

# A REGULAÇÃO DA SUPERINTELIGÊNCIA SOB AS ÓTICAS DO PENSAMENTO TRANSHUMANISTA E DOS DIREITOS FUNDAMENTAIS

## THE REGULATION OF SUPERINTELLIGENCE UNDER THE PERSPECTIVE OF TRANSHUMANIST THOUGHT AND FUNDAMENTAL RIGHTS

*Lucas Pacheco Vieira*<sup>1</sup>

**RESUMO:** O presente artigo versa sobre os caminhos para a regulação dos fenômenos capazes de conduzir ao desenvolvimento da superinteligência, entendida como qualquer intelecto que exceda em muito o desempenho cognitivo dos seres humanos em, virtualmente, todos os domínios de interesse. Os campos do conhecimento que provavelmente permitirão tal desenvolvimento, segundo os estudiosos, são a inteligência artificial e as neurotecnologias. Em face disso, aborda-se, de um lado, uma visão regulatória derivada do pensamento transhumanista (autorregulação), e, de outro lado, formula-se uma visão regulatória alicerçada nos direitos fundamentais, mediante a qual é sustentada a regulação através de novo marco legislativo infraconstitucional e alterações em leis já existentes no ordenamento jurídico brasileiro.

**PALAVRAS-CHAVE:** Superinteligência; transumanismo; inteligência artificial; neurotecnologias; direitos fundamentais.

**ABSTRACT:** This article deals with the ways to regulate phenomena capable of leading to the development of superintelligence, understood as any intellect that greatly exceeds the cognitive performance of human beings in virtually all domains of interest. The fields of knowledge that will probably allow such development, according to scholars, are artificial intelligence and neurotechnologies. In view of this, on the one hand, a regulatory vision derived from transhumanist thinking (self-regulation) is approached, and, on the other hand, a regulatory vision based on fundamental rights is formulated, through which regulation is sustained through a new infraconstitutional legislative framework and amendments to existing laws in the Brazilian legal system.

**KEY-WORDS:** Superintelligence; transhumanism; artificial intelligence; neurotechnologies; fundamental rights.

---

<sup>1</sup> Doutorando em Ciência Jurídica pela Università degli Studi di Perugia (UNIPG) e pela UNIVALI. Mestre em Direito da Empresa e dos Negócios pela UNISINOS. Especialista em Direito Tributário pela PUC-RS. Professor de Direito Empresarial, Direito Tributário e *Compliance* e LGPD da AMF (Antonio Meneghetti Faculdade).

## 1. INTRODUÇÃO

O rápido desenvolvimento da área de inteligência artificial, percebido pelo público em geral através de ferramentas como ChatGPT e Bard, vem suscitando questionamentos técnicos, éticos, econômicos, políticos, sociais e jurídicos.

Uma das temáticas controversas é a da superinteligência, entendida como qualquer intelecto que exceda em muito o desempenho cognitivo dos seres humanos em, virtualmente, todos os campos do conhecimento. A literatura especializada aponta que as áreas que provavelmente permitirão tal desenvolvimento são a inteligência artificial e as neurotecnologias.

A corrente de pensamento que demonstra maior entusiasmo e tem como filiados alguns dos maiores financiadores destes desdobramentos tecnológicos é a do transhumanismo. Cuida-se de uma nova ideologia nascida nos Estados Unidos, apoiada por grandes empresários (v. g., Elon Musk) e corporações transnacionais (v. g., Google) e dotada de centros de pesquisa com financiamentos elevadíssimos, cujo intuito está centrado em “melhorar” o ser humano, no sentido de trabalhar em seu *improvement* ou *enhancement*, mediante o uso convergente das novas tecnologias designadas sob o acrônimo NBIC: nanotecnologia, biotecnologia, informática (*big data*, internet das coisas) e cognitivismo (inteligência artificial e robótica).<sup>2</sup>

O transhumanistas consideram-se herdeiros do humanismo<sup>3</sup>, da filosofia iluminista<sup>4</sup>, pois privilegiam a razão, o progresso e o bem-estar, com a diferença de que “levam o humanismo mais longe, desafiando os limites humanos por meio da ciência e da tecnologia combinadas com o pensamento crítico e criativo”<sup>5</sup>. A melhoria do ser humano, na ótica do transhumanismo, é

---

<sup>2</sup> FERRY, Luc. **A Revolução Transumanista**. Barueri: Manole, 2018. p. XXVIII a XXX. ROCO, Mihail e BAINBRIDGE, William. **Converging Technologies for improving human performance – nanotechnology, biotechnology, information technology and cognitive Science**. Kluwer Academic Publishers, 2003.

<sup>3</sup> MOORE, Max. **The Extropian Principles. A Transhumanist Declaration**.

<sup>4</sup> JOTTERAND, F. “At the roots of transhumanism: from the Enlightenment to a posthuman future.” *In: Journal of Medicine and Philosophy*, v. 35, n. 6, p. 617. 2010. Murilo Vilaça e Maria Clara Marques Dias sustentam que a base do transhumanismo está calcada em um “modo de Iluminismo humanista de raízes biológicas”. VILAÇA, Murilo e DIAS, Maria. “Transumanismo e o futuro (pós-) humano.” *In: Physis Revista de Saúde Coletiva*, Rio de Janeiro, 24 [2]: 341-362, 2014. p. 341.

<sup>5</sup> MOORE, Max. **The Extropian Principles. A Transhumanist Declaration**.

concebida como a possibilidade de aumentar as capacidades intelectuais, físicas, emocionais e espirituais das pessoas, através do manejo da ciência e da tecnologia.<sup>6</sup>

Pretende-se, com isso, questionar a inevitabilidade do envelhecimento, das doenças e da morte, ultrapassando o paradigma da terapêutica (recuperação diante de uma patologia) em prol da melhoria contínua (*continual improvement*), de maneira a justificar o uso da ciência para acelerar a transição da condição humana para a transhumana ou pós-humana.<sup>7</sup> Em suma, o intento central é mudar “do acaso para a escolha” (*from chance to choice*), da loteria genética que se impõe sobre a humanidade para um melhoramento livremente consentido e ativamente procurado, principalmente via engenharia genética e interfaces cérebro-máquina, redesenhando as fronteiras da existência humana.<sup>8</sup>

Apesar do otimismo prático<sup>9</sup> com o progresso da ciência e da tecnologia, os pensadores do transumanismo, em especial Max Moore, Nick

---

<sup>6</sup> BOSTROM, Nick. **Human reproductive cloning from the perspective of the future**, dez/2002.

<sup>7</sup> MOORE, Max. *Op. Cit.* FERRY, Luc. *Op. Cit.* Pp. XXIX a XXX.

<sup>8</sup> FERRY, Luc. *Op. Cit.* p. XLVI. JOTTERAND, F. “At the roots of transhumanism: from the Enlightenment to a posthuman future.” *In: Journal of Medicine and Philosophy*, v. 35, n. 6, p. 620. 2010. PORTER, Allen. “Bioethics and Transhumanism”. *In: Journal of Medicine and Philosophy*, v. 42, n. 3, p. 237, 2017.

<sup>9</sup> MOORE, Max. *Op. Cit.* A expressão ‘otimismo prático’ veio, na Versão 3.0 dos ‘Princípios Extropianos – Uma Declaração Transhumanista’, de 1998, em substituição ao “otimismo dinâmico”, inicialmente descrito em *Extropy #8*. Os principais trechos da Declaração sobre o ‘otimismo prático’ assim o enunciam: “Extropianos adotam uma atitude positiva, dinâmica e fortalecedora. Procuramos realizar nossos ideais neste mundo, hoje e amanhã. Em vez de suportar uma vida insatisfatória sustentada por fantasias de outra vida (seja em devaneios ou em uma “vida após a morte”), direcionamos nossas energias com entusiasmo para nos movermos em direção à nossa visão em constante evolução. (...) Questionamos os limites que os outros consideram dados. Observando a aprendizagem científica e técnica acelerada, padrões de vida ascendentes e práticas sociais e morais em evolução, projetamos e encorajamos o progresso contínuo. Hoje, há mais pesquisadores estudando envelhecimento, medicina, computadores, biotecnologia, nanotecnologia e outras disciplinas habilitadoras do que em toda a história. O desenvolvimento tecnológico e social continua a acelerar. Extropianos se esforçam para manter o ritmo do progresso, incentivando o apoio para pesquisas cruciais e sendo pioneiros na implementação de seus resultados. Mantemos um ceticismo construtivo em relação às crenças limitantes de nossos associados, de nossa sociedade e de nós mesmos. Vemos os obstáculos atuais do passado ao reter uma abertura criativa fundamental para as possibilidades. (...) Otimismo prático e fé passiva são incompatíveis. Otimismo prático significa otimismo crítico. A fé em um futuro melhor é a confiança de que uma força externa, seja Deus, o Estado ou até mesmo extraterrestres, resolverá nossos problemas. A fé gera passividade ao prometer progresso como um dom concedido a nós por forças superiores. Mas, em troca do presente, a fé requer uma crença fixa e súplica às forças externas, criando assim crenças dogmáticas e comportamento irracional. O otimismo prático fomenta a iniciativa e a inteligência, garantindo-nos que somos capazes de melhorar a vida com o nosso próprio esforço. Oportunidades e possibilidades estão por toda parte, chamando-nos para aproveitá-las e construí-las. Alcançar

Bostrom, Fabrice Jotterand, Allen Buchanan, Julian Savulescu e James Hughes, reconhecem diversos problemas sociais - como a desigualdade no acesso às melhorias transumanistas, que pode acabar reproduzindo o cenário do convívio entre os homens de Neanderthal e os homens de Cro-Magnon, que terminou selando a extinção dos últimos -, danos individuais em experimentos genéticos e até mesmo desastres de grande magnitude a colocar em risco até a sobrevivência da espécie.<sup>10</sup>

Neste contexto, mesmo sabendo dos riscos e ameaças, o posicionamento transhumanista sobre regulação das novas tecnologias é praticamente unânime na linha da autorregulação pura.

Contudo, uma visão baseada nos direitos fundamentais positivados nas Cartas Constitucionais atuais pode levar a uma postura diferente. O balanceamento dos riscos/ameaças e dos avanços de novas tecnologias envolvendo IA e neurotecnologias aponta para construída uma principiologia e um catálogo de direitos capazes de proteger os cidadãos contra as consequências negativas que podem sobrevir como um sistema superinteligente em desconformidade com padrões civilizatórios elementares, ou até voltado para eliminar a humanidade.

Assim, no presente artigo serão examinados o fenômeno das superinteligências, especificamente quando oriundas da conjugação de esforços entre ferramentas de Inteligência Artificial e Neurotecnologia, e as possibilidades regulatórias, tanto as sugeridas pelo pensamento transhumanista quanto as decorrentes de uma visão lastreada nos direitos fundamentais.

## 2. CONCEITO E FORMAS DE SUPERINTELIGÊNCIA

A superinteligência é definida, segundo Nick Bostrom, “como *qualquer intelecto que exceda em muito o desempenho cognitivo dos seres humanos em, virtualmente, todos os domínios de interesse*”<sup>11</sup>. O termo faz referência a níveis sobre-humanos de inteligência *geral* - não sendo aplicável a inteligências

---

nossos objetivos exige que acreditemos em nós mesmos, trabalhemos diligentemente e que estejamos dispostos a revisar nossas estratégias.”

<sup>10</sup> VILAÇA, Murilo e DIAS, Maria. *Op. Cit.* p. 348-350.

<sup>11</sup> BOSTROM, Nick. **Superinteligência: caminhos, perigos, estratégias**. Darkside: Cajamar, 2018. p. 59.

superiores restritas a operações específicas, tais como programas que venceram campeões mundiais em jogos como gamão (BKG, em 1979), damas (CHINOOK, em 1994), xadrez (DEEP BLUE, 1997), Jeopardy! (WATSON, em 2010) e Go (AlphaGo, em 2017)<sup>12</sup>.

