

## O viés de otimismo na era da IA

Tatiana Negry<sup>1</sup>

**Resumo:** Este artigo nasce da inquietação diante do otimismo excessivo em relação ao avanço da inteligência artificial (IA). Em tempos marcados por instabilidades políticas, colapsos ambientais e excesso de informação, o viés de otimismo funciona como um mecanismo de defesa psíquico, uma forma de manter a esperança em meio ao caos. Com base em estudos de psicologia cognitiva, neurociência e ciências sociais, o artigo investiga as raízes desse viés e sua manifestação nas narrativas sobre IA. Além da perspectiva social, também são analisados interesses econômicos e políticos que motivam demasiado otimismo, e acabam por moldar decisões públicas e corporativas em detrimento de medidas de precaução. Mais do que trazer à tona os perigos do otimismo excessivo, o artigo almeja contribuir para a construção de um imaginário coletivo que reconheça, com lucidez, tanto as promessas quanto os limites da IA.

**Palavras-chave:** inteligência artificial; viés de otimismo; risco; precaução.

---

<sup>1</sup> Mestranda no programa de Tecnologias da Inteligência e Design Digital (TIDD) na Pontifícia Universidade Católica de São Paulo (PUC-SP). É graduada em Ciência da Computação pelo Centro Universitário de Brasília (UniCEUB), pós-graduada em Estratégia e Governança de TI pela USP e em Relações Internacionais pela FGV RI. Pesquisadora colaborativa em grupos de pesquisa no que se refere a inteligência artificial, gestão e riscos. Contato: [tnegry@gmail.com](mailto:tnegry@gmail.com). Orcid: <https://orcid.org/0009-0007-4009-6552>.

## **The optimism bias in the age of AI**

**Abstract:** This paper stems from a growing concern regarding the widespread optimism surrounding the development of Artificial Intelligence (AI). In times shaped by political instability, environmental collapse, and information overload, optimism bias acts as a psychological defense mechanism, a way to sustain hope amid chaos. Grounded in research from cognitive psychology, neuroscience, and the social sciences, the article investigates the roots of this bias and how it is reflected in AI related narratives. Beyond the social dimension, it also examines the economic and political interests that encourage overly optimistic views and influence public and corporate decision-making at the expense of precautionary measures. More than exposing the risks of excessive optimism, the article seeks to contribute to the construction of a collective imagination that can recognize, with clarity, both the promises and the limits of AI.

**Keywords:** Artificial intelligence; optimism bias; risk; precaution.

## Introdução<sup>1</sup>

O século XXI é marcado por confluências de crises que abalam profundamente as estruturas sociais, políticas e ambientais. O aumento da instabilidade política, marcado pela perda de confiança nas instituições democráticas, a ascensão de movimentos populistas e a polarização ideológica (Mounk, 2018; Mudde; Kaltwasser, 2017) não apenas fragiliza o discurso democrático, mas também pavimenta o caminho para explicações simplificadas, que encobrem realidades complexas.

Em paralelo, a crise climática global se intensifica a cada ano, desafiando as referências civilizatórias que nortearam a compreensão humana acerca do progresso e riscos a ele associados. Segundo relatório publicado pelo Painel Intergovernamental sobre mudança do Clima (IPCC, 2021), as alterações ambientais observadas desde a década de 1970, como o deslocamento das zonas climáticas em direção aos polos e o aumento da duração das estações, apontam para o aquecimento global contínuo. O relatório adverte que, caso o aquecimento persista, eventos extremos outrora raros se tornarão cada vez mais frequentes e intensos, ultrapassando os registros históricos até então observados.

Não obstante, a humanidade está atravessando uma crise informacional sem precedentes. Embora tenha democratizado o acesso à informação, a era digital também deu origem à disseminação generalizada de desinformação, às câmaras de eco e à fragmentação da esfera pública (Lazer *et al.*, 2018; Papacharissi, 2010). Acontecimentos dessa natureza, perturbam o debate público e comprometem a capacidade coletiva de responder com eficácia a desafios heterogêneos (Habermas, 1989). Juntas, a crise política, climática e informacional têm catalisado uma profunda reavaliação dos valores e prioridades que orientam a ação humana.

