Brasília, 6 dez. 2022. Disponível em: <https://www12.senado.leg.br/noticias/materias/2022/12/06/comissao-conclui-texto-sobre-regulacao-da-inteligencia-artificial-no-brasil>. Acesso em: 10 jan. 2023.

ARISTÓTELES. Ética a Nicômaco. Tradução Mário de Gama Kury. 4. ed. Brasília: Editora Universidade de Brasília, 2001.

_____. Da Alma (De Anima). Tradutor Edson Bini. São Paulo: Edipro, 2011.

ARRUDA, Wellington. Engenheiro do Google que afirmou que IA tem vida é demitido. Tecmundo, [s. l.], 23 jul. 2022. Disponível em: <https://www.tecmundo.com.br/mercado/242193-engenheiro-google-afirmou-que-ia-tem-vida-demitido.htm>. Acesso em: 30 out. 2022.

BACON, Francis. Novum Organum ou Verdadeiras Indicações Acerca da Interpretação da Natureza. Tradução e notas de José Aluysio Reis de Andrade. Pará de Minas: M&M Editores Ltda, 2003.

_____. Nova Atlântida. 1626. Disponível em: http://www.hs-augsburg.de/~harsch/Chronologia/Lspost17/Bacon/bac_atla.html. Acesso em: 3 set. 2022.

BARBOSA, Rui. Oração aos Moços. Brasília: Senado Federal - Conselho Editorial, 2019. Disponível em: https://www2.senado.leg.br/bdsf/bitstream/handle/id/564558/Oracao_aos_mocos_Rui_Barbosa.pdf. Acesso em: 27 fev. 2023.

BELLMAN, R. E. An Introduction to Artificial Intelligence: Can Computers Think. New Jersey: Boyd & Fraser Publishing Company, 1978.

BENEDICTUS PP. XVI. Carta Encíclica SPE SALVI. Libreria Editrice Vaticana, Roma, novembro 2007. Disponível em: https://www.vatican.va/content/benedict-xvi/pt/encyclicals/documents/hf_ben-xvi_enc_20071130_spe-salvi.html. Acesso em: 3 set. 2022.

BENJAMIN, Walter. Capitalismo e religião. Tradução Nélcio Schneider. São Paulo: Boitempo, 2013.

BIONI, Bruno Ricardo; LUCIANO, Maria. O princípio da precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada. In: FRAZÃO, Ana (coord.); MULHOLLAND, Caitlin (coord.). Revista de Direito da Inteligência Artificial, Ética e Regulação. São Paulo: Revista dos Tribunais, 2019.

BONALDO, Frederico et al. A Renascença do Jusnaturalismo no Brasil e no Mundo. 1. ed. Londrina: Thoth Editora, 2021.

BONAVIDES, Paulo. Curso de direito constitucional. 10. ed. São Paulo: Malheiros, 2000.

BORELLA, Augusto. Inteligência Operacional auxilia empresas no cumprimento de tarefas e regulações. Inforchannel, [s. l.], novembro 2021. Seção Opinião. Disponível em: <https://inforchannel.com.br/2021/11/11/inteligencia-operacional-auxilia-empresas-no-cumprimento-de-tarefas-e-regulacoes/>. Acesso em: 25 nov. 2021.

BOSTROM, N. Superintelligence: Paths, Dangers, Strategies. Oxford: Oxford University Press, 2014.

BRANDÃO, Antônio José. Moralidade administrativa. Revista de Direito Administrativo, Rio de Janeiro, v. 25, n. 99, jul. 1951.

BRASIL. Constituição (1988). Constituição da República Federativa do Brasil de 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 7 dez. 2021.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). 15 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm. Acesso em: 9 dez. 2021.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. [S. l.], 18 nov. 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: 7 fev. 2023.

BRASIL. Lei nº 8.429, de 2 de junho de 1992. Dispõe sobre as sanções aplicáveis em virtude da prática de atos de improbidade administrativa, de que trata o § 4º do art. 37 da Constituição Federal; e dá outras providências. [S. l.], 3 jun. 1992. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/L8429compilada.htm. Acesso em: 07 out. 2022.

BRASIL. Decreto-lei nº 2.848, de 7 de dezembro de 1940. Institui o Código Penal: Poder Executivo, 1940. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em: 7 dez. 2021.

