

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS
ESCOLA DE COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E CULTURA

Paula Cardoso Pereira

Futuros Maquímicos:
racionalidade e temporalidade nos algoritmos da Inteligência Artificial

Rio de Janeiro, RJ
Agosto de 2024

PAULA CARDOSO PEREIRA

Futuros Maquínicos:
racionalidade e temporalidade nos algoritmos da Inteligência Artificial

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Comunicação e Cultura (Tecnologias da Comunicação e Estética), Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Doutora em Comunicação e Cultura.

Orientadora: Prof^ª. Dra. Fernanda Glória Bruno

Rio de Janeiro, RJ

Agosto de 2024

PAULA CARDOSO PEREIRA

Futuros Maquínicos:

racionalidade e temporalidade nos algoritmos da Inteligência Artificial

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Comunicação e Cultura (Tecnologias da Comunicação e Estética), Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Doutora em Comunicação e Cultura.

Fernanda Glória Bruno, Universidade Federal do Rio de Janeiro

Flavia Costa, Universidad de Buenos Aires

Claudio Celis Bueno, University of Amsterdam

Icaro Ferraz Vidal Junior, MediaLab.UFRJ

Giselle Beiguelman, Universidade de São Paulo

CIP - Catalogação na Publicação

C436f Cardoso Pereira, Paula
 Futuros Maquímicos: racionalidade e
 temporalidade nos algoritmos da Inteligência
 Artificial / Paula Cardoso Pereira. -- Rio de
 Janeiro, 2024.
 244 f.

 Orientadora: Fernanda Glória Bruno.
 Tese (doutorado) - Universidade Federal do Rio
 de Janeiro, Escola da Comunicação, Programa de Pós
 Graduação em Comunicação, 2024.

 1. algoritmos. 2. Inteligência Artificial. 3.
 racionalidade. 4. temporalidade. 5. futuro. I.
 Glória Bruno, Fernanda , orient. II. Título.

Às mulheres pesquisadoras que perturbam fronteiras

A nossa feroz insistência em outros possíveis

*O tempo não é um panorama geral,
mas antes o resultado provisório da ligação entre os seres.*

Bruno Latour

*Na realidade, ficar com o problema requer aprender a estar
verdadeiramente presente; não como um eixo que se desvanece
entre passados terríveis ou edênicos e entre futuros apocalípticos
ou salvadores - mas como bichos mortais entrelaçados em uma
miríade de configurações inacabadas de lugares, tempos, matérias,
significados.*

Donna Haraway

Agradecimentos

Às pesquisadoras e pesquisadores do MediaLab.UFRJ, pelas trocas generosas, enriquecedoras e alegres, pelo prazer de pensar e fazer mundos juntas. A Fernanda Bruno, pela oportunidade de fazer deste lugar em que desenvolvi verdadeira paixão pela pesquisa sobre tecnologia e pude participar de projetos que foram fundamentais para esta tese, pela parceria nos pensamentos tentaculares, pela orientação sempre precisa e preciosa, pelos incentivos, pela amizade. A Paulo Faltay e Anna Bentes, pelo companheirismo acadêmico e afetivo, pelas trocas, *insights* e sugestões generosas. A Debora Pio, Mariana Antoun, Helena Strecker, Manuella Caputo, Alice Lerner, Laís Sebben e Juliana Monteiro e a todas companheiras e companheiros de pesquisa deste período. Vida longa ao “lab” e às individualizações coletivas que ele produz!

A Flavia Costa, Claudio Celis Bueno, Icaro Ferraz Vidal Junior, Giselle Beiguelman e Paulo Faltay, por aceitarem generosamente compor a banca. A Claudio e Icaro, também, pelas valiosas contribuições na qualificação.

Ao CNPq, pelo financiamento da bolsa de pesquisa, e ao Programa de Pós-Graduação em Comunicação da UFRJ, por ser um espaço tão acolhedor e instigante ao pensamento.

A minhas irmãs e especialmente minha mãe, mulheres da minha vida, pelo apoio, carinho, incentivo e paciência irrestritos, por serem o testemunho de que não há trabalho intelectual sem trabalho de cuidado e que esse segue sendo um trabalho majoritariamente feminino. Ao meu pai, pelo apoio, torcida constante e acolhimento afetivo nos momentos de incerteza.

A Paulo Germano, Tatiana Kito e Hélder Vieira, por terem se tornado família estendida, pela presença duradoura e alegre em minha vida. A Julie Maciel, Mariana Bonadio, Daniel Ranzi Werle, Joaquín Zerené, pela amizade fortalecedora e fiel.

Resumo

PEREIRA, Paula C. **Futuros Maquínicos**: racionalidade e temporalidade nos algoritmos da Inteligência Artificial. Tese (Doutorado em Comunicação e Cultura) — Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2024.

Esta tese investiga questões de ordem *epistemológica* e, sobretudo, *cronopolítica* nos algoritmos preditivos da Inteligência Artificial, ou seja, aqueles que utilizam métodos de aprendizado de máquina. A hipótese central que nos orienta é que, em virtude da centralidade do conhecimento e intervenção sobre a dimensão temporal para a racionalidade algorítmica, sob a forma da predição de probabilidades, os algoritmos da IA podem ser considerados *tecnologias de governo do tempo*. Argumentamos que seus efeitos mais poderosos, marcadamente performativos, residem no modo com que vêm, progressivamente, modulando o curso das ações e eventos possíveis, ou simplesmente, o que chamamos de *campo do possível*. Seja na regulação personalizada do espectro visível e legível nas plataformas digitais, na automatização de interpretações e decisões nas mais variadas esferas ou nos mundos sintéticos e cada vez mais abundantes fabricados pelos modelos generativos, vem sendo gestada uma *algoritmização do tempo* que tem como uma de suas principais características uma *redução das possibilidades às probabilidades*. Além desta dimensão, nos debruçamos sobre a noção de *futuros maquínicos*, situando essa reflexão num panorama mais amplo sobre futuros históricos na modernidade e na contemporaneidade. Propomos tal noção para designar a *modalidade de futuro* produzido por meio das técnicas, práticas, discursos, promessas, imaginários e ideologias sobre a Inteligência Artificial na contemporaneidade. Deste modo, uma das perguntas que atravessa esta tese se desdobra num duplo movimento complementar: por um lado, o questionamento por que forças (sociais, epistêmicas, materiais, históricas, etc) dão forma à *algoritmização do tempo e do futuro* e, recursivamente, quais os efeitos e implicações de dita algoritmização. A tese está estruturada em dois eixos: um *epistemológico* e um *cronopolítico*. O primeiro está centrado sobre o modelo de racionalidade e cognição algorítmicos e sobre o regime de saber que sustenta tal modelo. O segundo, por sua vez, desdobra-se em duas vias: num aprofundamento sobre os *regimes de temporalidade algorítmicos*, incluindo dois estudos de caso, um sobre o TikTok, outro sobre o ChatGPT, e em reflexões sobre os *futuros maquínicos*.

Palavras-chave: Inteligência Artificial; racionalidade algorítmica; temporalidades algorítmicas; futuros maquínicos.

Abstract

PEREIRA, Paula C. **Machinic Futures: rationality and temporality in Artificial Intelligence algorithms.** Thesis (Doctorate in Communication and Culture) – Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2024.