O conceito clássico de ‘máquina ultrainteligente’ e a sua decorrência de um evento denominado ‘explosão de inteligência’ foi elaborado originalmente por Irving John Good, matemático britânico que trabalhou com Alan Turing na *Government Code and Cypher School* (GC&CS), em Bletchley Park, no Reino Unido, durante a Segunda Guerra Mundial, e lecionou no Virginia Tech posteriormente. De acordo com Good, uma máquina ultrainteligente pode ser definida como “uma máquina que pode muito ultrapassar todas as atividades intelectuais de qualquer homem, por mais inteligente que seja.”<sup>13</sup> O cientista já vislumbrava a possibilidade de a máquina ultrainteligente projetar máquinas ainda melhores, causando uma ‘*explosão de inteligência*’, fenômeno que deixaria a inteligência humana para trás<sup>14</sup>. Em face disso, alertava:

“Assim, a primeira máquina ultrainteligente é a última invenção que o homem precisaria fazer, desde que a máquina seja dócil o suficiente para dizer como mantê-lo sob controle. É curioso que este ponto seja levantado tão raramente fora da ficção científica. Às vezes vale a pena levar a ficção científica a sério.”<sup>15</sup>

Stuart Russell, Professor de Ciência da Computação da Universidade da Califórnia em Berkeley, fundador do *Center for Human-Compatible Artificial Intelligence* (CHAI) e coautor do manual mais disseminado no campo da Inteligência Artificial<sup>16</sup>, defende concepção semelhante à de Nick Bostrom e I.J. Good, caracterizando a superinteligência, em matéria de Inteligência Artificial, como uma IA forte *geral*, ou de *propósito geral*.<sup>17</sup>

---

<sup>12</sup> Idem. p. 42-44. SINGH, S., OKUN, A. & JACKSON, A. “Learning to play Go from scratch.” *Nature* 550, 336–337 (2017). <https://doi.org/10.1038/550336a>. Acesso em: 01/04/2021.

<sup>13</sup> GOOD, Irving John. 1965. “Speculations Concerning the First Ultraintelligent Machine.” *In: Advances in Computers*, edited by Franz L. Alt and Morris Rubinoff, 31–88. Vol. 6. New York: Academic Press. Disponível em: <https://vtechworks.lib.vt.edu/bitstream/handle/10919/89424/TechReport05-3.pdf>. Acesso em: 02/04/2021. p. 33. Tradução livre.

<sup>14</sup> Idem. p. 33.

<sup>15</sup> Idem. p. 33.

<sup>16</sup> UNIVERSIDADE DA CALIFÓRNIA EM BERKELEY. **1525 Schools Worldwide That Have Adopted AIMA**. Disponível em: <http://aima.cs.berkeley.edu/adoptions.html>. Acesso em: 01/04/2021.

<sup>17</sup> RUSSELL, Stuart. **Human Compatible**. Nova York: Penguin Publishing Group, 2020. p. 76.

Esse sistema ainda não existe. Mas recentes pesquisas demonstram que a maioria dos pesquisadores ativos em IA projeta que a IA de nível humano chegue por volta da metade do século XXI. Russell aponta uma trajetória de oitenta anos até esse estágio, porém alerta para a possibilidade de a IA superinteligente chegar repentinamente. Como exemplo, recorda o que ocorreu com a ideia de Szilard para uma reação nuclear em cadeia induzida por nêutrons na década de 30, horas depois de Lord Ernest Rutherford ter consignado publicamente a impossibilidade desse avanço.<sup>18</sup>

Outra situação dessa natureza na comunidade científica, desta vez no ramo da edição genética, aconteceu com a revolucionária tecnologia CRISPR-Cas9, inventada por Jennifer Doudna e Emmanuelle Charpentier – que foram agraciadas com o Prêmio Nobel de Química em 2020 por esta contribuição. A técnica foi utilizada, poucos anos depois de surgir, na segunda década do século XXI, sem que tenham sido observadas as devidas medidas de segurança e precaução, ou mesmo maiores reflexões éticas, pelo cientista chinês He Jiankui, que levou a cabo o desenvolvimento dos primeiros bebês geneticamente editados.

O experimento, súbito e não imaginado pela comunidade científica internacional, pretendia livrar totalmente os bebês do vírus HIV, portado pelo pai, bloqueando inclusive a possibilidade de infecções futuras.<sup>19</sup> A justificativa apresentada por Jiankui para a sua realização estava ancorada na obra do transhumanista Allen Buchanan, então Professor de Filosofia da Universidade Duke<sup>20</sup>. O resultado, assim como a ideia de Szilard, que conduziu às armas atômicas, foi bastante negativo, tendo o cientista sido condenado criminalmente na China e o país iniciado discussões, num clima desfavorável à engenharia genética, para introduzir normas jurídicas estabelecendo a responsabilização de profissionais que manipulassem genes humanos em adultos ou embriões por eventuais consequências danosas.<sup>21</sup>

---

<sup>18</sup> Idem. p. 77-78.

<sup>19</sup> ISAACSON, Walter. **A decodificadora**: Jennifer Doudna, edição de genes e o futuro da espécie humana. Rio de Janeiro: Editora Intrínseca, 2021. p. 391.

<sup>20</sup> Idem. p. 396-397.

<sup>21</sup> CYRANOSKI, David. “China set to introduce gene-editing regulation following CRISPR-baby furor”. In: **Nature**, 20 de maio de 2019. Disponível em: <https://www.nature.com/articles/d41586-019-01580-1>. Acesso em: 01/04/2021.

Volvendo à superinteligência, cabe registrar que sérias preocupações com seu desenvolvimento foram objeto de emblemática carta aberta firmada, em janeiro de 2015, pelos empresários Bill Gates e Elon Musk, por cientistas e professores universitários renomados como Stephen Hawking, Stuart Russell, Vernor Vinge, Max Tegmark, Joshua Greene, Nick Bostrom, Daniel Dewey, Eric Horvitz e centenas de especialistas em IA.<sup>22</sup> No artigo científico anexo à carta, redigido por Stuart Russell, Max Tegmark e Daniel Dewey, a privacidade foi um dos seis tópicos mencionados entre as prioridades sugeridas para as pesquisas sobre direito e ética em matéria de inteligência artificial, ao lado da responsabilidade civil e legislação para veículos autônomos, ética de máquina, armas autônomas, ética profissional e políticas públicas.<sup>23</sup>

Uma iniciativa proeminente que considera necessários estudos aprofundados sobre a superinteligência é o “*The One Hundred Year Study on Artificial Intelligence*”, fundado pelo *Chief Scientific Officer* da Microsoft, Eric Horvitz, na Universidade Stanford. Em uma das primeiras publicações desse empreendimento acadêmico, que inclusive influenciou a carta aberta supracitada, Eric Horvitz colocou entre os temas centrais a perda de controle sobre sistemas de IA, com a ascensão de uma perigosa superinteligência, e os desafios impostos pela inteligência de máquina no campo da privacidade.<sup>24</sup>

Quanto às formas de superinteligência, Nick Bostrom sustenta a existência de três principais: rápida, coletiva e de qualidade<sup>25</sup>. A superinteligência rápida é um intelecto capaz de fazer tudo o que um intelecto humano realiza, porém muito mais rapidamente.<sup>26</sup> Como exemplo, pode-se mencionar uma emulação completa do cérebro executada em um *hardware*

---

<sup>22</sup> RUSSELL, S.; DIETTERICH, T.; HORVITZ, E.; SELMAN, B.; ROSSI, F.; HASSABIS, D.; LEGG, S.; SULEYMAN, M.; GEORGE, D.; PHOENIX, S. “Letter to the Editor: Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter.” *In: AI Magazine*, [S. l.], v. 36, n. 4, 2015. Disponível em: <https://ojs.aaai.org/index.php/aimagazine/article/view/2621>. Acesso em: 3 apr. 2021. P. 3-4.

<sup>23</sup> RUSSELL, Stuart; TEGMARK, Max; DEWEY, Daniel. “Research priorities for robust and beneficial artificial intelligence.” *In: AI Magazine*, V. 36, Issue 4, Inverno de 2015, Janeiro de 2015. Disponível em: <https://ojs.aaai.org/index.php/aimagazine/article/view/2577>. Acesso em: 02/04/2021. p. 107.

<sup>24</sup> HORVITZ, Eric. **One-Hundred Year Study of Artificial Intelligence: Reflections and Framing**, White paper, Stanford University, Stanford, CA (ai100.stanford.edu). Disponível em: [https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100\\_framing\\_memo\\_0.pdf](https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100_framing_memo_0.pdf). Acesso em: 03/04/2021. p. 3-5.

<sup>25</sup> BOSTROM, Nick. *Op. Cit.* p. 109.

<sup>26</sup> Idem. p. 109.

veloz, que operaria com velocidade 10 mil vezes maior que a de um cérebro biológico, tendo aptidão para ler um livro em segundos ou escrever uma tese doutoral em poucas horas. Projeta-se uma preferência pelo convívio com outras mentes rápidas, ao invés de se comunicar com “as lerdas mentes humanas”<sup>27</sup>.

A superinteligência coletiva é “um sistema composto de um grande número de intelectos menores, de forma que o seu desempenho total supere significativamente, em diversas áreas gerais o conhecimento, qualquer sistema cognitivo atual”<sup>28</sup>. Empresas, grupos de trabalho e comunidades acadêmicas são exemplos de ‘sistemas’ capazes de resolver diferentes categorias de problemas, como operar negócios complexos, inventar novas tecnologias e otimizar processos, mediante inteligências coletivas.<sup>29</sup> No caso da superinteligência coletiva, depara-se com um conjunto de mentes humanas coordenadas ou unificadas (‘uma única e grande “mente”’) através do aprimoramento das tecnologias de comunicação para que muitos indivíduos possam trabalhar ao mesmo tempo em qualquer problema intelectual complexo.<sup>30</sup>

Por sua vez, a superinteligência de qualidade consubstancia “um sistema que é no mínimo tão rápido quanto uma mente humana e qualitativamente muito mais inteligente.”<sup>31</sup> A concepção de qualidade pode ser extraída, de um lado, por comparação com os demais animais. Bostrom afirma que se trata de “uma inteligência no mínimo superior à inteligência humana, assim como a qualidade da inteligência humana é superior à dos elefantes, golfinhos e chimpanzés.”<sup>32</sup> De outro lado, a noção de qualidade vincula-se à ideia de “*talentos cognitivos possíveis, mas não realizados*”, que outros sistemas inteligentes poderiam viabilizar mediante um novo conjunto de módulos capaz de potencializar as habilidades de realizar uma ampla gama de tarefas estrategicamente relevantes.<sup>33</sup>

Como se pode deduzir das modalidades descritas, a superinteligência pode ser hospedada em (i) máquinas com softwares potentes ou redes alicerçadas puramente em inteligência artificial; (ii) em híbridos de seres

---

<sup>27</sup> Idem. p. 110.

<sup>28</sup> Idem. p. 111.

<sup>29</sup> Idem. p. 111 e 113.

<sup>30</sup> Idem. p. 114.

<sup>31</sup> Idem. p. 114.

<sup>32</sup> Idem. p. 115.

<sup>33</sup> Idem. p. 116.

humanos e máquinas com base em implantes de dispositivos no cérebro ou de nanorrobôs no corpo humano; e/ou (iii) em seres humanos melhorados por técnicas biológicas/genéticas.

### 3. O CAMINHO PROVÁVEL PARA A SUPERINTELIGÊNCIA: INTELIGÊNCIA ARTIFICIAL E NEUROTECNOLOGIA

A superinteligência, de acordo com a literatura especializada, provavelmente será alcançada por meio da combinação de ferramentas de Inteligência Artificial e Neurotecnologia. Bostrom aponta a IA e a 'emulação completa do cérebro' como os caminhos mais prováveis, embora reconheça que a cognição biológica - via instrumentos genéticos -, as interfaces cérebro-computador e as redes e organizações também possam levar à superinteligência, mesmo que as chances sejam aparentemente menores.<sup>34</sup>

No que toca à IA, cumpre apresentar o conceito formulado pelo Grupo de Especialistas de Alto Nível em Inteligência Artificial, formado pela Comissão Europeia:

Os sistemas de inteligência artificial (IA) são sistemas de software (e possivelmente também hardware) projetados por humanos que, atendendo a um objetivo complexo, atuam na dimensão física ou digital percebendo seu ambiente por meio da aquisição de dados, interpretando os dados estruturados e não estruturados coletados, raciocinando sobre o conhecimento, ou processamento das informações, derivado desses dados e decidindo a(s) melhor(es) ação (ões) para atingir o objetivo determinado. Os sistemas de IA podem usar regras simbólicas ou aprender um modelo numérico, e também podem adaptar seu comportamento analisando como o ambiente é afetado por suas ações anteriores.”<sup>35</sup>

Diversos *experts* têm feito previsões sobre avanços extraordinários pelo menos desde os anos 1960. Marvin Minsky, um dos pioneiros, coorganizador do *workshop* de Dartmouth de 1956 que originou este campo do conhecimento, asseverou, em 1967, que, dentro de uma geração, o problema de criar inteligência artificial estaria substancialmente resolvido, restando poucos

---

<sup>34</sup> Idem. p. 104-105.