No desenrolar desse cenário turbulento, o otimismo tem desempenhado o papel de amortecedor psicológico, promovendo esperança e resiliência mesmo diante das adversidades (Sharot, 2011). Todavia, embora o otimismo motive a perseverança e a inovação, seu excesso acarreta riscos de negligência de ameaças e adiamento de intervenções necessárias. No âmbito das tecnologias emergentes, principalmente a inteligência artificial (IA), a perspectiva otimista ocupa posição de destaque, especialmente no âmbito corporativo. O discurso tecno-otimista que predomina em torno da IA promete benefícios transformadores e avanços sociais (Bostrom,

---

<sup>1</sup> Artigo escrito sob orientação da Profa. Lucia Santaella, orientadora de mestrado da autora.

2014; Yampolskiy, 2016), ao mesmo tempo em que encobre discussões críticas sobre riscos potenciais e dilemas éticos.

Por intervenção do tsunami de crises e expectativas que se avizinha, é necessário compreender os mecanismos psicológicos e culturais que moldam as reações sociais a ameaças emergentes. No que se refere à IA, investigar as raízes do viés de otimismo e seus efeitos sobre a percepção dos riscos tecnológicos, é fator mandatário para que o futuro seja pensado de forma crítica. Este artigo se posiciona na interseção entre psicologia cognitiva e inovação tecnológica, buscando oferecer subsídios teóricos e práticos para uma postura mais reflexiva e proativa sobre os riscos na Era da IA.

### **O viés de otimismo e a percepção dos riscos da IA**

A resistência a previsões pessimistas sobre IA pode ser explicada por mecanismos psicológicos enraizados na cognição humana. Estudos na área de psicologia cognitiva, demonstram que indivíduos tendem a valorizar resultados positivos e minimizar riscos futuros devido a processos heurísticos e vieses sistemáticos no julgamento e na tomada de decisões (Kahneman; Tversky, 1979; Sharot, 2011). Esse tipo de comportamento ocorre em razão de uma tendência mental denominada viés de otimismo, que se caracteriza pela propensão humana a superestimar a probabilidade de viver eventos positivos no futuro e subestimar a probabilidade de viver eventos negativos (Weinstein, 1980).

As raízes do conceito de otimismo podem ser remontadas ao pensamento europeu do século XVII. Descartes, foi um dos primeiros filósofos a expressar a idealização otimista ao manifestar a confiança de que os seres humanos poderiam dominar o próprio universo e assim, aproveitar os frutos da terra e a manutenção da boa saúde (Sharot, 2016). A introdução de otimismo como termo técnico foi cunhada em 1710 pelo filósofo alemão Gottfried Wilhelm Leibniz, que afirmou que a humanidade vive no melhor de todos os tempos, num mundo criado por Deus de forma a ser o mais perfeito possível, apesar das imperfeições aparentes (*ibid.*, 2016).

Assim como ocorre com outras ilusões cognitivas da mente humana, o viés de otimismo tem uma função específica: proteger o cérebro de perceber com precisão a dor e as dificuldades que o futuro reserva e defendê-lo de ver suas opções na vida como um tanto ou quanto limitadas. Como resultado, o estresse e a ansiedade são reduzidos, a saúde física e mental melhora e a motivação para a ação e a produtividade, aumentam

(Sharot, 2016). Entretanto, enquanto o otimismo moderado pode estar vinculado a mais resultados positivos do que negativos, níveis mais extremos de otimismo parecem levar a comportamentos demasiados arriscados, aumentando as chances de fracasso (Puri; Robinson, 2007).

Pesquisas em neurociência cognitiva, corroboram a predisposição otimista dos humanos. Através da análise dos resultados de exames de ressonância magnética funcional (RMF), cientistas identificaram que o viés de otimismo está associado a padrões específicos de ativação neural, mais notadamente no córtex pré-frontal medial e na amígdala, regiões responsáveis pela regulação emocional e pela resposta a ameaças (Sharot *et al.*, 2007). Já a teoria da perspectiva, desenvolvida por Kahneman e Tversky (1979), mostra que indivíduos processam ganhos e perdas de maneira assimétrica, sendo mais propensos a tomar decisões baseadas em expectativas otimistas. Esse tipo de comportamento se manifesta em diversas áreas, que vão desde previsões econômicas e financeiras (Puri; Robinson, 2007) à adoção de inovações tecnológicas (Salazar *et al.*, 2007).