This thesis investigates *epistemological* and, above all, *chronopolitical* issues in predictive Artificial Intelligence algorithms, specifically those utilizing machine learning methods. The central hypothesis guiding this research is that, due to the centrality of knowledge and intervention in the temporal dimension for algorithmic rationality — manifested through probability prediction — AI algorithms can be understood as *time-governing technologies*. We argue that their most powerful effects, markedly performative in nature, lie in how they are progressively modulating the course of possible actions and events or, simply put, what we call *the field of the possible*. Whether through the personalized regulation of the visible and legible spectrum on digital platforms, the automation of interpretations and decisions in various spheres, or the increasingly abundant synthetic worlds created by generative models, an *algorithmization of time* is emerging, characterized primarily by the *reduction of possibilities to probabilities*. Beyond this dimension, we delve into the notion of machinic futures, situating this reflection within a broader panorama of historical futures in modernity and contemporaneity. We propose this notion to designate the *modality of future* produced through the techniques, practices, discourses, promises, imaginaries, and ideologies surrounding Artificial Intelligence in the contemporary era. Thus, one of the central questions of this thesis unfolds in a complementary dual movement: on one hand, examining the forces (social, epistemic, material, historical, etc.) shaping *the algorithmization of time and the future*; and, on the other hand, exploring the recursive effects and implications of this algorithmization. The thesis is structured around two axes: an *epistemological* and a *chronopolitical* one. The first focuses on the model of algorithmic rationality and cognition, as well as the knowledge regime underpinning this model. The second unfolds in two directions: a deep dive into algorithmic temporality regimes, including two case studies — one on TikTok and the other on ChatGPT — and reflections on *machinic futures*.

Keywords: Artificial Intelligence; algorithmic rationality; algorithmic temporalities; machinic futures.

Lista de figuras

Figura 1: Rosenblatt manipulando o Perceptron (esquerda) e parte da máquina conexionista (direita)	50
Figura 2: Arquitetura de camadas de uma rede neural	51
Figura 3: Gráfico da correlação espúria entre diplomas concedidos em artes liberais x pesquisas no Google por 'dor de barriga'	69
Figura 4: Ranking de plataformas sociais por tempo de uso referentes ao período entre junho e setembro de 2023 (horas/mês)	111
Figura 5: Ranking de plataformas sociais por número de usuários ativos em abril de 2024.	112
Figura 6: Aba Para Você.	113
Figura 7: Tela de <i>onboarding</i> do TikTok.	120
Figura 8: Tuíte em que usuária relata o nível de intimidade e confiança que desenvolveu com o ChatGPT.	136
Figura 9: Timeline com os LLM's lançados desde 2018 e seus respectivos tamanhos. As cores agrupam as empresas produtoras dos modelos.	139
Figura 10: Mapa de calor global que mede a representação linguística em 1.800 datasets de texto para treinamento de IA.	147
Figura 11: Tabela de probabilidades da próxima palavra para a frase “ <i>The best thing about AI is its ability to</i> ”.	149
Figura 12: Mapa de incorporações de palavras produzido a partir da palavra “genealogy” gerado com a ferramenta Embedding Projector.	152
Figura 13: Tabela do livro <i>The Precipice</i> com estimativas sobre ‘catástrofes existenciais’ para os próximos 100 anos.	204

Sumário

Prólogo: do protagonismo do algoritmo ao ‘hype da IA’	12
Introdução	16
EIXO EPISTEMOLÓGICO	
Capítulo 1. Racionalidade e cognição algorítmica	28
1.1 Racionalidade algorítmica e o governo do possível	28
1.2 Racionalidade algorítmica e a crise da razão reflexiva	34
1.3 Algoritmos enquanto cognoscentes e o não-consciente cognitivo	37
1.4 Inteligência Artificial: definições e paradigmas	43
1.4.1 <i>O que é e como raciocina uma IA?</i>	40
1.4.2 <i>Simbólicos e conexionistas: duas abordagens da IA</i>	47
Capítulo 2. Regime de saber algorítmico: três deslocamentos na ‘ordem das coisas’	55
2.1 Performatividade x representação	56
2.2 Correlação x causalidade	63
2.3 Provável x possível	74
EIXO CRONOPOLÍTICO	
Capítulo 3: Regimes de temporalidade algorítmicos	87
3.2 Matrizes da cronopolítica algorítmica	89
3.2.1 <i>A matriz cibernética: temporalidade recursiva e causalidade circular</i>	89
3.2.2 <i>A matriz militar-securitária: temporalidade antecipatória</i>	97
3.2.3 <i>A matriz financeira: temporalidade acelerada e os mercados de futuros</i>	102
3.3 Temporalidades algorítmicas: dois estudos de caso	109
3.3.1 <i>Predição, aceleração, fragmentação: TikTok como ambiente de alta frequência</i>	109
3.3.2 <i>ChatGPT como megamáquina probabilística</i>	133
Capítulo 4. Futuros maquínicos	160
4.1 História(s) do futuro: do futuro utópico ao futuro maquínico	164
4.2 O futuro como repetição recursiva e as margens de indeterminação das máquinas computacionais	179
4.3 Antropoceno e Tecnoceno: a dupla aceleração da nossa época	187
4.4 Habitando o cronotopo das ameaças e promessas finais: mitologias do fim	195
Considerações finais	206
Referências	219

Prólogo: do protagonismo do algoritmo ao ‘hype da IA’

No período de desenvolvimento desta tese, entre os anos de 2019 e 2024, mudanças radicais marcaram as culturas e práticas algorítmicas, seja no âmbito da pesquisa e desenvolvimento tecnológico, na academia ou no debate público. Naquele momento inicial, o termo *algoritmo* começava a se popularizar, migrando de contextos estritamente técnicos e debates intelectuais nichados para um público mais amplo e não especializado. Casos como o escândalo da *Cambridge Analytica*¹ e os impactos da crescente personalização algorítmica de conteúdo em plataformas como Facebook, Netflix, YouTube, Instagram e Spotify, como observa Faltay (2020), forjavam a passagem da prevalência de termos que remetiam a processos (*big data*, *datamining*, *data analytics*, *profiling*, etc) para um entendimento do algoritmo enquanto “ente único (...) e dotado de ação própria e diretiva” (ibid., p. 9), seguidamente associado a um “sentimento persecutório difuso” (ibid., p. 10).

No plano teórico, trabalhos como *The Black Box Society* (Pasquale, 2015), *Weapons of math destruction* (O’Neil, [2016] 2021), *Algorithms of Opression* (Noble, [2018] 2021) denunciavam o caráter simultaneamente *opaco* e *opressor* da mediação algorítmica e seus efeitos na automatização e intensificação de desigualdades históricas, atualizando práticas racistas, sexistas, classistas, eugenistas e excluindo minorias do acesso a oportunidades progressivamente reguladas por algoritmos (ver também Eubanks, 2019; Silva, 2020; Chun, 2021). Além disso, o contundente trabalho de Zuboff ([2019] 2021) sobre a consolidação de um *Capitalismo de Vigilância* nomeou e deu contorno aos pressupostos, instrumentos e objetivos de um complexo entramado que envolve um novo modelo econômico, de conhecimento, de controle comportamental e de poder.

Assim, progressivamente ao longo destes anos, os algoritmos foram assumindo “protagonismo nas narrativas tecnológicas, culturais e políticas contemporâneas” (Faltay, op. cit., p. 9). Segundo Mbembe (2019b), já não vivemos na era da máquina, mas na *era do algoritmo* ou, nos termos de Berardi (2019), passamos, entre a modernidade e a contemporaneidade, da era da *máquina externa* para a da *máquina interna*².

¹ Revelado em março de 2018, o conhecido escândalo da *Cambridge Analytica* expôs o uso não consentido de dados comportamentais de milhões de usuários do Facebook para influenciar campanhas políticas, como a eleição presidencial dos EUA em 2016 e o referendo do Brexit. O caso gerou grande repercussão global, levantando questões sobre privacidade, manipulação comportamental e regulamentação das plataformas digitais.