<sup>35</sup> COMISSÃO EUROPEIA. **A definition of Artificial Intelligence: main capabilities and scientific disciplines.** Março/2021. Disponível em: <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>. Acesso em: 08/05/2021. Tradução livre. p. 6.

compartimentos do intelecto fora do 'reino da máquina'<sup>36</sup>. O economista Herbert Simon, vencedor do Prêmio Nobel, havia projetado, em 1960, que, em 20 anos, as máquinas seriam capazes de fazer qualquer trabalho que um homem executa.<sup>37</sup>

Inobstante não se tenha alcançado a superinteligência geral, o fato é que as máquinas já excedem as capacidades humanas em várias áreas, conforme exemplificado anteriormente. Ao longo do tempo, esses desenvolvimentos serão ampliados e aprofundados, o que torna provável, segundo Stuart Russell, que surjam sistemas de conhecimento geral sobre-humanos, sistemas de pesquisa biomédica sobre-humanos, robôs super-humanos destros e ágeis, sistemas sobre-humanos de planejamento corporativo e assim por diante muito antes de termos um sistema de IA superinteligente completamente geral.<sup>38</sup>

A criação da IA superinteligente provavelmente sobrevirá, diz o citado Professor da UC Berkeley, de pesquisas sobre 'ferramentas de IA' ou 'IA estreita', o que significa sistemas de IA agradáveis, seguros e enfadonhos, projetados para problemas específicos, como jogar Go ou reconhecer dígitos manuscritos.<sup>39</sup> O time que inventou o AlphaGo, no interior da *Google DeepMind*, por exemplo, não arquitetou um código específico sobre o Go. Os protocolos decisórios não funcionam somente para o Go. As melhorias realizadas pertinem a técnicas de propósito geral (*general-purpose techniques*), aplicando-se a muitos outros problemas.<sup>40</sup>

Vale frisar que o desenvolvimento da IA superinteligente não pode ocultar os relevantes mudanças e avanços já promovidos pela IA na vida humana em geral. Conforme destacam Peter Diamandis e Steven Kotler, a resposta instantânea do mecanismo de pesquisa do Google e o reconhecimento de fala nas chamadas de consulta à lista telefônica evidenciam que já somos codependentes da IA.<sup>41</sup>

---

<sup>36</sup> MINSKY, Marvin APUD STUART, Russel. *Op. Cit.* p. 76.

<sup>37</sup> RUSSELL, Stuart. *Op. Cit.* p. 76.

<sup>38</sup> *Idem.* p. 76.

<sup>39</sup> *Idem.* p. 46.

<sup>40</sup> *Idem.* p. 46-47.

<sup>41</sup> DIAMANDIS, Peter e KOTLER, Steven. **Abundância: o futuro é melhor do que você imagina.** São Paulo: HSM Editora, 2012. p. 72.

Outro exemplo pode ser extraído do anúncio feito, em 2018, pelo Facebook de que havia conseguido que suas máquinas reconhecessem rostos “em estado natural” com 97,35% de precisão, chegando bem perto do desempenho em nível humano.<sup>42</sup> Embora a corporação tenha enfatizado as vantagens disso para a publicidade direcionada, Shoshana Zuboff afirma que “um crescimento ainda maior viria das imensas oportunidades de treinamento de máquina representadas por tantas fotografias”<sup>43</sup>. No mesmo ano, as máquinas do Facebook “estavam aprendendo a discernir atividades, interesses, estados de espírito, olhares, roupas, o caminhar, o cabelo, o tipo corporal e a postura”, segundo a referida Professora aposentada da *Harvard Business School*.<sup>44</sup>

Apesar desse cenário, Nick Bostrom consigna que persistem obstáculos técnicos relevantes para a IA superinteligente, notadamente no que atine aos “recursos computacionais necessários para que sejamos capazes de simplesmente replicar os processos evolutivos relevantes que produziram a inteligência de nível humano”<sup>45</sup>.

Um caminho provável para superar essas dificuldades é utilizar as ferramentas de IA em cotejo com as descobertas feitas a partir de instrumentos de neurotecnologia, que podem viabilizar uma emulação completa do cérebro e desvendar os princípios gerais de funcionamento do órgão, o que traria um modelo para a formulação da inteligência de máquina e guiar os esforços da IA para produzir uma superinteligência.<sup>46</sup>

Raymond Kurzweill, Diretor de Engenharia do Google, filiado à ideologia transhumanista, relatou, na primeira década do século XXI, que a maioria dos modelos baseados em pesquisa de escaneamento cerebral, utilizando métodos como fMRI (Ressonância Magnética Funcional), MEG e outros, ofereceram apenas sugestões dos mecanismos subjacentes ao funcionamento do cérebro. A resolução espacial e temporal rudimentar, contudo, não permitia a engenharia reversa das principais características do cérebro.<sup>47</sup>

---

<sup>42</sup> ZUBOFF, Shoshana. **A era do capitalismo de vigilância**: a luta por um futuro humano na nova fronteira de poder. Rio de Janeiro: Editora Intrínseca, 2021. p. 376.

<sup>43</sup> Idem. p. 376.

<sup>44</sup> Idem. p. 376.

<sup>45</sup> BOSTROM, Nick. *Op. Cit.* p. 63.

<sup>46</sup> KURZWEILL, Raymond. **A singularidade está próxima**: quando os humanos transcendem a biologia. São Paulo: Iluminuras, 2018. p. 240 e 699. BOSTROM, Nick. *Op. Cit.* p. 68.

<sup>47</sup> KURZWEILL, Raymond. *Op. Cit.* p. 263-264.

As recentes gerações da fMRI, ferramenta de varredura do cérebro mais utilizada – que ostenta resolução espacial relativamente alta de um a três milímetros - fornecem resoluções temporais de cerca de um segundo, ou um décimo de segundo para uma fatia fina do cérebro, gerando um aprimoramento em comparação com a geração anterior, que tinha resolução temporal de alguns segundos.<sup>48</sup> A principal vantagem da técnica é a possibilidade de obter imagens do sistema nervoso central durante a execução de uma atividade realizada pelo paciente que se queira analisar.<sup>49</sup> Captam-se imagens, portanto, de forma dinâmica.<sup>50</sup>

Outra técnica de varredura é a Tomografia por Emissão Positrônica (PET). Através desta ferramenta, átomos de determinados isótopos emissores de pósitrons são usados para identificar moléculas de interesse, sendo injetados na corrente sanguínea do organismo para rastrear tanto o metabolismo como o fluxo sanguíneo.<sup>51</sup> Esta técnica e a Tomografia por Emissão de Fóton Único (SPECT) permitem o desenvolvimento dos mapas tridimensionais da atividade do sistema nervoso central.<sup>52</sup>

Com o tempo, a PET e a SPECT passaram a ser usadas para obter imagens durante a execução de tarefas neuropsicológicas, abrangendo memória, atenção, linguagem e outras operações mentais, viabilizando a obtenção de dados sobre os padrões de funcionamento cerebral, associados a tarefas relevantes para a fisiopatologia de transtornos mentais específicos.<sup>53</sup>

Outra abordagem é a Estimulação Magnética Transcraniana (TMS), que consiste na aplicação de um campo magnético de pulso forte na parte de fora do crânio, colocando com precisão uma bobina magnética sobre a cabeça. Com a estimulação ou indução de uma 'lesão virtual' em pequenas regiões do cérebro, as habilidades podem ser diminuídas ou reforçadas.<sup>54</sup> O cientista Allan Snyder relatou que cerca de 40% dos sujeitos de teste ligados à TMS exibem novas habilidades significativas, tal como habilidade para desenhar.<sup>55</sup>

---

<sup>48</sup> Idem. p. 266.

<sup>49</sup> SILVA, Cláudio Messias. **Neurotecnologia: Fundamentos Teóricos, Tecnológicos e Mercadológicos**. Dissertação de Mestrado. São Leopoldo: Unisinos, 2016. p. 74.

<sup>50</sup> Idem. p. 74.

<sup>51</sup> Idem. p. 72.

<sup>52</sup> Idem. p. 73.

<sup>53</sup> Idem. p. 73.

<sup>54</sup> KURZWEILL, Raymond. *Op. Cit.* p. 268.

<sup>55</sup> Idem. p. 268.

Pesquisas na Universidade Carnegie Mellon, na Universidade A&M do Texas e na Universidade da Pensilvânia estão indo além. A primeira, conduzida por Andreas Nowatzky, estava realizando escaneamento do sistema nervoso central do cérebro e do corpo de um rato com resolução menor do que 200 nanômetros, o que se aproximava do necessário para a engenharia reversa total, segundo Kurzweill. A segunda, também de escaneamento, seria capaz de digitalizar um cérebro inteiro de um camundongo em um mês, com resolução de 250 nanômetros.<sup>56</sup> A terceira consiste numa câmera de escaneamento capaz de gerar imagens de neurônios individuais, suficientes para registrar o disparo de cada neurônio.<sup>57</sup>

Outras técnicas contemporâneas como a TPLSM (microscopia de varredura a laser de dois fótons) – utilizada para fazer cirurgias intracelulares ultraprecisas -, SHG (microscopia de segunda geração harmônica), OCI (geração de imagens com coerência óptica), e inclusive a possibilidade de uso de nanorrobôs para escanear o cérebro igualmente fazem parte dos esforços para decifrar digitalmente o cérebro.<sup>58</sup>

No estudo '*Whole Brain Emulation*', feito no âmbito do *Future of Humanity Institute*, da Universidade de Oxford, Anders Sandberg, membro do referido Instituto e PhD em Neurociência Computacional, e Nick Bostrom desenharam o seguinte gráfico para mostrar as capacidades tecnológicas necessárias para uma emulação completa do cérebro, incluindo, resalte-se, a formação de uma robusta base de dados para a interpretação do que foi escaneado<sup>59</sup>:

---

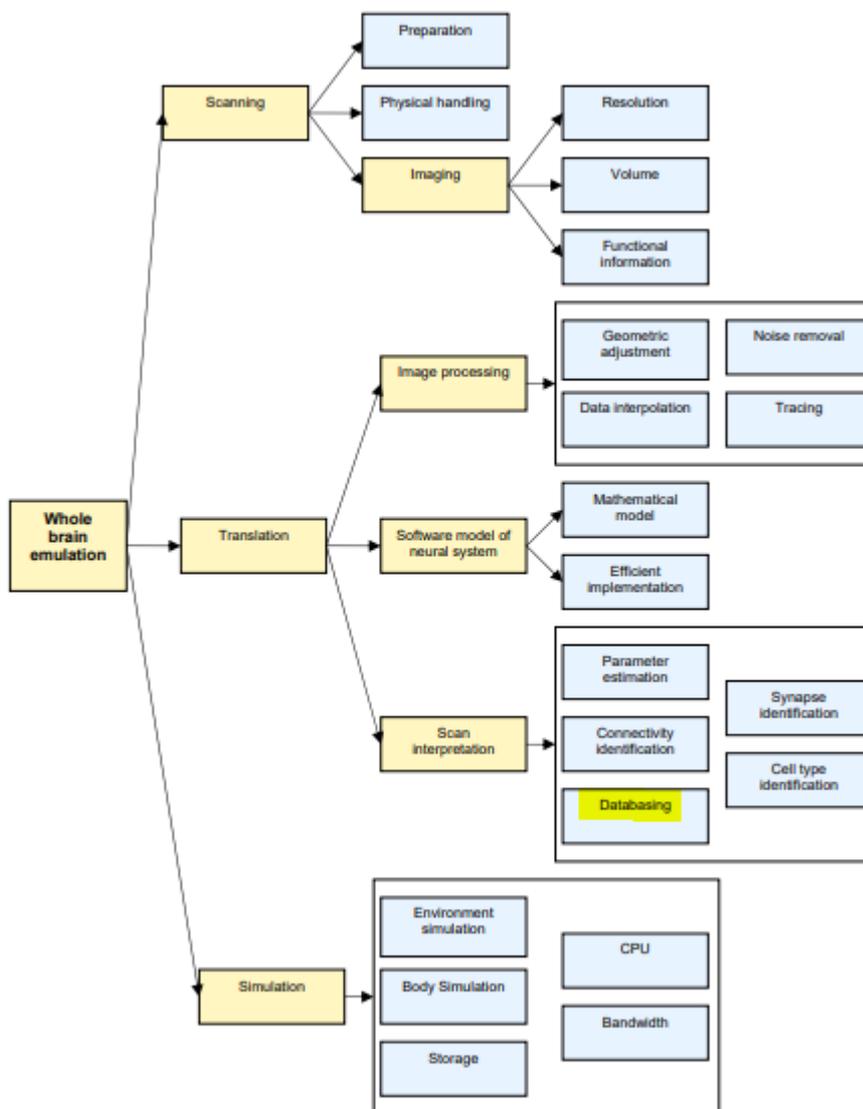
<sup>56</sup> Idem. p. 269.