BRASIL. Decreto nº 9854, de 25 de junho de 2019. Institui o Plano Nacional de Internet das Coisas e dispõe sobre a Câmara de Gestão e Acompanhamento do Desenvolvimento de

Sistemas de Comunicação Máquina a Máquina e Internet das Coisas. Diário Oficial da União, Brasília, DF, 26 jun. 2019. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/d9854.htm. Acesso em: 4 mar. 2023.

BRASIL. Decreto nº 9.637, de 26 de dezembro de 2018. Institui a Política Nacional de Segurança da Informação, dispõe sobre a governança da segurança da informação, e altera o Decreto nº 2.295, de 4 de agosto de 1997, que regulamenta o disposto no art. 24, caput, inciso IX, da Lei nº 8.666, de 21 de junho de 1993, e dispõe sobre a dispensa de licitação nos casos que possam comprometer a segurança nacional. [S. 1.], 27 dez. 2018. Art. 12, IX. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/decreto/D9637.htm. Acesso em: 28 fev. 2023.

BRASIL. Ministério da Ciência, Tecnologia e Inovações/Gabinete do Ministro. Portaria GM nº 4.617, de 6 de abril de 2021. Institui a Estratégia Brasileira de Inteligência Artificial e seus eixos temáticos. [S. 1.], 12 abr. 2021. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-gm-n-4.617-de-6-de-abril-de-2021-*-313212172. Acesso em: 22 nov. 2021.

BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº 6387. Ementa. Medida Cautelar em Ação Direta de Inconstitucionalidade. Referendo. Medida Provisória Nº 954/2020. Emergência de Saúde Pública de Importância Internacional Decorrente do Novo Coronavírus (covid-19). Compartilhamento de Dados dos usuários do serviço telefônico fixo comutado e do serviço móvel pessoal, pelas empresas prestadoras, com o Instituto Brasileiro de Geografia e Estatística. Fumus Boni Juris. Periculum in Mora. Deferimento. Recorrente: Conselho Federal da Ordem dos Advogados do Brasil - CFOAB. Intimado: Presidente da República. Relator: min. Rosa Weber. Julgamento em 6 e 7.5.2020. DJE nº 137, divulgado em 02/06/2020. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=754>.

BRASIL. Supremo Tribunal Federal. ADI 6649 e ADPF 695, rel. min. Gilmar Mendes, Ata nº 26, de 14/09/2022. DJE nº 191, divulgado em 23/09/2022.

BRASIL. Supremo Tribunal Federal. ADI 2.661 MC, rel. min. Celso de Mello, julgamento em 5-6-2002, Plenário, DJ de 23-8-2002.

BRASIL. Supremo Tribunal Federal. MS 32.494 MC, rel. min. Celso de Mello, j. 11-11-2013, dec. monocrática, DJE de 13-11-2013.

BRASIL. Supremo Tribunal Federal. RE 221.291, rel. min. Moreira Alves, julgado em 11 de abril de 2000, 1ª Turma, DJ de 9 de junho de 2000.

BRASIL. Supremo Tribunal Federal. RE 206.889, rel. min. Carlos Velloso, julgado em 25 de março de 1997, 2ª Turma, DJ de 13 de junho de 1997.

BRASIL. Supremo Tribunal Federal. AO 1.833, rel. min. Alexandre de Moraes, julgado em 10-4-2018, 1ª T, DJE de 8-5-2018.

BRASIL. Supremo Tribunal Federal. Recurso Extraordinário n. 579.951, Relator o Ministro Ricardo Lewandowski, Plenário, DJE 23.10.2008, principal paradigma da Súmula Vinculante n. 13.

BRASIL. Supremo Tribunal Federal. RE 570.392, rel. min. Cármen Lúcia, j. 11-12-2014, Plenário, DJE de 19-2-2015, Tema 29.

BRASIL. Tribunal de Contas da União. Acórdão 2593/2017 - Plenário, 22 nov 2021. Disponível em: <https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/AC%25C3%2593RD%25C3%2583O%25202593%252F2017/%2520/DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0/%2520>. Acesso em: 25 jun. 2022.

BRASIL. Tribunal de Contas da União. Acórdão nº 1139/2022 – Plenário, de 25 de maio de 2022. Disponível em: https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/*/NUMACORDAO%253A1139%2520ANOACORDAO%253A2022/DTRELEVANCIA%2520desc%252C%2520NUMACORDAOINT%2520desc/0/%2520. Acesso em: 25 jun. 2022.