² Como veremos no Capítulo 4, Berardi propõe que a passagem do séc. XX ao XXI implicou também uma mudança no tipo de máquinas predominantes em cada momento histórico. Enquanto o “século que acreditou no futuro” (p. 13) se relaciona à Máquina Externa — a máquina pesada, ferruginosa e volumosa, externa em

A tecnologia que estava na base de todos esses processos já era isso que hoje se popularizou chamar de Inteligência Artificial, mas que corresponde, tecnicamente, à subárea deste campo concentrada no uso de técnicas de aprendizado de máquina (*machine learning*, em inglês), especialmente o aprendizado profundo (*deep learning*, em inglês). Conforme veremos em detalhes ao longo desta tese, a operação cognitiva fundamental desta vertente, também chamada de IA conexionista, consiste na elaboração de *previsões* a partir do *reconhecimento indutivo de padrões* em vastos conjuntos de dados. Daí a denominação *algoritmos preditivos da IA* (ou apenas *algoritmos preditivos*) que utilizamos em vários momentos do texto. Aqui já destacamos um ponto fundamental para as questões-chave desta tese: *a centralidade da previsão (e da dimensão temporal de forma mais ampla) para o modelo de racionalidade algorítmico*.

Como a própria trajetória da pesquisa vinha acompanhando, essa classe de algoritmos já integrava intensa e extensivamente as infraestruturas dos ecossistemas digitais há anos, invisivelmente incorporada a sistemas de uso cotidiano (em ferramentas de busca e tradução, reconhecimento de imagens, reconhecimento facial, assistentes virtuais, na recomendação algorítmica personalizada de conteúdo e publicidade em plataformas como o Facebook, Spotify, Netflix e Instagram) ou na automatização de decisões ligada a um novo regime de antecipação de eventos e gerenciamento de riscos (Benbouzid, Cardon, 2018) em áreas como a segurança, a educação, a guerra, a justiça, o trabalho, a saúde, as finanças, etc. Contudo, foi somente após o avassalador lançamento do ChatGPT, em novembro de 2022, e da subsequente explosão dos modelos generativos³, que o termo Inteligência Artificial invadiu o léxico e o imaginário das redes e da arena pública. Corrêa (2023) chamou-o de “evento de percepção”, um “acontecimento que altera a sensibilidade coletiva com relação à tecnologia”. Transformado numa espécie de ‘marco de entrada oficial na *era da Inteligência Artificial*’, o evento provocou um abalo sísmico nos imaginários e nas expectativas mercadológicas sobre a tecnologia, despertando também com força os antigos “fantasmas de superação, destruição e salvação do humano” (Bruno, Pereira, Faltay, 2023, p. 237) que sempre assombraram a história da IA.

relação ao corpo e à mente, sintetizada sobretudo na vanguarda futurista e na projeção progressiva do futuro —, nossa época seria marcada pela Máquina Interna — a máquina bioinformática, miniaturizada, que se entrelaça com o sistema nervoso social.

³ É interessante pontuar como o *boom* dos modelos generativos intensifica uma tendência que já vinha ocorrendo no desenvolvimento da IA de que subáreas cada vez mais específicas passem a dominar todo o mercado, tornando-se quase sinônimos de todo o campo. Deste modo, poderíamos traçar uma ‘evolução’ gradual e cada vez mais especializada de IA - Machine Learning - Deep Learning - IA Generativa.

Grandes modelos de linguagem como o GPT conferem à IA o status de *tecnologia de propósito geral*⁴, ou *metatecnologia* — uma tecnologia que pode ser utilizada como base para o desenvolvimento de diversas outras, tornando seu potencial de transformar transversalmente diversos setores da sociedade muito maior, como foi o caso de tecnologias como a máquina a vapor, a eletricidade e o computador (Kanashiro, 2023). Os alardes sobre uma transformação acelerada de *‘tudo em todo lugar ao mesmo tempo’* que marcaram o chamado *‘hype da IA’* alimentaram toda sorte de *IA-solucionismo*, corrida desenfreada por novos modelos e lucros sem precedentes do setor que passou a dominar os investimentos do mercado de tecnologia. Esse movimento provocou a entrada ao cenário de novas empresas disputando com as conhecidas gigantes da GAFAM, como a OpenAI e a Anthropic, mas também acelerou a necessária elaboração de marcos regulatórios em todo o mundo⁵.

É importante pontuar também como a pandemia de Covid-19, deflagrada em 2020, foi um ponto de inflexão crucial para as transformações que estamos descrevendo. Evento histórico, tão traumático quanto mal elaborado coletivamente quanto às suas causas, sentidos e implicações, a pandemia, paradoxalmente, pôs tudo em suspenso e acelerou vários processos que já estavam em curso (Bruno, Pereira, Faltay, 2023), incluindo a intensa digitalização e algoritmização da vida que criou as condições materiais para que o *boom* da IA emergisse.

Assim, mais do que um tema em si visado desde o princípio, a aceleração sem precedentes que marcou os últimos anos do desenvolvimento da IA, consolidando a hegemonia de um paradigma *preditivo-aceleracionista* pautado na produção de previsões estatísticas via aprendizado de máquina movido por uma otimização ininterrupta (Bruno, Bentes, Faltay, 2023) e na premissa de que quanto maiores os modelos, melhores e mais eficientes são (Varoquaux, Luccioni, Whittaker, 2024) — atravessou e modificou a trajetória desta pesquisa.

⁴ Esse traço, no entanto, não deve ser confundido com a chamada *Inteligência Artificial Geral*, o conceito de uma IA que seria capaz de realizar qualquer tarefa intelectual que um ser humano poderia fazer. Apesar de que esse objetivo segue guiando esforços, expectativas e sobretudo as promessas e temores mobilizados em torno da IA, em íntima relação com os imaginários sobre a Singularidade Tecnológica (Vinge, 1993; Kurzweil, 2005), mesmo os grandes modelos de linguagem como o GPT ainda estão longe de atingir esse estágio. Na prática, atualmente os modelos de IA ainda estão atrelados à realização de tarefas específicas ou resolução de problemas limitados, correspondendo à chamada *IA Estreita* ou *Fraca*.

⁵ Entre estes, o principal foi o *AI Act*, aprovado em março de 2024 pelo Parlamento Europeu e cuja maior inovação foi a classificação de sistemas de IA segundo grau de risco. Nos Estados Unidos, o presidente Joe Biden assinou uma ordem executiva em novembro de 2023 para a regulamentação da IA no país. No Brasil, neste momento, tramita o Projeto de Lei nº 2.338/23, conhecido como Marco Legal da IA no Brasil e que teve grande influência do AI Act da União Europeia.

Por um lado, mobilizando novos objetos e problemas, como o próprio ChatGPT e o TikTok — que, dado a relevância que adquiriram e o modo como materializam questões fundamentais para nossa investigação, tornaram-se estudos de caso — ou as relações imbricadas entre os imaginários apocalípticos suscitados pelo Antropoceno e pelo Tecnoceno (Costa, 2021) num contexto histórico de *crise do futuro* (Fisher, 2019; Berardi, 2019; Hartog, 2021). Por outro, tal aceleração consolidou ainda mais algumas de nossas intuições e hipóteses iniciais, como o entendimento dos algoritmos da IA enquanto *tecnologias de governo do tempo* e seus efeitos de redução das possibilidades às probabilidades, que descrevemos em mais detalhes na Introdução.