<sup>57</sup> Idem. p. 269.

<sup>58</sup> Idem. p. 270-271.

<sup>59</sup> BOSTROM, Nick e SANDBERG, Anders. **Whole Brain Emulation: A Roadmap**. Technical Report #2008-3. Oxford: Future of Humanity Institute, Oxford University, 2008. p. 18.

Figura 1 - Capacidades tecnológicas necessárias para uma emulação completa do cérebro



Percebe-se, deste modo, que todas as ferramentas neurotecnológicas citadas vêm promovendo uma acumulação massiva de dados sobre a dinâmica e as características precisas dos sistemas e das partes constituintes do cérebro, desde sinapses individuais até amplas regiões como o cerebelo.<sup>60</sup> Kurzweill, na linha do que se depreende do estudo de Bostrom e Sandberg, relata que “bancos extensos de dados estão catalogando metodicamente nosso exponencial crescente conhecimento do cérebro.”<sup>61</sup>

<sup>60</sup> Idem. p. 240.

<sup>61</sup> Idem. p. 240.

Com isso, resta delineado o caminho provável para o surgimento de futuras superinteligências, mediante a conjugação de esforços de cientistas atuantes nos campos da IA e da neurotecnologia.

#### **4. A ABORDAGEM REGULATÓRIA DA SUPERINTELIGÊNCIA SOB A ÓTICA TRANSHUMANISTA**

Visualizados o conceito, as formas e os caminhos prováveis para as futuras superinteligências, torna-se imperativo examinar o panorama regulatório, haja vista que oferecerá os incentivos e os limites para o desenvolvimento desse fenômeno. A moldura jurídica constitui um dos elementos institucionais de maior relevo para o sucesso ou o fracasso de novas empreitadas econômicas.

A ótica regulatória exposta neste capítulo será a do transhumanismo, ancorada especialmente no posicionamento de Nick Bostrom e Raymond Kurzweill. Outros pensadores e cientistas, inclusive de outras vertentes, também serão referenciados para traçar de maneira crítica o quadro institucional alvitado pela corrente transhumanista.

A tecnologia, assinala Raymond Kurzweill, tem sido uma faca de dois gumes.<sup>62</sup> Oferece, de um lado, “benefícios como uma vida mais longa e mais saudável, liberdade do trabalho físico estafante e mental, e muitas novas possibilidades criativa”<sup>63</sup>. Por outro lado, introduz novos perigos, evidenciando ambas as faces da natureza humana, a criativa e a destrutiva.

O dilema relativo à escolha da regulação, da autorregulação regulada ou da autorregulação é frequente em assuntos que envolvem inovações tecnológicas. O setor de mídia, por exemplo, que era liderado por conglomerados detentores de emissoras de televisão, jornais e revistas, foi substancialmente transformado pela revolução encetada pelas novas tecnologias de comunicação e informação, ganhando inclusive a robusta concorrência das redes sociais, e até hoje prevalece, em vários países, um ambiente normatizado via autorregulação ou autorregulação regulada.

Os produtos e serviços que surgiram com as novas tecnologias noutros setores, marcados pelo signo da economia colaborativa, também não

---

<sup>62</sup> KURZWEILL, Raymond. *Op. Cit.* p. 680.

<sup>63</sup> *Idem.* p. 680.

vêm sendo regulados de maneira paternalista, em um momento inicial, em vários sistemas jurídicos. Os aplicativos de transporte de passageiros (Uber, Cabify, etc), os aplicativos de reserva de acomodações/hospedagem (*Airbnb*), os produtos nanofarmacêuticos, os dispositivos de nanomedicina, entre outros, não estão sujeitos, via de regra, a marcos legais específicos, que os conformem detalhada e exaustivamente, nos ordenamentos nacionais.

Embora se esteja numa fase de amadurecimento das discussões na seara jurídica, seja no que toca à promulgação de diplomas legislativos, seja quanto à atuação judicial para o enquadramento nas regras vigentes, a autorregulação e a autorregulação regulada têm sido as primeiras alternativas utilizadas, confiando-se aos agentes econômicos privados e instituições privadas como ISO a normatização dos novos segmentos.<sup>64</sup>

No que tange ao caso da superinteligência, Nick Bostrom considera “humanamente impossível formular explicitamente um conjunto de regras altamente complexas de forma detalhada, aplicá-las em diversas circunstâncias e ainda conseguir fazer isso da maneira correta na primeira tentativa de implementação.”<sup>65</sup>

O controle e os limites seriam estabelecidos através de meios técnicos com propósito voltado à obtenção de segurança, ou seja, dentro da órbita da autorregulação, balizada pela questão da segurança. Fala-se especialmente de (i) métodos de controle de capacidade, ou seja, mecanismos de controle de capacidade que buscam prevenir resultados indesejáveis *limitando o que a superinteligência pode fazer*; e (ii) métodos de seleção de motivação, isto é, ferramentas que objetivam prevenir resultados indesejáveis por meio de uma *predefinição do que a superinteligência terá vontade de fazer*.<sup>66</sup>

A primeira categoria de métodos divide-se em (i.a) *métodos de confinamento*, mediante os quais a superinteligência seria posicionada em um ambiente onde seria incapaz de causar prejuízos, dado que estaria em confinamento físico e com restrição de acesso a informações; (i.b) *métodos de*

---

<sup>64</sup> Um exemplo de questão jurídica que está em fase de amadurecimento é a submissão à legislação trabalhista dos contratos firmados entre empresas proprietárias de aplicativos de transporte de pessoas (Uber, Cabify, etc) e motoristas, que se mostra controversa no mundo jurídico, havendo decisões conflitantes até no mesmo país.

<sup>65</sup> BOSTROM, Nick. *Op. Cit.* p. 248.

<sup>66</sup> *Idem.* p. 232.

*incentivo*, assim entendidas as razões instrumentais fortemente convergentes para que a superinteligência não se comporte de forma prejudicial; (i.c) *inibição*, considerada como o conjunto de ferramentas para limitação das capacidades internas da superinteligência; e (i.d) *detonadores*, conceituados como os mecanismos implantados para detectar e reagir automaticamente a vários tipos de falhas de contenção ou tentativas de transgressão, inclusive com desligamento automático em caso de apuração de atividade perigosa.<sup>67</sup>

A segunda categoria de métodos (de seleção de motivação) seria composta pelos seguintes: (ii.a) *especificação direta*, consistente na formulação explícita de um objetivo ou de um conjunto de regras, consequencialista ou específicas, que deverão ser obedecidas – restando afastado, aqui, o sistema legal; (ii.b) *normatividade indireta*, ou seja, a configuração do sistema de modo que ele possa descobrir um conjunto apropriado de valores por si mesmo, utilizando-se de critérios formulados implicitamente ou indiretamente; (ii.c) *domesticidade*, isto é, a construção de um sistema com escopo severamente limitado de atividades, de sorte que teria objetivos modestos e pouco ambiciosos; e (ii.d) *ampliação*, que espelha a hipótese de seleção de um sistema que já tenha motivações substantivamente humanas ou benevolentes e, a partir daí, promove-se o aumento de seus poderes cognitivos para torná-lo superinteligente, assegurando que não se tornaria corrupto durante o processo.<sup>68</sup>

Raymond Kurzweill<sup>69</sup>, diferentemente de Bostrom, considera que “a inteligência é inerentemente impossível de controlar”, sendo “impraticável hoje elaborar estratégias que irão garantir de modo absoluto que a IA futura encarne a ética e os direitos humanos”<sup>70</sup>. Apesar disso, entende que a mais alta prioridade deve ser conferida ao avanço contínuo das tecnologias defensivas, tanto para enfrentar os perigos da IA, quanto os derivados da biotecnologia.<sup>71</sup> A regulação, por sua vez, deveria ser aplicada no que tange à *segurança* no campo da biotecnologia, que o Diretor do Google avalia como extensa na atualidade,

---

<sup>67</sup> Idem. p. 232.

<sup>68</sup> Idem. p. 246 e 252-253.

<sup>69</sup> KURZWEILL, Ray. *Op. Cit.* p. 700.

<sup>70</sup> Idem. p. 702-703.

<sup>71</sup> Idem. p. 720.

dados os padrões exigidos pela FDA nos EUA, por exemplo.<sup>72</sup> De outro lado, na indústria de *software* a autorregulação deveria permanecer como modelo prevalecente.<sup>73</sup>

O uso exclusivo de instrumentos de autorregulação, notadamente aqueles de matriz técnica, já recebe críticas robustas de pensadores como Benoit Frydman, Professor da Universidade Livre de Bruxelas. De acordo com as lições do jurista belga, haveria um *déficit* de legitimidade e validade numa governança baseada em *standards* técnicos e indicadores.<sup>74</sup> Ao contrário do que fazem crer as aparências, os *standards* não seriam normas puramente técnicas, facultativas e neutras, mas sim modos de governança, que dirigem as condutas e produzem efeitos políticos, fato que coloca em causa a sua validade e legitimidade porque a sua adoção (i) não empregou procedimento de caráter democrático; (ii) não sobreveio a partir de decisões tomadas por instituições democráticas; e (iii) não se subordinaria à conformidade com o Estado de Direito, sob o controle das jurisdições.<sup>75</sup>

Retornando à proposta de Nick Bostrom, vale chamar a atenção para a denominada normatividade indireta. O filósofo de Oxford alega que nenhuma teoria ética tem o apoio majoritário dos filósofos. Há risco de que todos estejam errados sobre o que seja a moralidade; ou o que é bom para a humanidade; e até mesmo quanto ao que realmente queremos. Uma escolha errada seria a cristalização de equívocos morais graves, impedindo o progresso ético que a superinteligência poderia alcançar.<sup>76</sup>

A solução seria delegar à superinteligência alguns dos raciocínios necessários para selecionar o valor a ser atingido, uma vez que seria cognitivamente superior à humanidade e, por isso, conseguiria enxergar além dos erros e confusões do pensamento humano.<sup>77</sup> É o chamado 'princípio da deferência epistêmica'. Uma forma de atender a esse princípio seria a adoção de uma 'vontade extrapolada coerente', ou seja, desenhando um esboço do que se deseja alcançar e a superinteligência buscar materializá-lo de modo

---

<sup>72</sup> Idem. p. 715.

<sup>73</sup> Idem. p. 715.

<sup>74</sup> p. 81.

<sup>75</sup> Idem. p. 86.

<sup>76</sup> BOSTROM, Nick. *Op. Cit.* p. 358-359.

<sup>77</sup> Idem. p. 359-360.

específico e concreto nos eventos futuros.<sup>78</sup> Outra forma possível de normatividade indireta seria a construção de uma IA que tivesse como objetivo fazer o que é moralmente certo, acreditando nas capacidades cognitivas superiores da IA para descobrir exatamente quais são as ações que se encaixariam nesse cenário, enunciado como 'proposta de retidão moral'.<sup>79</sup>

Eliezer Yudkowsky, notável pesquisador norte-americano sobre inteligência artificial e um dos inspiradores da obra de Nick Bostrom sobre superinteligência, também defende uma posição de autorregulação da superinteligência nesses moldes, mediante a criação das bases e procedimentos necessários para o que denomina 'IA Amigável' (*Friendly AI*)<sup>80</sup>. Yudkowsky inclusive rechaça com maior vigor a regulação como elemento favorável ao processo de construção de uma IA Amigável.<sup>81</sup>

Os esforços realizados pelos mais brilhantes pesquisadores da IA teriam capacidade superior para lidar com os problemas da IA Amigável do que o Poder Público, em especial o Congresso.<sup>82</sup> Inserir a política no processo de criação dessa IA benevolente retardaria o seu progresso, pois traria o elemento da influência política, do *lobby*, em substituição à troca livre de ideias e à legítima confrontação de opiniões.<sup>83</sup> Outro fator nocivo seria o ingresso do princípio ético de que "eu tenho uma ideia brilhante, então me dê poder sobre os outros", que tende a não funcionar bem como processo social, uma vez que para de funcionar quando outra pessoa procura dizer a mesma coisa.<sup>84</sup>

Mesmo que superados esses obstáculos, Yudkowsky avalia que a regulação mais plausível que o Congresso (norte-americano) conseguiria aprovar seria balizada por critérios antropomórficos (incompatíveis com o funcionamento de uma IA avançada) e adversariais (ou seja, eivados de

---

<sup>78</sup> Idem. p. 361-362.