BRASIL. Tribunal de Contas da União. Regimento interno do Tribunal de Contas da União. Brasília, 20 set. 2022. Disponível em: <https://portal.tcu.gov.br/data/files/AA/97/D7/1C/631D28102DFE0FF7F18818A8/RITCU.pdf>. Acesso em: 20 set. 2022.

BRUNO, Fernanda. Economia psíquica dos algoritmos. *Nexo Jornal*, São Paulo, 24 jul. 2018. Disponível em: <https://www.nexojornal.com.br/ensaio/2018/A-economia-ps%C3%ADquica-dos-algoritmos-quando-o-laborat%C3%B3rio-%C3%A9-o-mundo>. Acesso em: 10 nov. 2020.

CÂMARA DO DEPUTADOS. PL 21/2020, 29 set. 2021. Disponível em: <https://www.camara.leg.br/propostas-legislativas/2236340>. Acesso em: 23 mar. 2023.

CANOTILHO, José Joaquim Gomes. Direito Constitucional e Teoria da Constituição. 7. ed. Coimbra: Almedina, 2003.

CARVALHO FILHO, José dos Santos. A discricionariedade: análise do seu delineamento jurídico. In: GARCIA, Emerson. Discricionariedade Administrativa. Rio de Janeiro: Lumen Juris, 2005.

CICERO, Marcus Tullius. De Officiis. Transl. Walter Miller. Cambridge, MA: Harvard University Press, 1913.

COLBROOK, Matthew J.; ANTUN, Vegard; HANSEN, Anders C. The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem. Proceedings of the National Academy of Sciences, v. 119 (12), out. 2022. Disponível em: <https://www.pnas.org/doi/10.1073/pnas.2107151119>. Acesso em: 30 out. 2022.

COMISSÃO DE JURISTAS RESPONSÁVEL POR SUBSIDIAR ELABORAÇÃO DE SUBSTITUTIVO SOBRE INTELIGÊNCIA ARTIFICIAL NO BRASIL. Relatório Final: Comissão de Juristas instituída pelo Ato do Presidente do Senado nº 4, de 2022, destinada a subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051, de 2019, 21, de 2020, e 872, de 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil. Brasília, 2022. Disponível em: <https://legis.senado.leg.br/sdleg->

getter/documento/download/777129a2-e659-4053-bf2e-e4b53edc3a04. Acesso em: 8 mar. 2023. nota 43

COOPER, S. Barry; VAN LEEUWEN, Jan (Ed.). Alan Turing: His work and impact. Elsevier, 2013.

COPELAND, Jack. Alan Turing: The codebreaker who saved ‘millions of lives’. BBC News, Christchurch, Nova Zelândia, 19 jun. 2012. Disponível em: <https://www.bbc.com/news/technology-18419691>. Acesso em: 30 out. 2022.

CORNELIS, Disco; BAREND, Van der Meulen. Getting New Technologies Together: Studies in Making Sociotechnical Order. New York: Walter de Gruyter, 1998. Disponível em: https://books.google.com.br/books?id=1khsIZ-jbgEC&pg=PA206&lpg=PA206&ots=D38v82mSkm&output=html&sig=ACfU3U2jPixZgKq-PYwVPHDpwO2Zt31puQ&redir_esc=y. Acesso em: 6 dez. 2021.

CRAWFORD, Angus. Instagram “helped kill my daughter.” BBC News, Inglaterra, 22 jan. 2019. Disponível em: <https://www.bbc.com/news/av/uk-46966009>. Acesso em: 7 nov. 2019.

DELEUZE, Gilles. Conversações. Trad. Peter Pál Pelbart. 3. ed. Rio de Janeiro: Editora 34, 2017.

DELOITTE UNIVERSITY PRESS. How much time and money can AI save government? Cognitive technologies could free up hundreds of millions of public sector worker hours. [S. l.]: s. n.], 2017. Disponível em: https://www2.deloitte.com/content/dam/insights/us/articles/3834_How-much-time-and-money-can-AI-save-government/DUP_How-much-time-and-money-can-AI-save-government.pdf. Acesso em: 30 out. 2022.

DONEDA, Danilo Cesar Maganhoto, SOUZA, Carlos Affonso Pereira de et al. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. Pensar, Fortaleza, v. 23, n. 4, p. 1-17, out-dez. 2018.