A leitora identificará ao longo do texto marcas desses distintos momentos das culturas algorítmicas que descrevemos brevemente acima, manifestações de uma certa anacronia inerente a um processo de pesquisa e escrita que abarcou, justamente, um período de mudanças tão profundas, condensadas e velozes para tais temas e questões.

Introdução

Ainda que as práticas e técnicas relacionados à *predição*, assim como seus sentidos culturais, sejam historicamente situados, há algo que parece sobreviver ao longo dessa história. É a presença de um olhar, identificado por Jean-Pierre Vernant em seus estudos sobre a psicologia da adivinhação na Grécia antiga, capaz de acessar as “relações íntimas e secretas das coisas, as correspondências e as analogias”⁶ (Didi-Huberman, 2010, p. 37). Segundo Vernant, esse olhar operaria uma transmutação do estatuto do objeto contemplado, que passaria de *coisa visível* para ser suporte de *coisas a adivinhar ou prever*. Delimitado enquanto *templum*, este objeto se tornava então capaz de revelar os *signos da predição*.

Se, como descreve o autor, na Antiguidade, esse olhar adivinhatório repousava na observação dos astros, das entranhas dos animais ou dos movimentos das aves, hoje ele parece repousar na observação de dados massivos através do *olho do algoritmo* (Pasquinelli, 2015). Fundado no reconhecimento de padrões inacessíveis para “nossa capacidade cognitiva desarmada” (Oliveira, 2023), esse olho maquínico opera uma transmutação da vastidão de dados incessantemente produzida, capturada e analisada nos ecossistemas digitais, em economicamente valiosos *signos da predição*. A convergência do *big data* com os algoritmos preditivos vem produzindo espécies de *oráculos digitais* para as indústrias do mercado, da ciência, da guerra, e da política, cada vez mais entrelaçadas. Nesse *novo regime preditivo*, a realidade e o próprio sujeito são datificados, interpretados e classificados segundo padrões e correlações ocultos⁷ a serem decifrados por este poderoso *cognoscente técnico*⁸ (Hayles, 2017) que agora ocupa lugar privilegiado numa série de processos em curso.

⁶ No original: “relaciones íntimas y secretas de las cosas, las correspondencias y las analogías”.

⁷ Que a realidade se converta num conjunto de padrões não é algo especialmente novo. O que nos parece novo em nossa atualidade é que o algoritmo se converta no agente privilegiado que reconhece, interpreta e classifica tais padrões.

⁸ Conforme veremos no tópico 1.3, no entendimento de Hayles cognoscentes são atores, biológicos ou técnicos, capazes de interpretação e escolha.

No contexto de um cada vez mais extensivo e intensivo *Capitalismo de Dados*⁹ (West, 2017) assentado sobre um *novo extrativismo* que converte o mundo num grande “empório de dados” (Mbembe, 2019b) e dominado pelo modelo de negócios das grandes plataformas digitais (Cf. Zuboff, 2021; Sadowski, 2019; Srnicek, 2017), os saberes e poderes algorítmicos ocupam um lugar cada vez mais central sobre os modos de conhecer e agir sobre a realidade, os fenômenos e os sujeitos. Sob a forma do aprendizado de máquina, o algoritmo se tornou fundamental para o capital, fazendo surgir novas formas de produção de valor e desencadeando uma modificação dos diagramas do poder (Cf. Pasquinelli, 2009; Morozov, 2018; Wark, 2019; Zuboff, op. cit., Markelj, Bueno, 2023).

Além disso, o algoritmo é um agente essencial da radical reorganização do visível e da verdade pela qual passamos, produzindo principalmente uma intensa *personalização de mundos* (Cesarino, 2022) nos ecossistemas digitais através de uma série de filtragens, recomendações e classificações em ambientes de alta entropia informacional, velocidade e disputa acirrada pela atenção dos usuários. De forma ampla, esses agentes maquínicos cada vez mais poderosos protagonizam um processo histórico de crescente *delegação e automação cognitivas*. A centralidade que os algoritmos adquirem contemporaneamente nos regimes de conhecimento e governo da realidade, dos fenômenos e dos sujeitos, influenciando profundamente processos epistemológicos, sociológicos, econômicos, políticos, estéticos e subjetivos, pode ser descrita nos termos de um novo modelo de *racionalidade* (Bruno, op. cit.) ou *governamentalidade algorítmica* (Rouvroy, Berns, 2015), como veremos no Capítulo 1.

Porém, apesar das crescentes implicações dos algoritmos no nosso atual ‘estado das coisas’, as visões que propagam um entendimento do algoritmo como ente único (Faltay, 2020) e desmaterializado dificultam reconhecer como os processos e sistemas algorítmicos, sobretudo sob a forma da *IA em grande escala* que dominou o mercado nos últimos anos, somente são viáveis através de uma exploração também em larga escala de dados, de recursos naturais e energéticos, de trabalho humano e de infraestruturas e logísticas que operam a nível planetário e cujo monopólio está mão das corporações que conseguem dominar essa cadeia

⁹ West (2017) utiliza o termo capitalismo de dados para descrever um sistema baseado na comoditização de nossos dados e que introduz uma distribuição assimétrica de poder, consolidando atores que com acesso privilegiado e capacidade de dar sentido a tais informações. A noção se aproxima da de *Capitalismo de Vigilância* de Zuboff (2021), sobre a qual discorreremos nas próximas páginas, ainda que esta descreva não apenas uma nova lógica acumulativa baseada em dados, mas também um modelo mais amplo de poder que visa prever e modificar comportamentos no contexto digital.

produtiva (Cf. Crawford, 2021). Assim, a consolidação e o exercício de uma racionalidade ou governamentalidade algorítmica implica a materialidade e articulação de uma complexa rede sociotécnica heterogênea, multiescalar e distribuída formada, além de algoritmos, por corporações, centros de pesquisa, cientistas, *crowdworkers*, infraestruturas técnicas, bancos de dados de treinamento, fontes de minérios e de água, etc.

Vários autores vêm apontando como a escalada tecnológica em curso corresponde a uma atualização das lógicas extrativistas coloniais. Enquanto o colonialismo tradicional estava assentado sobre a conquista territorial e a exploração de recursos naturais e de populações subalternizadas, o que estes autores vêm chamando de *colonialismo de dados* (Couldry, Mejías, 2019; Amadeu, Souza, Cassino, 2021), *colonialismo algorítmico* (Birhane, 2020) ou *colonialismo digital* (Faustino, Lippold, 2023) se baseia em operações extrativas que convertem os fluxos da vida em dados e os expropriam, assim como no controle privado das infraestruturas de processamento e comercialização desses dados por parte de grandes corporações de tecnologia do Norte Global. Ricaurte (2019) enfatiza como a epistemologia e racionalidade *data-driven* que subjaz a esse modelo econômico implica um arranjo de práticas, materialidades, territórios, corpos e subjetividades que impõem violentamente modos de ser, pensar e sentir que negam a existência de mundos e epistemologias alternativas. Ainda que as técnicas e táticas de exploração desse novo paradigma colonialista sejam aparentemente muito mais abstratas, reticulares e menos violentas que as do anterior, seu impulso predatório é tributário da colonialidade do poder (Cf. Quijano, 2007).