<sup>79</sup> Idem. p. 370.

<sup>80</sup> YUDKOWSKY, Eliezer. "Artificial Intelligence as a Positive and Negative Factor in Global Risk". In: BOSTROM, Nick e CIRKOVIC, Milan M. **Global Catastrophic Risks**. Nova York: Oxford University Press, 2008. Disponível em: <https://intelligence.org/files/AIPosNegFactor.pdf>. Acesso em: 02/04/2021. p. 12-13.

<sup>81</sup> YUDKOWSKY, Eliezer. **Creating Friendly AI 1.0: The Analysis and Design of Benevolent Goal Architectures**. São Francisco: The Singularity Institute, 2015. Disponível em: <https://intelligence.org/files/CFAI.pdf> Acesso em: 08/05/2021. p. 208.

<sup>82</sup> Idem. P. 208.

<sup>83</sup> Idem. p. 210.

<sup>84</sup> Idem. p. 211.

preconceitos contra a IA), tornando os diplomas normativos completamente impraticáveis.<sup>85</sup>

Como se deduz a partir das linhas de raciocínio de Nick Bostrom, Raymond Kurzweill e Eliezer Yudkowsky, a abordagem dominante entre os mais destacados pensadores transumanistas é caracterizada pela autorregulação. Os profissionais da área de IA e as empresas com maior poder neste campo resolveriam, *interna corporis*, os dilemas e riscos da superinteligência, até mesmo com a delegação da solução para a própria IA forte geral. Submeter o tema ao debate político e aprovar um marco legal seria desaconselhável – a primeira medida por causa dos *lobbys* e a segunda em virtude da impraticabilidade das normas que adviriam no melhor dos cenários. O máximo de regulação que seria admitida, consistiria em diploma legal para a seara da segurança das aplicações, conforme Kurzweill, mas isso não valeria para o desenvolvimento de tecnologias envolvendo *softwares*.

Vale salientar, como providência de honestidade acadêmica, que a ótica transhumanista exposta não reflete uma orientação geral, integral, desta corrente de pensamento. Trata-se da visão de alguns dos seus proeminentes defensores, que foram responsáveis por redigir linhas de relevo em livros e artigos sobre a questão versada, com notáveis conhecimento de causa e contato com a realidade do universo da IA.

## **5. A ABORDAGEM REGULATÓRIA DA SUPERINTELIGÊNCIA SOB A ÓTICA DOS DIREITOS FUNDAMENTAIS**

O enfoque baseado no direito fundamental à proteção de dados pessoais para a regulação da superinteligência recebe influxos de ordem diferente da ótica transhumanista, e está comprometido com propósitos dissonantes daqueles sustentados por essa corrente de pensamento, em especial quando se fala das ideias sintetizadas no capítulo anterior.

Mesmo que os autores reconheçam os riscos e perigos da superinteligência, que podem ameaçar a existência da humanidade<sup>86</sup>, a posição

---

<sup>85</sup> Idem. p. 211.

<sup>86</sup> Eliezer Yudkowsky, em publicação de 2008, colocava o desenvolvimento avançado da inteligência artificial como um risco catastrófico global. YUDKOWSKY, Eliezer. “Artificial

defendida afasta o tema do debate público realizado nas esferas políticas da sociedade. Almejam um diálogo aberto e público que não possa resultar em decisões políticas que desagüem na aprovação de marcos legais.

Ou seja, a sociedade pode acompanhar as discussões, mas a regulação, normalmente formulada pelos legisladores eleitos pela população nos regimes democráticos, é medida nociva, que deve ser evitada – Yudkowsky chega a mencionar expressamente o papel negativo que o Congresso norte-americano poderia desempenhar se colocasse em pauta as ideias sobre uma ‘IA Amigável’.

A abordagem ancorada nos direitos fundamentais é diametralmente oposta. Favorece, incentiva e defende a submissão de todos os temas sensíveis para o desenvolvimento econômico e social à apreciação pelas instituições constitucionais, mesmo que dotados de alguns elementos técnicos complexos. Rechaça-se a interdição, o veto, à discussão na arena política, em especial quando se trata do Poder Legislativo, de matérias que envolvam riscos elevados para o bem-estar e o futuro da sociedade, e até da humanidade como um todo.

Os regimes democráticos contemporâneos estão alicerçados no Estado de Direito e nos direitos fundamentais, cuja materialização depende de um qualificado e inclusivo diálogo público, conduzido nas arenas institucionais legitimadas nas Constituições Políticas.

O cerne da democracia é a preservação de um complexo e plural processo político de co-decisão, como dizia Pontes de Miranda<sup>87</sup>. Democracia, segundo Zagrebelsky, é discussão, raciocinar em conjunto, filologia, e não misologia<sup>88</sup>. Nesta engrenagem, prevalece a vontade da maioria no desenho das políticas públicas e das leis, com fulcro na atuação de representantes legitimados por pleitos eleitorais. Um regime político que consagra a soberania popular. O governo do povo, pelo povo e para o povo, na clássica acepção de Abraham Lincoln.

---

Intelligence as a Positive and Negative Factor in Global Risk”. In: BOSTROM, Nick e CIRKOVIC, Milan M. **Global Catastrophic Risks**. Nova York: Oxford University Press, 2008. Disponível em: <https://intelligence.org/files/AIPosNegFactor.pdf>. Acesso em: 02/04/2021.

<sup>87</sup> PONTES DE MIRANDA, Francisco Cavalcanti. **Democracia, liberdade e igualdade**. Os três caminhos. Campinas: Bookseller, 2002.p. 191.

<sup>88</sup> ZAGREBELSKY, Gustavo. **Imparare democrazia**. Torino: Einaudi, 2007. p. 21.

Com a ascensão dos direitos fundamentais e o reconhecimento da normatividade das Constituições Políticas depois da Segunda Guerra Mundial, o regime democrático recebeu influxos que conduziram ao hodierno Estado Constitucional, sob o qual as garantias fundamentais individuais e sociais foram alçadas a parâmetros de validade material do ordenamento jurídico positivo – que antes era apreciado somente pela ótica da legalidade formal – e das políticas públicas.

As bases das democracias foram reforçadas com o ideal da rigidez constitucional, amparada nas aludidas garantias magnas, a partir de um discurso originado nas declarações, convenções e tratados internacionais de direitos humanos, que foi respaldado pelas teorias dos direitos fundamentais e amplamente disseminado pelo mundo.

Neste paradigma, os direitos fundamentais passam a conferir legitimidade para toda a ordem jurídica. Sobressai um núcleo duro, uma esfera do indecível, nos regimes políticos, oriundo dos mandamentos constitucionais cardeais, denominados cláusulas pétreas no Brasil.<sup>89</sup> De acordo com Flávia Piovesan, “não há direitos humanos sem democracia, tampouco democracia sem direitos humanos”<sup>90</sup>.

Não se pode perder de vista, olvidar, contudo, que os direitos fundamentais e o próprio Estado Constitucional representam conquistas históricas, desde as revoluções liberais até a contemporaneidade.

Por serem históricos, cada geração deve se posicionar para manter o respeito a esses preceitos civilizatórios e, se possível, avançar na implementação de maiores camadas jurídicas protetivas para toda a população.

A proteção da vida, das liberdades fundamentais, da igualdade, da dignidade humana, da saúde, da segurança, da privacidade e da proteção de dados pessoais, todos direitos estampados no art. 5º da Constituição Federal devem guiar o legislador brasileiro a uma regulação ancorada nos direitos fundamentais.

---

<sup>89</sup> GARCIA, Marcos Leite. “A concepção tridimensional dos direitos fundamentais de Gregorio-Peces Barba: reflexões na busca de critérios para o conceito de direitos humanos”. In: CRUZ, Paulo; BRANDÃO, Paulo; OLIVIERO, Maurizio. **O Direito contemporâneo e diálogos científicos UNIVALI e Perugia**. Itajaí: UNIVALI, 2016. p. 14.

<sup>90</sup> PIOVESAN, Flávia. **Direitos humanos e justiça internacional**. 3ª Ed. São Paulo: Saraiva, 2012. p. 44.

Trata-se de direitos que provavelmente serão vulnerados pelas novas tecnologias ora versadas, mas que sobreviveram, por méritos das gerações anteriores, em face das ameaças trazidas pelas tecnologias desenvolvidas nas revoluções industriais anteriores.

Essa é a matriz, a inspiração, da abordagem de direitos fundamentais que deve orientar a busca pelo caminho mais apropriado para o enfrentamento dos desafios trazidos pelos acelerados desenvolvimentos em matéria de inteligência artificial e neurotecnologias.

O campo da superinteligência, seja ela atingida mediante esforços de IA ou de neurotecnologia (ou da combinação das duas áreas), está começando a ser desbravado por estudiosos e profissionais com conhecimentos altamente especializados em TI, neurotecnologia e biotecnologia, motivo pelo qual a regulação não tem sido objeto de discussões marcadas por maior profundidade jurídica.

A importância da questão, decorrente dos fantásticos benefícios e perigosos riscos dela advindos, reclama uma abordagem mais detalhada e comprometida com os vetores jurídicos dirigentes das democracias contemporâneas no que atine à regulação. O entendimento pela autorregulação pura, com sugestões até de interdição ao debate legislativo e à promulgação de novos marcos legais, sustentados pelos *experts* dos citados nichos tecnológicos, merece um exame sob outra perspectiva, iluminada pelos holofotes dos direitos fundamentais e, em particular, do Direito da Proteção de Dados Pessoais.

Como primeira medida, urgente e essencial para calibrar o enfoque com base nos direitos fundamentais, destaca-se a necessidade de levar o tema para o debate público, inclusive para as Casas Legislativas. A legitimação democrática é o que garante a validade material, substancial, das diretrizes – privadas ou públicas – que regem tecnologias e processos complexos e sensíveis do porte de uma superinteligência.

Em segundo lugar, impende analisar os impactos desta ferramenta sobre os direitos fundamentais como um todo. A livre escolha dos cidadãos a respeito do seu projeto de vida, do trabalho que executarão, dos alimentos comprarão e consumirão, do cônjuge e dos amigos que escolherão, da época em que terão filhos, entre outros aspectos chave da vida privada e da intimidade,

podem ser influenciados severamente – até manipulados – por uma superinteligência.

A fragilização da igualdade e da dignidade por conta de discriminações em matéria de acesso ao mundo do trabalho, de obtenção de crédito para financiamento da casa própria e bens da vida em geral, de oferta de planos de saúde e de seguros, são algumas das circunstâncias que podem ser acentuadas – haja vista que já ocorrem atualmente – no cenário de uma vindoura IA forte geral<sup>91</sup>.

A integridade corporal também estará em jogo com a progressiva instalação de dispositivos, implantes, nanorrobôs, no cérebro e noutros órgãos de pessoas que se disponham a colaborar com experimentos neurotecnológicos. Haverá riscos de corrupção destes dispositivos por vírus de *software* e técnicas de hackeamento<sup>92</sup>, quiçá por sistemas superinteligentes utilizados para fins nocivos, tornando imperativas as preocupações jurídicas com a autodeterminação, o livre desenvolvimento da personalidade, a privacidade, a segurança e demais direitos fundamentais destes cidadãos.

Existem, portanto, situações práticas concretas sensíveis, de ordem individual, que permitem afirmar que os direitos fundamentais de liberdade, dignidade e igualdade podem ser seriamente vulnerados pelo surgimento e atuação de uma superinteligência.