Em se tratando de IA, o viés de otimismo traz um olhar exacerbado aos potenciais benefícios dessa tecnologia, enquanto os riscos, principalmente aqueles de maior gravidade, acabam sendo minimizados ou até mesmo ignorados. Segundo (Bostrom, 2003) a crença generalizada de que as soluções tecnológicas sempre impulsionarão o progresso humano, reduz a preocupação com cenários adversos e fortalece o pressuposto de que problemas futuros serão resolvidos naturalmente pelo progresso computacional.

Em seu livro *A próxima onda*, Mustafa Suleyman e Bhaskar (2023) alertam para os riscos do otimismo excessivo, argumentando que a rápida evolução da IA e de outras tecnologias emergentes poderá gerar instabilidades para as quais a sociedade ainda não está preparada. O autor também critica o que denomina armadilha da aversão ao pessimismo, ou seja, a errônea análise que surge quando o ser humano é sobrecarregado pelo medo de confrontar realidades potencialmente sombrias, e a resultante tendência em olhar para o lado oposto. Ainda na mesma obra, Suleyman e Bhaskar afirmam que a espécie humana não foi programada para lidar com a escala de transformação trazida pela IA, quem dirá com possíveis falhas que tragam prejuízos irreversíveis à sociedade moderna.

O desconforto psicológico gerado por explicações negativas é um fator preponderante para se entender a armadilha da aversão ao pessimismo. De acordo com a teoria da dissonância cognitiva (Festinger, 1957), indivíduos tendem a rejeitar informações que contradizem suas crenças

pré-existentes, visto que a contradição gera um estado interno de tensão e desânimo. Assim, quando confrontados com previsões pessimistas sobre IA, muitos as minimizam ou desconsideram, com o intuito de preservar sua visão de mundo essencialmente otimista.

A negação implicatória, proposta pela socióloga estadunidense Kari Marie Norgaard, certifica a armadilha da aversão ao pessimismo quando propõe que a negação pode ser estruturada socialmente, com vias a diminuir a importância de informações desconfortáveis, sem rejeitá-las abertamente (Norgaard, 2011). Muito empregada no contexto das mudanças climáticas, a negação plausível não contesta a existência do aquecimento global, mas evita incorporar esse conhecimento na vida cotidiana ou na tomada de decisões do dia a dia. A mesma lógica pode ser aplicada à IA, uma vez que o otimismo sobre sua utilização e evolução se sobrepõe à consideração de riscos futuros.

Paul Slovic, psicólogo renomado por seus estudos sobre tomada de decisão e psicologia do risco, argumenta em seu artigo “Perception of Risk” (1987) que a avaliação de riscos não é um processo puramente racional, mas influenciado por fatores emocionais, sociais e cognitivos. Na perspectiva do pesquisador, riscos percebidos como distantes no tempo ou associados a impactos incertos tendem a ser relegados ao segundo plano. O modo como ameaças são entendidas e priorizadas ajuda a explicar por que a perda de controle sobre sistemas superinteligentes recebe menos atenção do que riscos mais imediatos, como ataques cibernéticos ou falhas em sistemas automatizados.

O viés de otimismo e os mecanismos psicológicos ligados à aversão ao pessimismo, desempenham um papel importante na forma como a sociedade percebe os riscos e benefícios das tecnologias emergentes como a IA. A tendência a minimizar os perigos e exagerar os benefícios dessa tecnologia, pode resultar em despreparo para se lidar para possíveis falhas, criando uma falsa sensação de segurança. Ainda que o otimismo moderado seja importante, o debate sobre IA deve incluir uma análise realista de seus riscos, para que as consequências imprevistas possam ser evitadas. O desafio está em equilibrar a inovação com avaliações de possíveis situações negativas e reconhecer que a transição para uma era tecnológica tão avançada, exige preparo tanto psicológico quanto estrutural.

### **Minimização de riscos pela indústria da IA**

Os mecanismos psicológicos que sustentam o viés de otimismo, influenciam diretamente a percepção e a conduta social perante os riscos advindos da IA. Um dos principais efeitos é a minimização de cenários de

risco extremo. Na medida em que tomadores de decisão e pesquisadores superestimam sua capacidade de controle sobre sistemas avançados de IA, a preocupação com ameaças de longo prazo é frequentemente negligenciada (Russell, 2019).