ENCICLOPÉDIA do Holocausto: A Solução Final. 19 maio 2022. Disponível em: <https://encyclopedia.ushmm.org/content/pt-br/article/the-final-solution#:~:text=A%20E2%80%9CSolu%C3%A7%C3%A3o%20Final%20E2%80%9D%20foi%20o,em%20massa%20dos%20judeus%20europeus>. Acesso em: 3 set. 2022.

EUROPEAN DIGITAL RIGHTS. An EU Artificial Intelligence Act for Fundamental Rights: A Civil Society Statement. Internet, 30 nov. 2021. Disponível em: <https://edri.org/wp-content/uploads/2021/11/Political-statement-on-AI-Act.pdf>. Acesso em: 5 dez. 2021.

FERRAZ, Anna Cândida da Cunha. Processos Informais de Mudança da Constituição: Mutações Constitucionais e Mutações Inconstitucionais. São Paulo: Max Limonad, 1986.

FIA. Indústria 4.0: o que é, consequências, impactos positivos e negativos [Guia Completo]. Disponível em: <https://fia.com.br/blog/industria-4-0/>. Acesso em: 28 dez. 2020.

FRAGALE FILHO, Roberto. Quando a empiria é necessária? Anais do XIV Congresso Nacional do Conpedi, Fortaleza: Conpedi, 2005. Disponível em:

<http://www.publicadireito.com.br/conpedi/manaus/arquivos/anais/XIVCongresso/135.pdf>. Acesso em: 6 out. 2021.

FROMM, Erich. The sane society. New York: Rinehart & Company, 1955.

GATES, Dominic. Flawed analysis, failed oversight: How Boeing and FAA certified the suspect 737 MAX flight control system. The Seattle Times. Disponível em: <https://www.seattletimes.com/business/boeing-aerospace/failed-certification-faa-missed-safety-issues-in-the-737-max-system-implicated-in-the-lion-air-crash/>. Acesso em: 9 jan. 2022.

GOVERNMENT ACCOUNTABILITY OFFICE - GAO. Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities. [S. 1.], 2021. Disponível em: <https://www.gao.gov/assets/gao-21-519sp.pdf>. Acesso em: 9 mar. 2023.

GROSSMANN, Luís Osvaldo. Comissão de juristas fará novo texto da Lei sobre Inteligência Artificial. Convergência Digital, [S. 1.], 18 fev. 2022. Disponível em: <https://www.convergenciadigital.com.br/Governo/Legislacao/Comissao-de-juristas-fara-novo-texto-da-Lei-sobre-Inteligencia-Artificial-59495.html>. Acesso em: 9 jan. 2023.

G 20. G20 Ministerial Statement on Trade and Digital Economy., 2019. Disponível em: <https://www.mofa.go.jp/files/000486596.pdf>. Acesso em: 27 dez. 2020.

HÄBERLE, Peter. Hermenêutica Constitucional: a sociedade aberta dos intérpretes da Constituição: contribuição para a interpretação pluralista e procedimental da Constituição. Porto Alegre: SAFE, 1997.

HAN, Byung-Chul. Psicopolítica: O neoliberalismo e as novas técnicas de poder. Tradução de Maurício Liesen. Belo Horizonte: Ayiné, 2018.

HAURIOU, Maurice. Précis Elémentaire de Droit Administraof. 10. ed. Paris: [S. n.].

HERVADA, Javier. Lições Propedêuticas de Filosofia do Direito. 1. ed. São Paulo: Martins Fontes. 2008.

HESSE, Konrad. A Força Normativa da Constituição. Gilmar Ferreira Mendes. Porto Alegre: SAFE, 1991.

HIGGINS, Chris. A Brief History of Deep Blue, IBM's Chess Computer. Mental Floss, 29 jul. 2017. Disponível em: <https://www.mentalfloss.com/article/503178/brief-history-deep-blue-ibms-chess-computer>. Acesso em: 30 out. 2022.

HIGH-LEVEL EXPERT GROUP ON AI. Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment. [S. 1.], 2020. Disponível em: <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>. Acesso em: 9 mar. 2023.

HOFFMANN-RIEM, Wolfgang. Inteligência Artificial Como Oportunidade para a Regulação Jurídica. Direito Público, Porto Alegre, v. 16, n. 90, p. 11-38, nov-dez, 2019.

HUXLEY, Aldous. Admirável Mundo Novo. Lino Vallandro, Vidal Serrano. 22. ed. São Paulo: Biblioteca Azul, 2014.