Eventos nocivos de ordem coletiva também podem sobrevir e causar lesões a direitos fundamentais de grupos específicos ou de toda a sociedade. Discriminações raciais ou étnicas genéricas promovidas por sistemas de crédito social superinteligentes são claramente possíveis, tendo em vista já estar em funcionamento um sistema dessa natureza na China. A má pontuação nesse

---

<sup>91</sup> De acordo com Borges e Graminho, “Atualmente, a problemática dos dados tendenciosos causa uma inquietude em relação aos softwares utilizados por departamentos de polícia, que podem resultar em um policiamento desigual ou discriminatório. Isso implicará em impactos nos direitos inerentes da dignidade humana, como o direito à liberdade de viver sem discriminação, direito à liberdade pessoal, o direito à vida e segurança pessoal e o direito à privacidade. (...)”

Dessa forma, também corre-se o risco de algoritmos reproduzirem comportamentos discriminatórios e não se pode esquecer que um pequeno número de empresas de tecnologia concentra o poder sobre as novas formas de tecnologias e beneficiam-se do desenvolvimento propiciado por elas.” BORGES, Gustavo e GRAMINHO, Vivian. “Inteligência artificial e direitos humanos: interfaces regulatórias e os desafios”. In: BARBOSA, Mafalda Miranda; BRAGA NETTO, Felipe; SILVA, Michael César; FALEIROS JÚNIOR, José Luiz de Moura. *Op. Cit.* 952-955.

<sup>92</sup> KURZWEILL, Raymond. *Op. Cit.* p. 689.

mecanismo de alcance coletivo, nos dias correntes, produz consequências severas, como a inclusão em “listas negras” de maus cidadãos, a recusa de crédito bancário, a proibição de utilizar determinados serviços sociais, acesso limitado a empregos e a serviços públicos, proibição de matricular os filhos em escolas particulares, vedação à compra de bens de luxo e até a emissão de mensagens para os ‘bons cidadãos’ alertando-os quando estão buscando se comunicar com cidadãos ou empresas que figuram nas listas negras.<sup>93</sup>

A Corte Europeia de Direitos Humanos, no caso Shimovolos c. Rússia (30194/09), consignando a ilicitude da vigilância estatal e inclusão do nome de um ativista de direitos humanos, Sergey Mikhaylovich Shimovolos, em uma *black list* da polícia russa com base em ato normativo não publicado, e, por consequência, desconhecido pelo titular dos dados, violando as garantias e os direitos fundamentais que devem ser respeitados nos ordenamentos internos, com previsibilidade, cognoscibilidade e publicidade do processo legislativo.<sup>94</sup>

A própria cultura, os hábitos e a história de povos tradicionalmente vulneráveis ou discriminados podem ser ameaçados pelas manipulações e providências severamente invasivas propiciadas por um sistema superinteligente, inclusive com possibilidade de extermínio, caso seja direcionada a IA forte geral para o campo bélico, mediante o uso de armas autônomas dotadas e controladas por IA.

As ameaças ao núcleo dos direitos fundamentais, por conseguinte, manifestam-se igualmente na órbita coletiva, estatal, não se limitando a possíveis atos ou procedimentos isolados que gerem violações da liberdade, da dignidade, da igualdade e de outros direitos magnos contra certos indivíduos.

Neste contexto, surge legítima a discussão sobre medidas que sejam capazes, nos dias de hoje, de prevenir a cidadania contra lesões de direitos fundamentais que advirão dos desdobramentos tecnológicos atuais que têm aptidão para conduzir à superinteligência.

---

<sup>93</sup> PRISQUE, Emmanuel. **O Sistema de Crédito Social chinês**. Fundação Fernando Henrique Cardoso, 17 de setembro de 2020. Disponível em: [O Sistema de Crédito Social chinês | by Fundação FHC | Fundação FHC | Medium](#). Acesso em: 05/05/2021.

<sup>94</sup> CORTE EUROPEIA DE DIREITOS HUMANOS. **Caso Shimovolos v. Rússia (30194/09)**. 21 de junho de 2011. Disponível em: <https://hudoc.echr.coe.int/eng#%7B%22appno%22:%5B%2230194/09%22%5D,%22itemid%22:%5B%22001-105217%22%5D%7D>. Acesso em: 14/07/2023.

Na esteira desse debate, revela-se necessária a formulação de um ou mais diplomas legais nacionais para regular, desde já, essas áreas, de sorte a viabilizar maior segurança e conformidade com os direitos fundamentais para os cidadãos quanto aos avanços da inteligência artificial e das neurotecnologias – embora se tenha ciência de que a abordagem pautada por princípios éticos, e não pela regulação legal, seja defendida por vários juristas especializados na temática ora enfrentada e possua subsídios fortes no panorama dos países ricos.<sup>95</sup>

Não se pretende interromper ou burocratizar o progresso tecnológico, mas tornar esse caminho mais seguro e civilizado até que se alcance o estágio da superinteligência. A vigilância democrática, cidadã, inclusive mediante a aprovação de leis e a criação de órgãos de supervisão, é o remédio mais salutar e legítimo para se garantir a liberdade de todos no futuro e evitar a instauração de autocracias digitais governadas por elites de burocratas e especialistas, avessos ao debate público e negligentes com os interesses e direitos básicos da população.

Na seara da economia contemporânea baseada em algoritmos, diretamente vinculada com o tema ora versado, a defesa de um enfoque regulatório por meio da promulgação de diploma legal foi solidamente promovida

---

<sup>95</sup> Segundo Mulholland e Frajhof, “na ausência de regulação jurídica sobre o assunto, a observação de princípios éticos pode ser importante para maximizar os benefícios da IA e diminuir seus riscos.<sup>41</sup> A questão que se coloca é: quais princípios éticos? A utilização de princípios éticos como instrumento para a regulação da Inteligência Artificial, apesar de não conter uma natureza coercitiva e sancionatória típica da regulação jurídica, permite a criação de guias deontológicos que serão constituídos como a razão prima facie e o fundamento para o desenvolvimento e implementação da IA. Nesse caso, a governança da Inteligência Artificial será baseada em princípios que imporão uma regulação leve (soft law ou soft regulation), na medida em que servirão de limite para a expansão da IA, sem, contudo, prever uma penalidade específica em caso de desvio do padrão constituído. Tanto a União Europeia quanto outros organismos internacionais - como, por exemplo, a OCDE e a Access Now - e instituições privadas - como, por exemplo, a IBM e a Microsoft - instituíram guias de recomendação para a adoção de princípios éticos para a regulação de IA. Os princípios éticos indicados pelos documentos consultados partem de um mesmo pressuposto, qual seja, a necessidade de fomentar o desenvolvimento de uma Inteligência Artificial confiável e que seja auditável, permitindo que seus processos sejam conhecidos e controlados - ou controláveis - pelo ser humano.” MULHOLLAND, Caitlin e FRAJHOF, Isabella. “Entre as leis da robótica e a ética: regulação para o adequado.” *In* BARBOSA, Mafalda Miranda; BRAGA NETTO, Felipe; SILVA, Michael César; FALEIROS JÚNIOR, José Luiz de Moura. **Direito Digital e Inteligência Artificial**. São Paulo: Editora Foco, 2021. Edição do Kindle. p. 139.

por Deven Desai e Joshua Kroll no artigo “*Trust but verify: a guide to algorithms and the Law*”, publicado no *Harvard Journal of Law & Technology*<sup>96</sup>.

Os Professores, atrelados respectivamente ao Departamento de Direito do *Virginia Institute of Technology* e à *UC Berkeley School of Information*, indicam quatro benefícios para justificar a regulação mediante a aprovação de um arcabouço legal: (1) a regulação legal pode evitar o problema de aplicar abordagens inadequadas de regimes regulatórios anteriores ou exigindo resultados que não são possíveis; (2) pode abordar o dinamismo da indústria e as dificuldades de análise de *software*, provendo requisitos que os técnicos e cientistas da computação entendem e podem implementar; (3) assim como ocorreu com a regulação anterior de habilitação, crédito e emprego, a regulamentação legal de *software* e algoritmos pode oferecer clareza sobre o que é realizável e o que deve ser oferecido aos reguladores para demonstrar conformidade; e (4) se aqueles que devem ser regulados se opõem, caberá a eles o ônus de mostrar por que as soluções propostas e tecnicamente informadas não funcionam. Com a regulação legal, argumentam os Professores norte-americanos, surgirão muitos ganhos de automação advirão com mais segurança, englobando melhores níveis de governança.

De fato, revela-se indispensável a consecução de esforços combinados entre profissionais e estudiosos dos vários campos do saber envolvidos com a regulação da superinteligência. Juristas, cientistas da computação, filósofos, economistas e empresários não podem buscar o monopólio do diagnóstico nem do prognóstico sobre o tema. Urge que se desenvolvam conjuntamente os marcos normativos para a inteligência artificial e as neurotecnologias, levando em consideração as contribuições de todas as áreas.

Estabelecer o perímetro do que é legítimo e consonante com os direitos fundamentais constitui tarefa multidisciplinar, de modo a oferecer segurança jurídica sobre quais são os experimentos, iniciativas e negócios proibidos, e aqueles que precisam cumprir certos requisitos para serem

---

<sup>96</sup> DESAI, Deven e KROLL, Joshua. “Trust but verify: a guide to algorithms and the Law”. In: **Harvard Journal of Law & Technology**, Vol. 31, n.1, 2017. Disponível em: <https://jolt.law.harvard.edu/assets/articlePDFs/v31/31HarvJLTech1.pdf>. Acesso em: 05/05/2021.

realizados, variando em conformidade com o risco gerado (inaceitável, alto, moderado ou baixo).

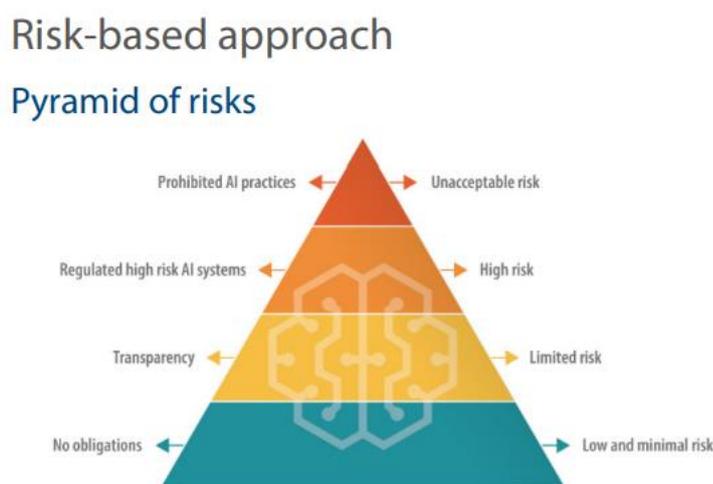
Assim, com fulcro na teoria dos direitos fundamentais e sua ligação íntima e essencial com o regime democrático e o Estado Constitucional, bem como aderindo às premissas e propostas elaboradas por Deven Desai e Joshua Kroll, surge como melhor alternativa para regular os campos que podem conduzir à superinteligência, notadamente a IA e a neurotecnologia, a discussão pública e aprovação de marcos legais, viabilizando o *enforcement* de padrões civilizatórios mínimos para proteger a sociedade em geral, com especial atenção aos mais vulneráveis.

Em relação aos futuros diplomas legais, a abordagem de direitos fundamentais indica a possibilidade de regulação separada de cada uma dessas áreas, inteligência artificial e neurotecnologias.

Na União Europeia, está em fase avançada o *AI Act*, aprovado em junho pelo Parlamento Europeu. Revela-se interessante, no que atine aos riscos extremos característicos da temática da superinteligência, expor as hipóteses de proibição de sistemas de IA.