A propensão em minimizar riscos não se limita apenas à comunidade acadêmica ou governamental, mas permeia o discurso da indústria de IA. Geoffrey Hinton, um dos precursores do deep learning, tem sido uma das vozes mais críticas sobre a aceleração do desenvolvimento de sistemas avançados de IA, sem um entendimento pleno de suas consequências. Em 2023, Hinton declarou que os avanços em IA podem representar uma ameaça existencial à humanidade, destacando que muitos pesquisadores e empresas maximizam sua capacidade de controle sobre esses sistemas. Segundo seu ponto de vista, o viés otimista dominante na comunidade científica e no setor tecnológico tem impedido discussões mais realistas sobre os riscos envolvidos, resultando na indiferença a cenários potencialmente catastróficos (Hinton, 2023).

Da mesma maneira, Yoshua Bengio, outro líder na área de aprendizado profundo, defende a necessidade de regulamentação mais rigorosa para evitar a proliferação de sistemas de IA sem os mecanismos de segurança necessários. Bengio enfatiza que pode haver um aumento exponencial dos riscos sistêmicos, incluindo a manipulação de informações em larga escala, a automação descontrolada e a possibilidade de sistemas autônomos tomarem decisões contrárias aos interesses humanos (Bengio, 2023).

No setor corporativo, a minimização dos riscos também se reflete na maneira como as grandes empresas de tecnologia lidam com segurança e ética. Enquanto organizações como OpenAI, Google DeepMind e Anthropic promovem a ideia de desenvolvimento responsável, na prática, a corrida pela liderança em IA acaba por priorizar o desempenho dos modelos de linguagem em detrimento de sua interpretabilidade e segurança (Tegmark, 2020). Modelos cada vez mais sofisticados, como o GPT-4 e o Gemini, são liberados ao público sem a total compreensão de suas limitações, aumentando a possibilidade de efeitos adversos. Embora a probabilidade de uma IA superinteligente se tornar incontrolável pareça ser pequena, o impacto potencial de um evento dessa magnitude pode ser devastador, justificando medidas preventivas rigorosas. Como mencionado por Bostrom (2014), a assimetria entre os custos de precaução e as consequências irreversíveis de riscos existenciais, requer ações de segurança que minimizem qualquer possibilidade de uma catástrofe global.

## Implicações políticas e econômicas do viés de otimismo

Além do fator psicológico, interesses políticos e econômicos são determinantes na construção de uma visão majoritariamente otimista sobre IA. Microsoft, Google, Nvidia e Meta são grandes corporações que investem massivamente em campanhas de marketing e lobbying para promover os benefícios da tecnologia, enfatizando seu impacto positivo na economia e na sociedade (Brynjolfsson; McAfee, 2014). A mídia, por sua vez, complementa essa visão, muitas vezes sem sequer questionar sobre os riscos associados à IA. Franklin Foer (2017), argumenta que os conglomerados tecnológicos exercem forte influência sobre a cobertura jornalística, direcionando a opinião pública para temas que favorecem seus interesses comerciais. Consequentemente, desafios éticos e impactos sociais trazidos pela IA são menos observados do que suas aplicações inovadoras e seu potencial econômico.

Na esfera regulatória, tanto o viés otimista quanto os interesses econômicos dificultam a adoção de medidas preventivas para mitigar riscos. Os formuladores de políticas públicas, afirmam com frequência que a IA trará apenas vantagens, minimizando externalidades negativas e efeitos colaterais (Cave; ÓhÉigeartaigh, 2018). A crença na inovação como um progresso intrinsecamente positivo gera resistência à implementação de regulamentações mais rígidas, resultando em decisões que priorizam o crescimento econômico em prejuízo à segurança e governança responsáveis (Weber, 2006).

A governança internacional da IA é impactada da mesma forma. O estabelecimento de restrições rigorosas pode gerar desvantagens competitivas entre os países que disputam a liderança tecnológica, favorecendo a flexibilidade regulatória. O exemplo da China, que investe massivamente em IA sem impor restrições éticas significativas, demonstra como incentivos econômicos podem suprimir preocupações com riscos sociais e políticos (Lee, 2018).