O extrativismo que alimenta o voraz modelo *data-driven* prototipado no Vale do Silício baseia-se, como se sabe, no monitoramento e captura incessante de nossas atividades nas redes. Esse modelo, simultaneamente econômico, epistemológico e político, visa tornar o máximo de aspectos da vida datificáveis, quantificáveis, classificáveis, previsíveis e, claro, monetizáveis. A mineração de padrões e correlações nesses gigantescos conjuntos de dados por meio de saberes estatísticos (que são o núcleo técnico do aprendizado de máquina) permite inferir *quais os futuros prováveis*: seja a probabilidade de conteúdos atraírem a atenção dos usuários ou viralizarem nas redes sociais, a próxima palavra de uma frase, a propensão de indivíduos quitarem suas dívidas, serem contratados, ficarem doentes, cometerem crimes ou de seus traços psíquicos e emocionais.

Aplicado ao comportamento humano, essa pretensiosa ‘maquinaria adivinhatória’ que visa nos conhecer melhor do que nós mesmos (Bruno, Bentes, Faltay, 2019) invade e expropria

aquilo que antes pertencia à esfera da vida psíquica privada — nossos hábitos, personalidades, emoções e, sobretudo, nossas propensões, vulnerabilidades e supostos desejos. Não necessariamente aquilo que *somos* hoje, mas principalmente *o que podemos vir a ser* no futuro. Esse verdadeiro extrativismo psíquico de nossos dados alimenta aquilo que Bruno, Bentes e Faltay (2019) chamam de uma *Economia Psíquica dos Algoritmos*.

A centralidade dos modelos preditivos de conhecimento e intervenção sobre os comportamentos na economia de dados é enfatizada por Zuboff (2021) ao descrever as engrenagens do que designou *Capitalismo de Vigilância*: “uma nova ordem econômica que reivindica a experiência humana como matéria-prima gratuita para práticas comerciais dissimuladas de extração, previsão e venda” (ibid., p. 2). A operação fundamental dessa nova fase do capitalismo, cujo precursor foi o Google e seu modelo de publicidade personalizada no começo dos anos 2000, reside na produção de valor a partir de *dados comportamentais*. Segundo a autora, a elaboração de “produtos de predição” (ibid., p. 117) por meio de inteligência de máquina se tornou a base de um mercado altamente lucrativo capitaneado pelas *big techs* que a autora chamou de “mercados de comportamentos futuros” (ibid.). Tais mercados, que refletem claramente a racionalidade da financeirização, conforme exploraremos no Capítulo 3, funcionam como espaços onde os dados comportamentais coletados das atividades dos usuários são analisados, transformados em predições, e vendidos para empresas e anunciantes interessados em prever e influenciar os comportamentos futuros dos consumidores¹⁰, como a previsão de engajamento de conteúdos e a personalidade dos usuários. Bruno, Bentes e Faltay (2019, p. 7) destacam que, “mais importante que o modelo de publicidade presente nestas plataformas, é promessa e capacidade de agir sobre os comportamentos enquanto eles acontecem” de modo a intervir sobre o ‘próximo passo’, ‘próximo clique’, ‘próxima interação’. A constante antecipação de nossas condutas online nessas plataformas resulta naquilo que Bruno (2020, p. 263) chama de “um sequestro, no nível cotidiano, do nosso *campo de ação possível*”, ou simplesmente, num “sequestro do futuro” (ibid.)

Para além de contextos estritamente plataformizados, essa *ação sobre o possível* viabilizada pelo *cálculo dos futuros prováveis*, seja em sistemas de decisão automatizada ou nos modelos generativos de IA, aponta para uma questão central de nossa pesquisa. É justamente esse

¹⁰ Os mercados de comportamentos futuros de que fala Zuboff ocorrem sobretudo através dos sistemas de recomendação de anúncios das plataformas.

poder preditivo, que envolve capacidade não apenas de prever, mas, como demonstraremos, também de *produzir performaticamente o futuro que supostamente antecipa*, que transforma o saber baseado em dados massivos e algoritmos preditivos num instrumento multiplamente valioso e disputado na atualidade. Em virtude da centralidade da predição e da intervenção sobre a dimensão temporal para a racionalidade algorítmica, como desenvolvo ao longo desta investigação, proponho entender os algoritmos da Inteligência Artificial enquanto *tecnologias de governo do tempo*. As dimensões mais poderosas destas tecnologias de governo do tempo, argumento, se relacionam ao modo como modulam o curso das ações e eventos possíveis, ou simplesmente, *o campo do possível*.

Esse *governo algorítmico do tempo*, suas matrizes, meios, finalidades e implicações, é um tema que cruzará de modo transversal nossa investigação sob distintos ângulos. Uma vez que estas tecnologias estão intrinsecamente relacionadas ao uso de saberes estatísticos, propomos que um dos principais objetivos e efeitos de seus usos é gerir e controlar os riscos, as incertezas, o acaso, *reduzindo o campo do possível ao campo do provável* e tornando o curso das ações, eventos e comportamentos dos sujeitos mais *previsíveis*. Nesse cenário, os algoritmos preditivos da Inteligência Artificial materializam e mobilizam uma série de questões relacionadas a novos modos de conhecer, ordenar e governar a realidade, de agir sobre o tempo e o futuro sobre as quais nos debruçaremos ao longo desta investigação.

Os fenômenos sociotécnicos que viemos descrevendo podem ser compreendidos na esteira de um projeto de *digitalização e dataficação da vida* cuja fronteira essencial passa a ser aquela que divide o *computável* do *incomputável*. Nesse novo regime, como pontua Mbembe (2019b), “*o que não é computável não existe*”¹¹. Ou seja, em tal contexto só adquire status ontológico (e conseqüentemente político) aquilo que pode ser digitalizado, datificado, processado e interpretado pelas tecnologias e protocolos computacionais. Como afirma Hong, a dataficação “*cria certos mundos e extingue outros*”¹² (Hong, 2022, p. 375), fazendo com que o que é visível e legível para as máquinas passe a ter prioridade (epistêmica e política) sobre o que permanece invisível ou ilegível para estas. Nos termos de Mbembe (2021, p. 88), “[c]omo a cifra se sobrepôs à palavra, o número se tornou o supremo fiador da realidade, em vez de seu indicador”. Condição para que a Inteligência Artificial opere simultaneamente como uma indústria extrativa e como instrumento de conhecimento, esse projeto traz

¹¹ No original: “which is not computable does not exist”.

¹² No original: “creates certain worlds and extinguishes others.”

implícito um impulso totalizante que pretende capturar o planeta numa forma computacionalmente legível (Crawford, 2021). Para Hui, a própria fantasia de uma superinteligência artificial é expressão de uma “forma extrema de computacionalismo, de acordo com a qual o mundo é calculável e poderia ser esgotado através de operações matemáticas” (Hui, 2020, p. 178-9).

A dataficação da vida e a delegação da produção de conhecimento e decisões a sistemas algorítmicos também desencadearam importantes inflexões nos métodos tradicionais de construção de evidências e interpretação dos fenômenos. No paradigmático artigo “O fim da teoria: o dilúvio de dados torna o método científico obsoleto”, o editor da revista *Wired*, Chris Anderson, propunha, ainda em 2008 (Anderson, 2008), que com a disponibilidade de dados massivos não haveria mais a necessidade da ‘velha’ teoria e o método científico — baseado em criar hipóteses e modelos e testá-los — se tornaria obsoleto. “Com dados suficientes, *os números falam por si mesmos*” (ibid.), defendia ele.