Como se sabe, o *AI Act* europeu foi desenhado sob a ótica da abordagem baseada em riscos (*risk-based approach*):

Figura 2 - *Risk-based approach*



97

<sup>97</sup> EUROPEAN PARLIAMENT. **Artificial intelligence act**. Disponível em: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS\\_BRI\(2021\)698792\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf). Acesso em: 15/07/2023

As práticas proibidas em matéria de IA previstas no diploma europeu são as seguintes, previstas no Título II, artigo 5 (*prohibited artificial intelligence practices*):

(1) A colocação no mercado, entrada em serviço ou utilização de determinados sistemas de IA com o objetivo ou o efeito de distorcer materialmente o comportamento humano, por meio do qual ou danos psicológicos possam ocorrer, devem ser proibidos. Esta limitação deve ser entendida como incluindo neurotecnologias assistidas por sistemas de IA que são usados para monitorar, usar ou influenciar dados neurais coletados por meio de interfaces cérebro-computador na medida em que distorcem materialmente o comportamento de uma pessoa física em uma maneira que cause ou possa causar a essa pessoa ou a outra pessoa dano significativo (art. 5, 1, 'a');

(2) a colocação no mercado, entrada em serviço ou utilização de um sistema de IA que explora qualquer uma das vulnerabilidades de uma pessoa ou de um grupo específico de pessoas, incluindo características de tal indivíduo ou grupo de pessoas conhecidas ou traços de personalidade previstos ou situação social ou econômica, idade, condição física ou capacidade mental, com o objetivo ou efeito de distorcer materialmente a comportamento dessa pessoa ou de uma pessoa pertencente a esse grupo de maneira que cause ou seja susceptível de causar a essa pessoa ou a outra pessoa um dano significativo (art. 5, 1, 'b');

(3) a colocação no mercado, colocação em serviço ou utilização de dados biométricos de sistemas de categorização que categorizam pessoas físicas de acordo com atributos ou características protegidos, ou com base na inferência desses atributos ou características. Esta proibição não se aplica a sistemas de IA destinado a ser utilizado para fins terapêuticos aprovados com base em consentimento informado específico dos indivíduos que estão expostos a eles ou, se for o caso, de seu responsável legal (art. 5, 1, 'ba');

(4) a colocação no mercado, colocação em serviço ou uso de sistemas de IA para pontuação, avaliação ou classificação social de pessoas físicas ou grupos delas durante um certo período de tempo com base em seu comportamento social ou conhecimento pessoal, inferido ou previsto ou características de personalidade, com a pontuação social levando a um ou ambos a seguir:

(4.i) tratamento prejudicial ou desfavorável de certas pessoas físicas ou seus grupos em contextos sociais que não estão relacionados com os contextos em quais os dados foram originalmente gerados ou coletados (art. 5, 1, 'c', i);

(4.ii) tratamento prejudicial ou desfavorável de certas pessoas físicas ou seus grupos injustificados ou desproporcionais à sua condição social comportamento ou sua gravidade (art. 5, 1, 'c', ii);

(5) a colocação no mercado, entrada em serviço ou utilização de um sistema de IA para fazer avaliações de risco de pessoas físicas ou grupos de pessoas físicas, a fim de avaliar o risco de uma pessoa física de ofender ou reincidir ou de prever a ocorrência ou recorrência de um crime real ou potencial ou contraordenação baseada na caracterização de uma pessoa singular ou em avaliação de traços e características de personalidade, incluindo a localização ou comportamento criminoso anterior de pessoas físicas ou grupos de pessoas físicas (art. 5, 1, 'da');

(6) A colocação no mercado, entrada em serviço ou utilização de sistemas de IA que criem ou expandir bancos de dados de reconhecimento facial por meio da raspagem não direcionada de imagens faciais da internet ou imagens de CCTV (art. 5, 1, 'db');

(7) a colocação no mercado, colocação em serviço ou uso de sistemas de IA para inferir emoções de uma pessoa física nas áreas de aplicação da lei, fronteira gestão, no local de trabalho e nas instituições de ensino (art. 5, 1, 'dc');  
e

(8) a colocação em serviço ou uso de sistemas de IA para a análise de registros filmagem de espaços acessíveis ao público por meio de biometria remota 'pós' sistemas de identificação, salvo se estiverem sujeitos a autorização pré-judicial em conformidade com o direito da União Europeia e estritamente necessário para a pesquisa direcionada relacionada com uma infração penal grave específica, conforme definido no artigo 83.º, n.º 1, do TFUE que já ocorreu para fins de aplicação da lei (art. 5, 1, 'e').

Na órbita brasileira, seguindo a linha europeia, o Projeto de Lei nº 2338/2023 estabelece, em seu art. 14, três vedações à implementação e uso de sistemas de inteligência artificial: (i) que empreguem técnicas subliminares que tenham por objetivo ou por efeito induzir a pessoa natural a se comportar de

forma prejudicial ou perigosa à sua saúde ou segurança ou contra os fundamentos da futura Lei da IA; (ii) que explorem quaisquer vulnerabilidades de grupos específicos de pessoas naturais, tais como as associadas a sua idade ou deficiência física ou mental, de modo a induzi-las a se comportar de forma prejudicial a sua saúde ou segurança ou contra os fundamentos da futura Lei da IA; e (iii) pelo poder público, para avaliar, classificar ou ranquear as pessoas naturais, com base no seu comportamento social ou em atributos da sua personalidade, por meio de pontuação universal, para o acesso a bens e serviços e políticas públicas, de forma ilegítima ou desproporcional.

Nos projetos europeu e brasileiro, identificam-se várias hipóteses de proibições ao uso e implementação de sistemas de IA. O ato normativo europeu tem maior amplitude no campo das vedações, demonstrando um nível maior de receio com os riscos e impactos que podem advir com o avanço sem qualquer regulação dos sistemas de IA.

Somente com a promulgação de um marco normativo dessa natureza é que se terá um ferramental jurídico mais apropriado para minorar as possibilidades de os sistemas de IA conduzirem de forma acrítica e altamente perigosa para a superinteligência.

Não obstante, cumpre alertar que a mera edição de uma lei não significa que a ordem das coisas mudará para o melhor caminho, da centralidade do ser humano. É indispensável que as normas jurídicas, como objetos culturais, tenham respaldo em uma visão ética crítica, sofisticada, da maior parte da sociedade, dos órgãos reguladores e das instâncias políticas a respeito do desenvolvimento da IA.

Sem um compromisso humanista nas esferas ética, cultural, educacional/pedagógica e política, as normas jurídicas se esvaziam e não induzirão os comportamentos para o que pretende o legislador. Mesmo que as normas éticas e as práticas culturais, no sentido da autorregulação pura, não sejam suficientes para bem conduzir os desdobramentos da IA forte, como evidenciado acima, as normas jurídicas, sejam princípios ou regras, precisam de patamares éticos e culturais elevados para atingirem os seus objetivos.

No campo das neurotecnologias, é mister enfatizar a necessidade de um desenvolvimento teórico mais robusto no campo jurídico para a formulação de direito fundamental à inviolabilidade cerebral, a ser posto na Constituição

Federal, como meio de colocar freios civilizatórios às neurotecnologias desenvolvidas e aplicadas por instituições públicas e corporações privadas.

O Senado do Chile foi pioneiro na aprovação de proposta de reforma constitucional para introduzir norma desse jaez no Texto Magno. O projeto, Boletín 13827-19, que ingressou na tramitação legislativa em 07/10/2020, foi cancelado em dezembro de 2020 pela Casa Alta do Legislativo Chileno, impulsionado pelo Senador Guido Girardi. O objetivo é alterar a redação do art. 19, incisos 1º e 4º, da Constituição Chilena, para o seguinte teor:

“1º.- El derecho a la vida, la integridad física y síquica y la inviolabilidad cerebral. La actividad neuronal sólo puede accederse, registrarse o intervenirse en los casos previstos por la ley con el consentimiento de la persona afectada o de quien lo represente.”

b) En el numeral 4º, intercálese entre la frase “datos personales.” y la frase “El tratamiento”, la siguiente frase: “El derecho a la protección de la intimidad mental y de los datos neuronales.”.<sup>98</sup>

A *Universidad de Chile* ofereceu apresentação com sólida fundamentação para justificar a proposição. Inicialmente, enfatizou o bem-estar social e produtivo que a ciência e a tecnologia podem proporcionar, bem como os riscos de que os avanços carreguem um potencial disruptivo que possa afetar direitos. Na sequência, apontou a ameaça de vulneração séria ao princípio da igualdade, uma vez que o progresso científico já permite aumentar a capacidade cognitiva ou motora das pessoas, sendo necessário promover acesso igualitário e sem discriminação a essas benesses, sob pena de exacerbar desigualdades.<sup>99</sup>

A privacidade foi abordada com destaque na apresentação. Consigna-se que os dados se converteram em elementos aptos a gerar significativo valor econômico, trazendo consigo a paulatina perda de controle sobre os dados pessoais e um debate sobre os limites da privacidade. Após referir aos dilemas sobre a decodificação do genoma humano, destaca-se que uma vulneração aos neurodados implica uma possível anulação da autonomia e da privacidade, afinal o cérebro é o único lugar onde se pode manter os pensamentos e tomar decisões. Cabe viabilizar resguardos e medidas de

---

<sup>98</sup> SENADO DA REPÚBLICA DO CHILE. **Boletín 13827-19**. Disponível em: [Senado - Tramitación de proyectos](#). Acesso em: 06/05/2021.

<sup>99</sup> Idem.

segurança com magnitude proporcional para evitar a imposição de lesões a esses dados.<sup>100</sup>

Sugeriu-se, ainda, que a aplicação da inteligência artificial à neurotecnologia seja avaliada com muita cautela, pois vieses e preconceitos poderiam ser implementados diretamente no cérebro humano com consequências por enquanto desconhecidas.<sup>101</sup> A perda de autonomia, ou seja, da capacidade de cada ser humano se governar, é uma das mais assustadoras.

No que diz respeito à identidade das pessoas, relatou-se que atualmente existem casos de pessoas tetraplégicas que têm eletrodos inseridos no cérebro para facilitar suas vidas e resolver alguns de seus problemas. Até o momento, os efeitos colaterais são mínimos. No entanto, o documento propõe que se deve garantir que a conexão de dispositivos ao cérebro humano não afete às pessoas, de sorte a que mantenham sua personalidade e não tenham seu livre arbítrio ameaçado.

O teor da proposta constitucional chilena e das razões legislativas que a sustentam demonstra que a manipulação da mente através das neurotecnologias não constitui problema das próximas décadas. Trata-se de questão do presente. Empresas e governos estão trabalhando com essas tecnologias e não se verifica um nível adequado de transparência sobre as iniciativas, seus objetivos e as consequências produzidas.

A temática da superinteligência vem para legitimar e robustecer o esforço pautado pela intenção de aumentar o perímetro de proteção e defesa dos direitos fundamentais na ordem constitucional. Sem transparência e segurança no desenvolvimento atual dessas tecnologias, quando sobrevierem sistemas de IA forte geral não se terá uma preparação e as cautelas indispensáveis para resguardar os indivíduos e a coletividade.

O texto da proposta de reforma constitucional chilena e suas razões demandam reflexão e debate público qualificado e aberto no Brasil. Os riscos, ameaças e lesões vislumbrados pelos legisladores chilenos estão presentes em *terrae brasilis*.

---

<sup>100</sup> Idem.

<sup>101</sup> Idem.

Urge que a doutrina jurídica, os parlamentares nas Casas Legislativas nacionais e os estudiosos e profissionais da área de neurotecnologia promovam discussões mais aprofundadas sobre o assunto.

Em primeiro lugar, a inviolabilidade do cérebro e dos dados neurais dos cidadãos deve ser introduzida no rol de direitos fundamentais da Constituição Federal, estabelecendo um norte para todo o ordenamento jurídico sobre o assunto. A proteção dentro da rede dos direitos individuais *magno* viabilizará sólido embasamento para uma posterior regulamentação em nível infraconstitucional.

Em segundo lugar, no patamar ordinário, tem-se atualmente no parlamento brasileiro com maior destaque o Projeto de Lei 522/2022, que tramita na Câmara dos Deputados. O objetivo declarado é regulamentar o uso de dados neurais, mediante alterações na Lei Geral de Proteção de Dados Pessoais. A proposta cria hipóteses de tratamento de dados neurais inclusive sem o consentimento dos cidadãos. Veda o uso desses dados que possa causar danos à identidade individual do titular dos dados, prejudicar sua autonomia ou sua integridade psicológica.

Não há, como ocorre com os marcos normativos de IA ora sob discussão em todo o mundo, um estudo ancorado em regulação por direitos fundamentais combinada com abordagem baseada em riscos no que atine aos dados neurais. Faltam indicações sobre operações com riscos proibitivos, com riscos elevados e com riscos moderados e leves.

Mostra-se necessário avançar para definir melhor as circunstâncias proibitivas e de alto risco, estabelecer obrigações pormenorizadas em face das grandes empresas do setor de neurotecnologia, exigir programas de *compliance* com especificidades dessa área, para que se possa ter condições de prevenir o surgimento superinteligências maléficas decorrentes das neurotecnologias.