A convergência entre o viés de otimismo e os incentivos políticos e econômicos contribui para uma percepção limitada dos riscos da IA. Além de reduzir a seriedade do tema, a lógica predominante compromete a implementação de políticas preventivas, expondo a sociedade a potenciais cenários de vulnerabilidade e instabilidade tecnológica (Bostrom, 2014).



## Estratégias para equilibrar otimismo e precaução

Em face aos desafios apresentados, se faz necessária a adoção de estratégias que atenuem os efeitos do otimismo desmedido e promovam mais equilíbrio ao desenvolvimento de sistemas de IA. Dentre as principais medidas, estão a educação pública e a conscientização social sobre riscos tecnológicos. Segundo (Sunstein, 2003; Slovic, 1987) indivíduos bem-informados tendem a apoiar políticas preventivas e a adotar comportamentos mais cautelosos para evitar situações futuras indesejáveis.

Além da educação comunitária sobre IA, também é preciso ampliar o debate acerca dos impactos dessa tecnologia e promover a participação social na elaboração de políticas públicas (Brynjolfsson; McAfee, 2014; Dafoe, 2018). Métodos como votação, pesquisas ou análise de comentários públicos, ajudam a incluir os cidadãos no processo de tomada de decisões e a conter o otimismo irrealista e as visões simplificadas ou excessivamente positivas das capacidades da IA (Ter-Minassian, 2025).

A atuação interdisciplinar no debate sobre governança da IA também é uma das medidas que colabora para o desenvolvimento de uma visão mais abrangente sobre o progresso da tecnologia. Especialistas em ética, filosofia, psicologia cognitiva e ciências sociais devem ser integrados às discussões sobre regulação, para promover uma visão mais abrangente dos impactos algorítmicos. A governança tecnológica eficaz depende da incorporação de múltiplas perspectivas, evitando que a regulamentação seja dominada exclusivamente por interesses corporativos ou reducionismos técnicos (Jasanoff, 2016).

Outra estratégia relevante, é a incorporação do princípio da precaução na formulação de políticas públicas sobre desenvolvimento e utilização de sistemas de IA. Originado em 1971 como parte do programa alemão de proteção ambiental, o princípio da precaução, tornou-se globalmente conhecido em 1992 durante a Conferência das Nações Unidas sobre meio ambiente e desenvolvimento (CNUMAD) ou ECO-92, ocorrida na cidade do Rio de Janeiro (Rio declaration, 1992). Por definição, tal princípio é a garantia contra os riscos potenciais que, de acordo com o estado atual do conhecimento, ainda não possam ser identificados (Martuzzi; Tickner, 2004).

Na ausência da certeza científica formal, a existência de um risco, que represente dano sério ou irreversível, requer a implementação de medidas preventivas que possam reverter potenciais danos (Goldim, 2002). Inicialmente adotado em contextos ambientais e de biossegurança, se

aplicado ao âmbito da IA, o princípio da precaução significa estabelecer regulações preventivas, mesmo que os cenários de risco futuro ainda não sejam totalmente compreendidos.

## Conclusão

Mediante o horizonte de transformações inevitáveis que se avizinha, equilibrar otimismo e responsabilidade não é somente um imperativo técnico, mas civilizatório. Reconhecer que a IA carrega consigo tanto promessas quanto perigos é o primeiro passo para escapar das armadilhas da negação e do otimismo excessivo.

Importa destacar que a ilusão cognitiva, que sustenta o viés otimista, não é apenas individual, mas ganha aspectos coletivos, em especial quando apoiada por narrativas midiáticas, interesses corporativos e discursos tecnocêntricos. Conforme alertado por cientistas da computação, como Bostrom (2014), Russell (2019), Hinton (2023) e Bengio (2023), a negação de riscos incertos, de longo prazo e potencialmente irreversíveis, compromete a capacidade de resposta institucional e adia a definição de formas eficazes de governança para antecipar e mitigar consequências catastróficas.