Apesar do artigo de Anderson soar simplista até mesmo para o que viria a se consolidar ao longo da última década como a Ciência de Dados, um dos pilares dos saberes algorítmicos, ele manifesta aspectos importantes da epistemologia *data-driven* (Ricaurte, 2019) que hoje orienta grande parte das análises dos fenômenos nos mais diversos campos¹³. Nesse *dataísmo* (Han, 2018), associado ao já velho *tecnosolucionismo*, promove-se a ilusão de que os dados existem em ‘estado natural’, são autoevidentes, neutros e objetivos (Gitelman, Jackson, 2013; Han, op. cit) e de que se os problemas ainda não foram solucionados é porque não há dados suficientes sobre o assunto (logo, é necessário obtê-los). Além de um modelo de gestão da realidade que considera apenas os ‘fatos’ computáveis por dados e algoritmos, o que esse discurso torna opaco é que “*dados brutos são um oxímoro*” (Gitelman, Jackson, op. cit., p. 3). Ainda que guiados por uma racionalidade extrativista, dados não são simplesmente extraídos, mas são sempre resultado de sofisticados processos e técnicas de produção e interpretação maquínicas e humanas. Como bem observa Strecker (2022), “dados não falam por si só, mas são *feitos para falar*”.

Para Mbembe (2021), o advento do que chama de *razão digital* deu vida à antiga fantasia do conhecimento pleno. Essa razão “vê o mundo como um imenso reservatório a ser haurido” (ibid., p. 89). Ao descrever nossa época como a de um *Brutalismo*, caracterizado

¹³ Dotada de um forte viés positivista, a epistemologia centrada em dados baseia-se para Ricaurte (op. cit.) em três suposições: (1) a de que os dados refletem a realidade, (2) a de que a análise desses dados gera um conhecimento valioso e preciso e (3) a de que os resultados do processamento desses dados permite melhorar as decisões sobre o mundo.

especialmente pela aceleração tecnológica e pela transição para uma *civilização computacional*, o filósofo afirma que, nesse novo regime, “mecanismos computacionais, modelagem algorítmica e expansão do capital pela totalidade da vida” (Mbembe, 2021, p. 74) passam a integrar um processo único. Para ele, o projeto de extensão indefinida dos horizontes de cálculo e as formas de extração de recursos que este requer visam, em última instância, “a conversão de qualquer substância em quantidades, o cálculo preemptivo das potencialidades, dos riscos e das contingências, visando sua financeirização, por um lado, e a conversão de propósitos orgânicos e vitais em meios técnicos, por outro” (ibid., grifo nosso). A entrada das tecnologias computacionais e algorítmicas em todas as esferas da vida, diretamente relacionada à infiltração do capital nessas mesmas esferas (sobretudo em sua versão financeirizada), produz o que poderíamos chamar de uma *conversão em larga escala de intensidades em quantidades*. Essa verdadeira *pulsão numérica* monitora, captura, datifica e classifica sujeitos, objetos e fenômenos de modo indiferenciado, criando uma “*equivalência sem igualdade*” (Zuboff, 2021, p. 452) que colapsa a separação moderna entre a representação das coisas e a representação dos sujeitos. A conversão de intensidades em quantidades e a indistinção ontológica entre objetos e sujeitos que esse regime de conhecimento e gestão da realidade introduz estão estreitamente relacionados à *cibernética* enquanto uma das principais matrizes da racionalidade e da temporalidade algorítmica e cujas implicações aprofundaremos em distintos pontos desta tese.

A escalada tecnológica atualmente em curso, que tem a racionalidade algorítmica como uma de suas principais facetas e a cibernética como uma de suas principais matrizes, acarreta dois movimentos complementares. Berardi (2019) descreve esse processo como o de uma importação da lógica do *bios* para as máquinas e uma importação da lógica da *tekné* para a vida humana e social. Ou seja, enquanto passamos por um *devir maquínico* e somos cada vez mais objetificados¹⁴ por meio da conversão e extração de nossos modos de existência em dados, métricas e algoritmos, as máquinas computacionais contemporâneas passam por aquilo que Hui (2019) chama de um *devir orgânico*. Conforme veremos, esse devir orgânico, intrinsecamente relacionado a como as máquinas cibernéticas lidam com a contingência (Hui, 2019, Fazi 2018), dota os algoritmos da IA de capacidades adaptativas que lhes permitem aprender e evoluir autonomamente a partir de novos dados sem a necessidade de serem

¹⁴ É claro que pode-se objetar que esse fenômeno não é novo, uma vez que o homem há muito converteu-se em objeto, como evidenciaram trabalhos como *As palavras e as coisas* de Foucault ([1966], 2000). O que nos parece novo, à semelhança da operação de reconhecimento de padrões que comentamos em nota anterior, é que essa objetificação seja feita por meio de saberes baseados em dados e algoritmos.

explicitamente programados para tanto. Esse é um dos grande diferenciais dos algoritmos aprendizes e o motivo pelo qual se tornaram tão vantajosos para o capitalismo de dados: eles transformam os oceanos de dados que inundam os ecossistemas digitais cotidianamente em produtos de predição (Zuboff, op. cit.) sem o trabalho (e o custo) de se definir regras, modelos ou hipóteses de antemão.

Mas esse *devir orgânico* das máquinas algorítmicas também aponta para o que autores como Mbembe (2021, p. 29) vem chamando de “espetacular retorno do animismo” na atual civilização computacional. Essa dimensão se tornou absolutamente tangível com as perturbações existenciais desencadeadas com a chegada do ChatGPT e seus companheiros e a ameaça ao excepcionalismo humano que materializam, descentrando o humano de privilégios epistêmicos e ontológicos até então relativamente resguardados¹⁵. Grande parte da reverberação das narrativas sobre *riscos existenciais em larga escala* associadas à IA e o retorno de questões como ‘aquilo que *realmente* nos faz humanos’ ou no que consiste ‘a *verdadeira* inteligência’, temas que abordaremos tanto no estudo de caso sobre o ChatGPT quanto no capítulo dedicado aos futuros maquínicos, são sintomas desse ‘abalo narcísico’. Convergindo com o supracitado diagnóstico de Mbembe, Viveiros de Castro (2023) aponta que a chegada dos sistemas de IA mais avançados configuram uma *substituição animista*. Para ele, essas “máquinas sobrenaturais” passam a ocupar um lugar outrora destinado aos espíritos nas cosmologias extramodernas e estão provocando “um enorme reboição” (p. 1) na tríplice fronteira antropológica formada por *humanos, animais e máquinas*¹⁶. Esse ‘reboição’ que já estava presente, por exemplo, no ciborgue de Haraway ([1985], 2019) e nas fronteiras que este fazia ruir, mas só agora essa perturbação parece reverberar de modo mais amplo.

Como exploraremos no Capítulo 4, a “confusão de fronteiras” (Haraway, op. cit.) desencadeada pela chegada dos últimos *habitantes maquínicos* desse triângulo antropológico — que carregam a ambivalência de, sob a superfície apresentarem comportamentos cada vez mais similares ao humano, mas em seu interior serem movidas por um “modelo de mundo” (Amoore et al, 2024) fundamentalmente distintos do ‘nosso’ — integra os efeitos do que Viveiros de Castro (op. cit.) chama de *dupla aceleração* de nossa época: a *aceleração climática* e a *tecnológica*. Assim, sugerimos que é no cruzamento histórico do *Antropoceno*

¹⁵ Talvez uma das questões fundamentais para pensarmos as implicações desse projeto de totalização tecnológica seja, como sugere Patricia Reed (2018), a distinção entre a *descentralização do humano* e *desumanização*, ainda que fazê-lo ainda pareça uma tarefa bastante complexa uma vez que nos encontramos demasiado dentro desse nevoeiro. Como pontua a autora, “a descentralização do humano não é igual à desumanização” (p. 14).