De outro lado, também não existe uma principiologia e um conjunto de direitos para resguardar a cidadania em matéria de dados neurais e neurotecnologias. A falta de uma maior sistematização normativa voltada para a proteção da pessoa humana representa uma relevante lacuna, que reduz o efetivo alcance regulatório em face de desenvolvimentos indevidos nessa área, tanto por entidades privadas quanto por entes públicos.

Revelam-se justificadas, assim, inovações jurídicas na Constituição Federal, mediante o acréscimo de direito fundamental, e na legislação infraconstitucional, através de diploma legal que seja capaz de articular princípios, direitos aos cidadãos, obrigações aos desenvolvedores, definição e regramento dos riscos, e regras gerais em matéria de neurotecnologias, de sorte a, entre outros objetivos, evitar esforço contra o desenvolvimento de superinteligências capazes de afrontar os direitos fundamentais e ameaçar a sobrevivência digna da própria humanidade.

Logo, os direitos constitucionais à vida, à liberdade, à igualdade, à dignidade, à saúde, à segurança, à privacidade e à proteção de dados pessoais, aliados (i) a um novo direito constitucional à inviolabilidade do cérebro e dos dados neurais e (ii) a dois marcos normativos infraconstitucionais – sobre IA e sobre neurotecnologias -, constituem instrumentos da ordem jurídica que merecem ter seus âmbitos protetivos alargados para que possam incidir nos campos da IA e das neurotecnologias para fazer valer os direitos fundamentais e garantir a construção de uma sociedade composta por cidadãos dotados de grau mínimo de liberdade, dignidade e igualdade, afastando-se, dentro do possível, os riscos e ameaças à real autonomia dos indivíduos para tomar decisões norteadoras das suas vidas, sem a manipulação de IAs ou de ferramentas neurotecnológicas abusivas.

## **7. CONCLUSÃO**

O presente trabalho abordou a temática da regulação do fenômeno da superinteligência, tomando como referências o pensamento transhumanista, em especial as contribuições de Nick Bostrom, Raymond Kurzweill e Eliezer Yudkowsky, e os influxos dos direitos fundamentais.

O assunto apresenta grande complexidade porque se trata de fenômeno do porvir, ou seja, de sistema que ainda não foi criado. Isso é agravado porque não se sabe exatamente qual(s) campo(s) de estudos e pesquisa viabilizará o desenvolvimento efetivo da superinteligência. Em vista disso, partiu-se do pressuposto, alicerçado nos autores transhumanistas, de que as áreas de inteligência artificial e de neurotecnologia são as candidatas favoritas a propiciar o surgimento dessa ferramenta.

Formado este cenário, o estudo adentrou na proposta regulatória sustentada pelos principais representantes do transhumanismo que enfrentaram o tema. A autorregulação prevaleceu como a solução dessa escola de pensamento. A sofisticada forma de implementação e suas consequências figuram em capítulo próprio com riqueza de detalhes, em especial quando se teoriza a delegação das decisões éticas para a própria superinteligência, num exercício de deferência epistêmica dos humanos em favor das máquinas.

Em contraposição, apresentou-se uma visão regulatória que parte dos direitos fundamentais, colocando o ser humano como o centro do desenvolvimento das novas tecnologias, inclusive nas searas da inteligência artificial e das neurotecnologias, que podem conduzir ao surgimento de sistemas superinteligentes.

À luz de contribuições europeias sobre inteligência artificial e chilenas sobre neurotecnologias, apontou-se a regulação baseada em direitos como o caminho mais apropriado no quadro democrático de um Estado de Direito contemporâneo, defendendo regulação, constitucional e legal, amparada na autonomia, na liberdade, na segurança, na saúde e na integridade física como forma de mitigar os riscos e danos que podem advir de futuros sistemas abusivos dotados de superinteligência geral.

## BIBLIOGRAFIA

- <sup>1</sup> PIOVESAN, Flávia. **Direitos humanos e justiça internacional**. 3ª Ed. São Paulo: Saraiva, 2012.
- BORGES, Gustavo e GRAMINHO, Vivian. “Inteligência artificial e direitos humanos: interfaces regulatórias e os desafios”. *In*: BARBOSA, Mafalda Miranda; BRAGA NETTO, Felipe; SILVA, Michael César; FALEIROS JÚNIOR, José Luiz de Moura. *Op. Cit.* 952-955.
- BOSTROM, Nick e SANDBERG, Anders. **Whole Brain Emulation: A Roadmap**. Technical Report #2008-3. Oxford: Future of Humanity Institute, Oxford University, 2008.
- BOSTROM, Nick. **Human reproductive cloning from the perspective of the future**, dez/2002.
- BOSTROM, Nick. **Superinteligência: caminhos, perigos, estratégias**. Darkside: Cajamar, 2018.
- COMISSÃO EUROPEIA. **A definition of Artificial Intelligence: main capabilities and scientific disciplines**. Março/2021. Disponível em: <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>. Acesso em: 08/05/2021. Tradução livre.
- CORTE EUROPEIA DE DIREITOS HUMANOS. **Caso Shimovolos v. Russia (30194/09)**. 21 de junho de 2011. Disponível em: <https://hudoc.echr.coe.int/eng#%7B%22appno%22:%5B%2230194/09%22%5D,%22itemid%22:%5B%22001-105217%22%5D%7D>. Acesso em: 14/07/2023.
- MULHOLLAND, Caitlin e FRAJHOF, Isabella. “Entre as leis da robótica e a ética: regulação para o adequado.” *In* BARBOSA, Mafalda Miranda; BRAGA

NETTO, Felipe; SILVA, Michael César; FALEIROS JÚNIOR, José Luiz de Moura. **Direito Digital e Inteligência Artificial**. São Paulo: Editora Foco, 2021. Edição do Kindle.

CYRANOSKI, David. "China set to introduce gene-editing regulation following CRISPR-baby furore". *In: Nature*, 20 de maio de 2019. Disponível em: <https://www.nature.com/articles/d41586-019-01580-1>. Acesso em: 01/04/2021.

DESAI, Deven e KROLL, Joshua. "Trust but verify: a guide to algorithms and the Law". *In: Harvard Journal of Law & Technology*, Vol. 31, n.1, 2017. Disponível em: <https://jolt.law.harvard.edu/assets/articlePDFs/v31/31HarvJLTech1.pdf>. Acesso em: 05/05/2021.

DIAMANDIS, Peter e KOTLER, Steven. **Abundância: o futuro é melhor do que você imagina**. São Paulo: HSM Editora, 2012.

EUROPEAN PARLIAMENT. **Artificial intelligence act**. Disponível em: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS\\_BRI\(2021\)698792\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf). Acesso em: 15/07/2023

FERRY, Luc. **A Revolução Transumanista**. Barueri: Manole, 2018.

GARCIA, Marcos Leite. "A concepção tridimensional dos direitos fundamentais de Gregorio-Peces Barba: reflexões na busca de critérios para o conceito de direitos humanos". *In: CRUZ, Paulo; BRANDÃO, Paulo; OLIVIERO, Maurizio. O Direito contemporâneo e diálogos científicos UNIVALI e Perugia*. Itajaí: UNIVALI, 2016.

GOOD, Irving John. "Speculations Concerning the First Ultraintelligent Machine." *In: Advances in Computers*, edited by Franz L. Alt and Morris Rubinoff, 31–88. Vol. 6. New York: Academic Press, 1965. Disponível em: <https://vtechworks.lib.vt.edu/bitstream/handle/10919/89424/TechReport05-3.pdf>. Acesso em: 02/04/2021.

HORVITZ, Eric. **One-Hundred Year Study of Artificial Intelligence: Reflections and Framing**, White paper, Stanford University, Stanford, CA (ai100.stanford.edu). Disponível em: [https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100\\_framing\\_memo\\_0.pdf](https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100_framing_memo_0.pdf). Acesso em: 03/04/2021.

ISAACSON, Walter. **A decodificadora: Jennifer Doudna, edição de genes e o futuro da espécie humana**. Rio de Janeiro: Editora Intrínseca, 2021.

JOTTERAND, F. "At the roots of transhumanism: from the Enlightenment to a posthuman future." *In: Journal of Medicine and Philosophy*, v. 35, n. 6, 2010.

KURZWEILL, Raymond. **A singularidade está próxima: quando os humanos transcendem a biologia**. São Paulo: Iluminuras, 2018.

MOORE, Max. **The Extropian Principles. A Transhumanist Declaration**.

PONTES DE MIRANDA, Francisco Cavalcanti. **Democracia, liberdade e igualdade**. Os três caminhos. Campinas: Bookseller, 2002.

PORTER, Allen. "Bioethics and Transhumanism". *In: Journal of Medicine and Philosophy*, v. 42, n. 3.

PRISQUE, Emmanuel. **O Sistema de Crédito Social chinês**. Fundação Fernando Henrique Cardoso, 17 de setembro de 2020. Disponível em: [O Sistema de Crédito Social chinês | by Fundação FHC | Fundação FHC | Medium](https://www.fundacaofhc.org.br/pt-br/2020/09/o-sistema-de-credito-social-chines/). Acesso em: 05/05/2021.

ROCO, Mihail e BAINBRIDGE, William. **Converging Technologies for improving human performance – nanotechnology, biotechnology, information technology and cognitive Science**. Kluwer Academic Publishers, 2003.

RUSSELL, S.; DIETTERICH, T.; HORVITZ, E.; SELMAN, B.; ROSSI, F.; HASSABIS, D.; LEGG, S.; SULEYMAN, M.; GEORGE, D.; PHOENIX, S. "Letter to the Editor: Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter." *In: AI Magazine*, [S. l.], v. 36, n. 4, 2015. Disponível em: <https://ojs.aaai.org/index.php/aimagazine/article/view/2621>. Acesso em: 3 apr. 2021.

RUSSELL, Stuart. **Human Compatible**. Nova York: Penguin Publishing Group, 2020.

RUSSELL, Stuart; TEGMARK, Max; DEWEY, Daniel. "Research priorities for robust and beneficial artificial intelligence." *In: AI Magazine*, V. 36, Issue 4, Inverno de 2015, Janeiro de 2015. Disponível em: <https://ojs.aaai.org/index.php/aimagazine/article/view/2577>. Acesso em: 02/04/2021.

SENADO DA REPÚBLICA DO CHILE. **Boletín 13827-19**. Disponível em: [Senado - Tramitación de proyectos](https://www.senado.cl/pt-br/2019/06/05/13827-19-senado-tramitacion-de-proyectos/). Acesso em: 06/05/2021.

SILVA, Cláudio Messias. **Neurotecnologia: Fundamentos Teóricos, Tecnológicos e Mercadológicos**. Dissertação de Mestrado. São Leopoldo: Unisinos, 2016.

SINGH, S., OKUN, A. & JACKSON, A. "Learning to play Go from scratch." *Nature* 550, 336–337 (2017). <https://doi.org/10.1038/550336a>. Acesso em: 01/04/2021.

UNIVERSIDADE DA CALIFÓRNIA EM BERKELEY. **1525 Schools Worldwide That Have Adopted AIMA**. Disponível em: <http://aima.cs.berkeley.edu/adoptions.html>. Acesso em: 01/04/2021.

VILAÇA, Murilo e DIAS, Maria. “Transumanismo e o futuro (pós-) humano.” *In: Physis Revista de Saúde Coletiva*, Rio de Janeiro, 24 [2]: 341-362, 2014.

YUDKOWSKY, Eliezer. “Artificial Intelligence as a Positive and Negative Factor in Global Risk”. *In: BOSTROM, Nick e CIRKOVIC, Milan M. Global Catastrophic Risks*. Nova York: Oxford University Press, 2008. Disponível em: <https://intelligence.org/files/AIPosNegFactor.pdf>. Acesso em: 02/04/2021.

YUDKOWSKY, Eliezer. **Creating Friendly AI 1.0: The Analysis and Design of Benevolent Goal Architectures**. São Francisco: The Singularity Institute, 2015. Disponível em: <https://intelligence.org/files/CFAI.pdf> Acesso em: 08/05/2021.

ZAGREBELSKY, Gustavo. **Imparare democrazia**. Torino: Einaudi, 2007.

ZUBOFF, Shoshana. **A era do capitalismo de vigilância**: a luta por um futuro humano na nova fronteira de poder. Rio de Janeiro: Editora Intrínseca, 2021.