Em meio a tantas incertezas, urge a necessidade de que governos e sociedades ressignifiquem o imaginário dominante sobre IA. O rompimento com a euforia tecnocrática demanda mais do que reconhecer os riscos, exige a incorporação de pontos de vista interdisciplinares que possam lidar com os efeitos contraditórios das tecnologias emergentes. Não se trata apenas de considerar o princípio da precaução nas políticas públicas, mas de reconstruir os fundamentos regulatórios à luz do conhecimento da psicologia, filosofia, ciências sociais e ética, áreas indispensáveis para orientar decisões alinhadas ao bem comum.

Como destaca Max Tegmark (2020), o mundo está diante de uma bifurcação civilizatória que exige prudência, lucidez e coragem moral. Reconhecer os limites do otimismo e buscar um ponto de equilíbrio entre inovação e responsabilidade é um passo necessário para que o futuro não seja construído apenas de inovação, mas de segurança, justiça, ética e principalmente, de humanidade.<sup>2</sup>

---

<sup>2</sup> O chat GPT 4o foi utilizado para revisão geral de língua portuguesa ao longo do texto e para dispor as referências em ordem alfabética.

## Referências

- BENGIO, Yoshua. AI and catastrophic risk. *Journal of Democracy*, v. 34, n. 3, p. 5–19, set. 2023. Disponível em: <https://www.journalofdemocracy.org/ai-and-catastrophic-risk/>. Acesso em: 15 mar. 2025.
- BOSTROM, Nick. Existential risks: analyzing human extinction scenarios and related hazards. *Journal of Evolution and Technology*, v. 9, n. 1, 2002. Disponível em: < <https://nickbostrom.com/existential/risks.pdf> >. Acesso em: 02 fev. 2025
- BOSTROM, Nick. *Superintelligence: paths, dangers, strategies*. Oxford: Oxford University Press, 2014.
- BRYNJOLFSSON, Erik; McAFFEE, Andrew. *The second machine age: work, progress, and prosperity in a time of brilliant technologies*. New York: Norton, 2014.
- CAVE, Stephen; ÓHÉIGEARTAIGH, Seán. *An AI race for strategic advantage: Rhetoric and risks*. In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2018.
- DAFOE, Allan. *AI governance: a research agenda*. Oxford: Centre for the Governance of AI, Future of Humanity Institute, University of Oxford, 2018.
- FESTINGER, Leon. *A teoria da dissonância cognitiva*. Rio de Janeiro: Zahar, 1975 [1957].
- FOER, Franklin. *World without mind: the existential threat of big tech*. New York: Penguin Press, 2017.
- GOLDIM, José Roberto. Princípio da precaução. *UFRGS – Núcleo Interinstitucional de Bioética*, 2002. Disponível em: <https://www.ufrgs.br/bioetica/precau.htm>. Acesso em: 15 mar. 2025.
- HABERMAS, Jürgen. *The structural transformation of the public sphere: an inquiry into a category of bourgeois society*, transl. Thomas Burger with the assistance of Frederick Lawrence. Cambridge, MA: MIT Press, 1989.
- HINTON, Geoffrey. Geoffrey Hinton on the existential threat of AI | Amanpour and Company. [Entrevista concedida a Christiane Amanpour]. *YouTube*, 2 maio 2023. Disponível em: <https://www.youtube.com/watch?v=Y6Sgp7y178k>. Acesso em: 12 mar. 2025.

IPCC – Intergovernmental Panel on Climate Change. Mudança do clima 2022: Impactos, adaptação e vulnerabilidade: contribuição do grupo de trabalho II ao sexto relatório de avaliação do painel intergovernamental sobre mudanças climáticas. Resumo para formuladores de políticas. Geneva: IPCC, 2022. Disponível em: [https://report.ipcc.ch/ar6wg2/pdf/IPCC\\_AR6\\_WGII\\_SummaryForPolicymakers.pdf](https://report.ipcc.ch/ar6wg2/pdf/IPCC_AR6_WGII_SummaryForPolicymakers.pdf). Acesso em: 3 abr. 2025.

JASANOFF, Sheila. *The ethics of invention: Technology and the human future*. New York: Norton, 2016.

KAHNEMAN, Daniel; TVERSKY, Amos. Prospect theory: An analysis of decision under risk. *Bethesda, MD*, v. 47, n. 2, 1979.

LAZER, David *et al.* The science of fake news. *Science*, New York, v. 359, n. 6380, 2018.