_____. Brave New World Revisited. 1. ed. New York: Harper Perennial modern classics, 2006.

IBM. Global AI Adoption Index 2022: New research commissioned by IBM in partnership with Morning Consult. 2022. Disponível em: <https://www.ibm.com/watson/resources/ai-adoption>. Acesso em: 30 out. 2022.

INFORMATION COMMISSIONER'S OFFICE (ICO). Guidance on the AI auditing framework: Draft guidance for consultation. [S. l.], 2020. Disponível em: <https://ico.org.uk/media/2617219/guidance-on-the-ai-auditing-framework-draft-for-consultation.pdf>. Acesso em: 9 mar. 2023.

INOVAÇÃO TECNOLÓGICA. Paradoxo matemático expõe limites da Inteligência Artificial. [S.l.], 28 mar. 2022. Disponível em: <https://www.inovacaotecnologica.com.br/noticias/noticia.php?artigo=paradoxo-matematico-expoe-limites-inteligencia-artificial&id=010150220328#.Y18AMXbMKU1>. Acesso em: 30 out. 2022.

KANTAYYA, Shalini. CODED Bias. Netflix, 2020. Disponível em: <https://www.netflix.com/title/81328723>. Acesso em: 03 mar. 2023.

KASPAROV, Garry. Palestra proferida no TED Taks, Vancouver Bc, abr. 2017. Disponível em: https://www.youtube.com/watch?v=NP8xt8o4_5Q. Acesso em: 30 out. 2012.

KELSEN, Hans. Teoria pura do direito. Tradutor João Baptista Machado. 6. ed. São Paulo: Martins Fontes, 1998.

KUBRICK, Stanley. 2001: Uma Odisséia no Espaço. Produção da Metro-Goldwyn-Mayer (MGM), 1968. 1 filme (148 min), son., color.

LACHARRIÈRE, René Ladreit de. Le Controle hierarchique de l'administration dans la forme jurisdictionnelle. [S. l.]: Sirey, 1937.

LAMB, Luís C. O Futuro do Trabalho Pós-Pandemia de COVID-19: Reflexões sobre os Impactos da Inteligência Artificial, Ciência e Educação. Métricas, São Paulo, 2020. Disponível em: <https://metricas.usp.br/wp-content/uploads/2020/07/M%C3%A9tricasUSP-revisado14Jul2020.pdf>. Acesso em: 27 dez. 2020.

MAGRANI, Eduardo. Entre dados e robôs. 2. ed. Porto Alegre: Arquipélago Editorial, 2019.

MARÇAL, Euclides Evangelista. Areté e Paideia: a excelência e a educação na era de Homero. Pensamento Extemporâneo, [S.l.], 20 jan. 2013. Disponível em: <https://pensamentoextemporaneo.com.br/?p=2498>. Acesso em: 14 out. 2021.

MARTINS JÚNIOR, Wallace Paiva. Probidade Administrativa. 3. ed. São Paulo: Saraiva, 2006.

MATTIUZZO, Marcela. Distribuição de UTIs a pacientes de Covid-19 por algoritmo não elimina dilema ético: para advogada, direitos serão violados em qualquer hipótese; resta adotar critérios técnicos e transparentes. Folha de São Paulo, [S.l.], 13 abr. 2020. Disponível em: <https://www1.folha.uol.com.br/ilustrissima/2020/04/distribuicao-de-utis-a-pacientes-de-covid-19-por-algoritmo-nao-elimina-dilema-etico.shtml>. Acesso em: 25 out. 2021.

MAYBIN, Simon. Sistema de algoritmo que determina pena de condenados cria polêmica nos EUA. BBC News Brasil, 31 out. 2016. Disponível em: <https://www.bbc.com/portuguese/brasil-37677421>. Acesso em: 25 out. 2021.

MCCARTHY, John, MINSKY, Marvin L., ROCHESTER, Nathaniel, SHANNON, Claude E. A proposal for the Dartmouth summer research project on artificial intelligence. [S.l.]: AI magazine, v. 27, n. 4, p. 12-14, 2006.

MEIRELLES, Hely Lopes. Direito Administrativo Brasileiro. 37. ed. São Paulo: Malheiros, 2012.

MELLO, Celso Antônio Bandeira de. O conteúdo jurídico do princípio da igualdade. 3. ed. [S.l.]: Malheiros editores, 2007. Disponível em: https://edisciplinas.usp.br/pluginfile.php/5367569/mod_resource/content/2/MELLO_Princ%C3%ADpio%20da%20Igualdade.pdf. Acesso em: 27 fev. 2023.