¹⁶ Nessa interpretação, uma versão atualizada da tríade *humanos, animais e espíritos* das culturas extramodernas.

com o *Tecnoceno* (Costa, 2021) que marca nosso turbulento *Tempo do Fim* (Anders, 2007) que os *futuros maquínicos* agenciados pela IA devem ser interpretados, sobretudo em relação às promessas salvacionistas e ameaças catastrofistas que vêm suscitando. Afinal, como especula Lapoujade, “talvez a ideia de futuro não seja outra coisa senão um *diagnóstico do presente*” (Lapoujade, 2015, p. 289).

A noção de *futuros maquínicos* designa aqui uma modalidade específica de *futuro histórico* (Simon, Tamm, 2021) ligada às tecnologias, práticas, discursos, imaginários e ideologias da IA na contemporaneidade. Propomos que estes futuros maquínicos previstos e produzidos pela IA são simultaneamente materiais (manifestos nas próprias previsões algorítmicas) e simbólicos (manifestos nos imaginários, discursos e ideologias promovidos sobre tais tecnologias) e estão hoje programando nossos futuros individuais e coletivos de múltiplas formas.

*

A hipótese central que orienta esta investigação é, como já enunciada, de que em virtude da centralidade da predição e da intervenção sobre a dimensão temporal para a racionalidade algorítmica, os algoritmos da Inteligência Artificial se transformaram em *tecnologias de governo do tempo* na atualidade. Ao longo da tese, essa hipótese se desdobra em múltiplas camadas que vão lhe dotando de distintas nuances. Nesse sentido, uma das questões fundamentais que nos mobilizam é o modo como o regime de conhecimento baseado em dados massivos e algoritmos preditivos está vinculado a um regime de temporalidade que materializa uma relação específica com o tempo e o futuro, ou uma determinada *cronopolítica*. Ao mesmo tempo, essa cronopolítica, marcadamente preditiva, é sintomática das mudanças históricas que marcam nossos afetos atuais em relação ao tempo, como a crescente instabilidade, incerteza e insegurança em relação ao futuro na esfera do comum.

Como sinalizamos no início desta Introdução, um dos principais objetivos ou efeitos deste *governo algorítmico do tempo* é reduzir o campo do possível ao campo do provável, introduzindo uma gestão calculada do acaso ou da incerteza que visa tornar o curso dos eventos, ações e comportamentos mais previsíveis. Mas além da matriz estatística, veremos que outras matrizes compõem a cronopolítica algorítmica, como a matriz cibernética, a matriz securitária-militar e a matriz financeira, produzindo um regime de temporalidade ancorado na *probabilidade*, na *recursividade*, na *antecipação* e na *aceleração*.

Uma das principais perguntas que sustenta nosso percurso se desdobra num duplo movimento complementar: por um lado, o questionamento por que forças (sociais, epistêmicas, materiais, históricas, etc) dão forma à *algoritmização do tempo e do futuro* e, recursivamente (para utilizar a própria estrutura temporal destes dispositivos), quais os efeitos e implicações dessa algoritmização?

*

Antes de detalhar o percurso de cada capítulo da tese, faremos uma breve diferenciação entre dois significantes importantes para as discussões propostas: o *possível* e o *futuro*. Apesar de intimamente relacionados, cada um enfatiza distintas facetas de nossas questões-chave.

Quando falamos do *possível*, ou do *campo do possível*, nos referimos às potencialidades e virtualidades que compõem as multiplicidades disso que usualmente costuma se chamar futuro. Aparentemente abstrato, mas manifesto de modo muito concreto nos processos de algoritmização que descrevemos, o possível tem relação com aquilo que confere *abertura e indeterminação* às trajetórias de atualização do curso dos fenômenos e processos de individuação ao longo do tempo (Simondon, 2009). Já ao utilizar o termo *futuro* nos referimos, preferencialmente, ao *artefato histórico e cultural* através do qual apreendemos aquilo que Koselleck (2006) chama de *horizontes de expectativa* ao longo de distintos estratos históricos. Nesse entender, o futuro manifesta-se sob distintas modalidades históricas e culturais (Cf. Simon, Tamm, 2021) que apontam para diagnósticos específicos do presente (Lapoujade, 2015).

Ou seja, enquanto o *possível* aponta para uma *dimensão mais processual* a partir da qual os eventos, fenômenos e processos de individuação se atualizam (ou não) no tempo, o *futuro* aponta preferencialmente para uma *dimensão mais cultural e histórica*. Apesar disso, a leitora notará que em alguns momentos essas noções se sobrepõem, sobretudo em função de um emprego mais amplo da expressão ‘futuro’ por alguns interlocutores com os quais dialogamos.

*

A tese estrutura-se a partir de dois eixos: um *epistemológico* e um *cronopolítico*. Apesar dessa divisão dupla, destacamos que, em virtude dos problemas centrais que nos mobilizam, o eixo cronopolítico possui uma hierarquia superior ao epistemológico.

O eixo *epistemológico* está centrado sobre o modelo de racionalidade e cognição algorítmicos e sobre o regime de saber que sustenta tal modelo. Este eixo visa estabelecer algumas bases, definições e pressupostos a partir dos quais desdobramos as reflexões que lhe seguem.

No Capítulo 1, começamos explorando a noção de racionalidade algorítmica (Bruno, 2021) e a *modulação algorítmica do possível* que esse modelo de conhecimento e gestão da realidade introduz, enfatizando sua relação com os deslocamentos epistemológicos do pós-guerra e a crise da razão reflexiva moderna. A seguir, baseando-nos sobretudo nas proposições de Hayles (2017), propomos um entendimento dos algoritmos enquanto cognoscentes e o *cognitivo não-consciente* como uma instância fundamental para compreender como a racionalidade algorítmica opera. Finalizamos este capítulo com definições técnicas e teóricas sobre *o que é e como raciocina uma IA* (mais precisamente, algoritmos baseados em aprendizado de máquina e aprendizado profundo) e uma breve incursão sobre a história da Inteligência Artificial, enfatizando a disputa entre os dois principais paradigmas da IA, o *simbólico*, baseado na dedução simbólica, e o *conexionista*, baseado na indução estatística (Cardon, Cointet, Mazieres, 2018) e atual vertente hegemônica.

O Capítulo 2 concentra-se em alguns aspectos do regime de saber algorítmico a partir de *três deslocamentos* que este produz na *ordem das coisas* (Foucault, 2000), materializando modos específicos de apreender, interpretar e produzir realidades: a prevalência da *performatividade* sobre a *representação*; a prevalência da *correlação* sobre a *causalidade*; e a prevalência do *provável* sobre o *possível*.

O eixo *cronopolítico* desdobra-se, por sua vez, em duas vias: num aprofundamento sobre os *regimes de temporalidade algorítmicos* e em reflexões sobre os *futuros maquínicos*.

No Capítulo 3, o mais extenso da tese, exploramos num primeiro momento o que chamamos de *matrizes da cronopolítica algorítmica*, a saber, a *cibernética*, a *militar-securitária* e a *tecnofinanceira*. Tais matrizes nos ajudarão a dar visibilidade às forças (históricas, sociais, econômicas e políticas) que dão forma ao desenvolvimento técnico dos algoritmos ao longo do tempo e a outros aspectos de seu regime de temporalidade, como a recursividade, a antecipação e a aceleração.