LEE, Kai-Fu. *AI superpowers: China, Silicon Valley, and the new world order*. Boston: Houghton Mifflin Harcourt, 2018.

MARTUZZI, Marco; TICKNER, Joel A. Introduction: The precautionary principle – protecting public health, the environment and the future of our children. Copenhagen: World Health Organization, Regional Office for Europe, 2004, p. 7–15.

MOUNK, Yascha. *O povo contra a democracia*, trad. Cássio de Arantes Leite e Débora Landsberg. São Paulo: Companhia das Letras, 2018.

MUDDE, Cas; KALTWASSER, Cristóbal Rovira. *Populism: A very short Introduction*. Oxford: Oxford University Press, 2017.

NORGAARD, Kari Marie. *Living in denial: climate change, emotions, and everyday life*. Cambridge, MA: MIT Press, 2011.

PAPACHARISSI, Zizi. *A private sphere: democracy in a digital age*. Cambridge: Polity Press, 2010.

PURI, Manju; ROBINSON, David T. Optimism and economic choice. *Journal of Financial Economics*, 86 n. 1, p. 71–99, 2007.

RUSSELL, Stuart. *Human compatible: Artificial intelligence and the problem of control*. New York: Viking, 2019.

SHAROT, Tali; KORN, Christoph W.; DOLAN, Raymond J. How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, New York, v. 14, p. 1475–1479, 2011.

SHAROT, Tali. *O viés otimista: porque somos programados para ver o mundo pelo lado positivo*, trad. Ana Beatriz Rodrigues. Rio de Janeiro: Rocco, 2016.

SHAROT, Tali. *The optimism bias: a tour of the irrationally positive brain*. New York: Pantheon Books, 2011.

SHAROT, Tali; RICCARDI, Alison M.; RAIIO, Candace M.; PHELPS, Elizabeth A. Neural mechanisms mediating optimism bias. *Nature*, v. 450, n. 7166, p. 102–105, 2007.

SLOVIC, Paul. Perception of risk. *Science*, v. 236, p.280–285, 1987.

SULEYMAN, Mustafa; BHASKAR, Michael. *The coming wave: Technology, power, and the twenty-first century's greatest dilemma*. New York: Crown, 2023.

SUNSTEIN, Cass R. Beyond the precautionary principle. *University of Pennsylvania Law Review*, v. 151, n. 3, p. 1003–1058, 2003. Disponível em: [https://scholarship.law.upenn.edu/penn\\_law\\_review/vol151/iss3/10/](https://scholarship.law.upenn.edu/penn_law_review/vol151/iss3/10/). Acesso em: 20 mar 2025.

TEGMARK, Max. *Vida 3.0: O ser humano na era da inteligência artificial*, trad. Petê Rissatti. São Paulo: Benvirá, 2018.

TER-MINASSIAN, Lucile. Democratizing AI governance: Balancing expertise and public participation. *arXiv*, 2025. Disponível em: <https://arxiv.org/abs/2502.08651v1>. Acesso em: 3 abr. 2025.

TORRES-SALAZAR, Cristina.; MORETA-HERRERA, Rodrigo; RAMOS-RAMÍREZ, Martha; LÓPEZ-CASTRO, Javier. Sesgo cognitivo de optimismo y percepción de bienestar en una muestra de universitarios ecuatorianos. *Revista Colombiana de Psicología*, v. 29, n. 1, p. 61–72, 2020. Disponível em: <https://revistas.unal.edu.co/index.php/psicologia/article/view/75853>. Acesso em: 28 mar. 2025.

UNITED NATIONS. Rio Declaration on environment and development. United Nations Environment Programme, 1992. Disponível em: <https://web.archive.org/web/20090518143206/http://www.unep.org/Documents.Multilingual/Default.Print.asp?DocumentID=78&ArticleID=1163&l=en>. Acesso em: 10 Mar. 2025.

WEBER, Max. *A ética protestante e o “espírito” do capitalismo*, trad. José Marcos Mariani de Macedo. São Paulo: Companhia das Letras, 2006.

WEINSTEIN, Neil D. Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, v. 39, n. 5, p. 806–820, 1980.

YAMPOLSKIY, Roman V. *Artificial superintelligence: a futuristic approach*. Boca Raton: CRC Press, 2016.