MELO, Vinicius Holanda; PEREIRA JÚNIOR, Antonio Jorge. Os limites da inteligência artificial no exercício da prudência: as atividades jurídicas correm risco? Revista dos Tribunais, v. 1015, p. 107 - 127, 2020. Disponível em: <https://www.thomsonreuters.com.br/content/dam/openweb/documents/pdf/Brazil/revistas-especializadas/rt-1015-vinicius-holanda-melo-e-antonio-jorge-pereira-junior-os-limites-da-inteligencia-artificial.pdf>. Acesso em: 4 jan. 2022.

MOLNAR, Christoph. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. Victoria: Leanpub, 2022

MORIN, Edgar; KERN, Anne-Brigitte. Terra-Pátria. Porto Alegre: Sulina, 2003.

NTOUSI, Eirini et al. Bias in data-driven artificial intelligence systems – an introductory survey. WIRES Data Mining and Knowledge Discovery, v. 10, n. 3, e1356, p. 1-14, 2020. Disponível em: <https://wires.onlinelibrary.wiley.com/doi/epdf/10.1002/widm.1356>. Acesso em: 06 fev. 2023.

NÚCLEO DE INFORMAÇÃO E COORDENAÇÃO DO PONTO BR. Pesquisa sobre o uso das tecnologias de informação e comunicação no setor público brasileiro: TIC Governo Eletrônico 2022. São Paulo, SP: Comitê Gestor da Internet no Brasil, 2022. Disponível em: https://cetic.br/media/docs/publicacoes/2/20220725170710/tic_governo_eletronico_2021_livro_eletronico.pdf. Acesso em: 17 set. 2022.

OCDE. OECD Principles on AI. Paris, 2019. Disponível em: <https://www.oecd.org/going-digital/ai/principles/>. Acesso em: 9 dez. 2021.

_____. Artificial Intelligence in Society, OECD Publishing, Paris, 2019. Disponível em: <https://doi.org/10.1787/eedfee77-en>. Acesso em: 6 mar. 2023.

_____. A caminho da era digital no Brasil. Paris: OCDE Publishing, 2020. Disponível em: <http://www.oecd.org/publications/a-caminho-da-era-digital-no-brasil-45a84b29-pt.htm>. Acesso em: 27 dez. 2020.

ORLOWSKI, Jeff. O Dilema das Redes. Netflix, 2020. Disponível em: <https://www.netflix.com/title/81254224>. Acesso em: 7 nov. 2020.

ORWELL, George. 1984. São Paulo: Companhia das Letras, 2009.

PARISER, Eli. The Filter Bubble: What the Internet Is Hiding from You. Nova York: Penguin Press, 2011.

PEREIRA, João Sergio dos Santos Soares. As decisões administrativas robóticas: das possibilidades aos limites. In: SADDY, André et al. Inteligência Artificial e Direito Administrativo. Rio de Janeiro: Centro para Estudos Empírico-Jurídicos (CEEJ), 2022.

PLATÃO. Protágoras. Carlos Alberto Nunes. Pará: Universidade Federal do Pará, 1980.

RACCA, Gabriela M. PERIN, Roberto Cavallo. Corrupção como quebra da confiança na administração pública e violação aos direitos fundamentais. RCJ – Revista Culturas Jurídicas. Niterói, v. 2, 3 ed., 2015.

RATTI, Carlo, CLAUDEL, Matthew. The City of Tomorrow: Sensors, Networks, Hackers, and the Future of Urban Life. [S. l.]: Yale University Press, 2016.

REALE, Giovanni, ANTISERI, Dario. História da Filosofia: Antiguidade e Idade Média. 3. ed. São Paulo: Paulus, 1990.

_____. O saber dos antigos: Terapia para os tempos atuais. 3. ed. São Paulo: Loyola, 2011.

REALE, Miguel. Lições preliminares de direito. 27. ed. São Paulo: Saraiva, 2002.

RICH, Elaine; KNIGHT, Kevin. Artificial Intelligence Second Edition. Nova Iorque: McGraw-Hill, 1991.

RODRIGUES, Marcelo. Microsoft explica episódio com chatbot racista e diz que Tay deve voltar. Tecmundo, Página web, 26 mar. 2016. Disponível em: <https://www.tecmundo.com.br/inteligencia-artificial/102835-microsoft-explica-episodio-chatbot-racista-diz-tay-deve-voltar.htm>. Acesso em: 6 fev. 2023.

RUSSELL, Stuart; NORVIG, Peter. Artificial Intelligence: A Modern Approach. 3. ed. New Jersey: Prentice Hall, 2010.

SADDY, André. Apreciatividade e discricionariedade administrativa. 2. ed. Rio de Janeiro: CEEJ, 2020.

_____. Curso de direito administrativo brasileiro. 2. ed. Rio de Janeiro: CEEJ, 2023, v.1

_____. Curso de direito administrativo brasileiro. 2. ed. Rio de Janeiro: CEEJ, 2023, v.2

_____.; et al. Aspectos Jurídicos das Novas Tecnologias (inovações) Disruptivas. 2. Ed. Rio de Janeiro: CEEJ, 2020.

_____.; GALIL, João Victor Tavares. O processo de tomada de decisão administrativa e o uso da inteligência artificial. In: SADDY, André et al. Inteligência Artificial e Direito Administrativo. Rio de Janeiro: Centro para Estudos Empírico-Jurídicos (CEEJ), 2022.

SEARLE, John R. Minds, brains, and programs. The behavioral and brain sciences, Estados Unidos, 1980. Disponível em: <https://home.csulb.edu/~cwallis/382/readings/482/searle.minds.brains.programs.bbs.1980.pdf>. Acesso em: 07 dez. 2021.

SEIFFERT, A. S. Os robôs de Asimov e o futuro da humanidade. Revista Eletrônica Da ANPHLAC, (24), 374–393. 2018. Disponível em: <https://revista.anphlac.org.br/anphlac/article/view/2877>. Acesso em: 7 dez. 2021.

SIMON, Herbert A. Designing Organizations for an Information-Rich World. In: GREENBERGER, Martin (Ed.). Computers, communications, and the public interest. Baltimore, MD: Johns Hopkins Press, 1971.

SOCIAL Credit: China's Digital Dystopia In: The Making. Direção: Janis Mackey Frayer. NBC News, 2018. Disponível em: <https://www.youtube.com/watch?v=0cGB8dCDf3c>. Acesso em: 30 out. 2022.

SOUSA, Antônio Francisco de. Os "conceitos legais indeterminados" no direito administrativo alemão. Revista de Direito Administrativo. Rio de Janeiro, v. 166, p. 276-291, 5 fev. 1986.

TRANSPARÊNCIA BRASIL. Estrutura de Avaliação de Riscos, 2020. Disponível em: https://www.transparencia.org.br/downloads/publicacoes/Estrutura_Avaliacao_Risco.pdf. Acesso em: 9 mar. 2023.

TURING, Alan. Computing Machinery and Intelligence. Mind, v. LIX, n. 236, p. 433-460, out. 1950. Disponível em: <https://archive.nytimes.com/www.nytimes.com/library/cyber/surf/1106surf-turing.html>. Acesso em: 30 out. 2022.

UNIÃO EUROPEIA. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho. Relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679&qid=1677953718324&from=PT>. Acesso em: 4 mar. 2022.

_____. Inteligência artificial para a Europa, COM/2018/237. Bruxelas, 25 abr. 2018. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52018DC0237>. Acesso em: 12 fev. 2023.

_____. Aumentar a confiança numa inteligência artificial centrada no ser humano, COM(2019) 168 final. Bruxelas, 8 abr. 2019. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52019DC0168&from=EN>. Acesso em: 12 fev. 2023.

VALLE, Vanice Regina Lírio do. Administração Pública, viés algorítmico e desvio de finalidade: existe conciliação possível? In: ZOCKUN, Maurício, GABARDO, Emerson. (coord.). Direito Administrativo e Inovação: crises e soluções. Curitiba: Íthala, 2022.

WAKKA, Wagner. Brasileiro integra Comitê de Supervisão de Liberdade no Facebook. Canaltech.com.br, [S. l.], 7 maio 2020. Disponível em: <https://canaltech.com.br/redes-sociais/brasileiro-integra-comite-de-supervisao-de-liberdade-no-facebook-164485/>. Acesso em: 20 out. 2021.

WINSTON, P. H. Artificial Intelligence. Massachusetts: Addison-Wesley, 1992.

ZUBOFF, Shoshana. A Era do Capitalismo de Vigilância. 1. ed. Rio de Janeiro: Intrínseca, 2021.