A segunda parte deste capítulo dedica-se a dois estudos de casos, um sobre o *TikTok* e outro sobre o *ChatGPT*. Ambos são manifestações contundentes da vertente *preditivo-aceleracionista* (Bruno, Pereira, Faltay, 2023) que passou a dominar a IA nos

últimos anos. Tais casos foram selecionados por evidenciarem problemas de ordem cronopolítica que não apenas materializam questões previamente abordadas ao longo da investigação, mas também introduzem novas, complexificando nossos argumentos e hipóteses. Dentre outras questões, propomos que o TikTok opera enquanto um *ambiente digital de alta frequência*, regido por *microtemporalidades aceleradas e fragmentadas* nas quais identificamos ressonâncias da racionalidade e da temporalidade tecnofinanceira. Já o ChatGPT, além de ter se tornado um objeto paradigmático da entrada oficial na ‘era da IA’ e da ameaça ao excepcionalismo humano que esse novo momento produz, é uma *megamáquina probabilística* que materializa de forma contundente a redução das possibilidades às probabilidades (bem como seus efeitos homogeneizantes) e, paradoxalmente, manifesta comportamentos que tensionam suas próprias lógicas programáticas.

O quarto e último capítulo da tese, de tom mais ensaístico que os demais, se debruça sobre a noção de *futuros maquínicos*, situando essa reflexão num panorama mais amplo sobre futuros históricos (Simon, Tamm, 2021) e sobre as complexidades do nosso denso e turbulento presente (Haraway, 2016). Observamos que os *futuros maquínicos* produzidos pela IA, concretizados tanto nas previsões algorítmicas quanto nos imaginários sobre essas tecnologias na contemporaneidade, são atravessados por uma série de tensões e ambiguidades. Guiando-nos por tais tensões, propomos a hipótese de que os futuros maquínicos da IA produzem uma dupla relação com o futuro. Por um lado, ao nível programático, o futuro materializado nas previsões algorítmicas se manifesta enquanto uma *repetição recursiva* dos passados datificados nos modelos algorítmicos, tornando improvável a ruptura em relação a tais padrões históricos. Por outro, principalmente ao nível dos discursos e imaginários alavancados pelos últimos desdobramentos no desenvolvimento da IA e promovidos por seus ‘tecnoevangelistas’, o futuro é anunciado como uma reiterada *promessa de ruptura* cujo evento prototípico é a Singularidade Tecnológica. Como mostram autoras como Haraway (2018) e Le Guin (2021), na história da tecnociência, promessa e ameaça, salvação e catástrofe, triunfo e tragédia sempre andaram lado a lado. Entretanto, argumentamos que esse “cronotopo das ameaças e promessas finais” (Haraway, op. cit.) mobilizado pela tecnologia ganha outros contornos na confluência da aceleração tecnológica e da aceleração climática que marca nosso presente histórico, levando a um entrelaçamento escatológico entre as “mitologias do fim” (Danowski, Viveiros de Castro, 2017) fabuladas em ambas esferas.

EIXO EPISTEMOLÓGICO

Capítulo 1. Racionalidade e cognição algorítmica

1.1 Racionalidade algorítmica e o governo do possível

A noção de *racionalidade algorítmica* da qual lançaremos mão ao longo deste trabalho pode ser descrita, junto à Fernanda Bruno, como “um modelo de racionalidade onde os algoritmos ocupam um lugar central nos processos de conhecimento de uma certa realidade, bem como nos processos de tomada de decisão e de gestão dessa mesma realidade.” (Bruno, 2021, p. 155). Inspirando-se na noção de racionalidade foucaultiana (Foucault, 2006, 1995), a autora esclarece que “[u]m modelo de racionalidade implica simultaneamente produzir conhecimento e intervir sobre um determinado contexto, problema, fenômeno ou realidade”. (ibid.).

Para Foucault (2006), a racionalidade tem, acima de tudo, um caráter *instrumental*. Trata-se de um modo de organizar os meios para alcançar um fim. Ao refletir sobre as relações entre racionalização e poder no conhecido texto *O sujeito e o poder*, o autor declara: “[c]onsidero a palavra racionalização perigosa” (Foucault, 1995, p. 233). Assim, ao invés de ‘perseguir’ um suposto processo crescente de racionalização, o que se deve fazer, orienta, é “analisar as racionalidades específicas, mais do que evocar constantemente o progresso da racionalização em geral” (ibid.). Referindo-se à abordagem foucaultiana da racionalidade, Castro (2004) adverte que “não se trata de abordar a história da razão como um processo que, apesar de seus retrocessos, adquire um caráter global e unitário, mas *como uma análise das diferentes formas de racionalidade que organizam a ordem das práticas*” (p. 304, grifo nosso).

Também orientados por entendimentos foucaultianos da racionalidade, Aradau e Blanke (2022) afirmam que “a razão algorítmica cria as condições de possibilidade de implementação de algoritmos para governar a conduta de indivíduos e populações, de amigos e inimigos, de normalidade e anormalidade através dos mundos sociais e das fronteiras políticas.”¹⁷ (p. 3). Para eles, trata-se de “uma racionalidade distintiva, que possibilita práticas

¹⁷ No original: “Algorithmic reason renders the conditions of possibility of rolling out algorithms for governing the conduct of individuals and populations, of friends and enemies, of normality and abnormality across social worlds and political boundaries.”

de governo e a produção de sujeitos datificados através da promessa de conhecimento mais preciso e de tomadas de decisão mais eficientes.”¹⁸ (ibid., p.3-4).

Deste modo, no presente contexto, a noção de *racionalidade algorítmica* designará tanto a centralidade que os algoritmos adquirem na contemporaneidade sobre os modos de conhecer, decidir e gerir a realidade, quanto os métodos que utiliza para tanto. Esse modelo de racionalidade ‘algoritmocêntrico’ abarca, simultaneamente, modos específicos de ordenar, classificar, interpretar e capturar a realidade, e modos específicos de agir sobre esta visando intervir sobre comportamentos, eventos e fenômenos.

Nosso entendimento da racionalidade algorítmica também é fortemente influenciado pela noção de *governamentalidade algorítmica*, tal como definida por Rouvroy e Berns (2015), também de inspiração foucaultiana¹⁹, e atualizada noutros desdobramentos de Rouvroy sobre o tema (Rouvroy, 2013, 2020, 2021). A noção visa descrever o surgimento de um modelo de governo do mundo social que não estaria mais baseado em políticas, leis ou normas sociais, mas no processamento algorítmico de grandes volumes de dados. Para os autores, esse novo modelo de ‘apreensão’ da realidade baseado na agregação, análise e correlações estatísticas de dados através do uso do que chamam de algoritmos autodidatas (referindo-se ao aprendizado de máquina) produziu um “novo regime de verdade digital” (ibid., p. 37) que dispensa normas ou hipóteses prévias, uma vez que opera através de métodos indutivos, conforme veremos em mais detalhes ao longo deste capítulo.

Alimentando-se de *dados infra-individuais* para produzir *modelos supra-individuais* (padrões, perfis, etc), segundo os autores, a governamentalidade algorítmica contorna e evita os sujeitos reflexivos, privilegiando as *relações* ao invés de *sujeitos*. Como assinala Rouvroy,

¹⁸ No original: “a distinctive rationality, which makes possible governing practices and the production of datafied subjects through the promise of more precise knowledge and more efficient decision making.”

¹⁹ Ainda que sejam bastante pontuais as passagens em que estabelecem uma relação mais direta entre a hipótese da governamentalidade algorítmica e o conceito de dispositivo de segurança em Foucault (2008), para Rouvroy e Berns, a governamentalidade algorítmica corresponderia ao dispositivo de segurança definido pelo autor. Segundo a hipótese foucaultiana, o dispositivo da segurança surge no século XVIII e se distingue tanto do mecanismo jurídico quanto disciplinar, e pode ser definido como a regulação “de um meio no qual não se trata tanto de determinar localizações, mas, sobretudo, essencialmente de possibilitar, garantir, assegurar circulações: circulação de pessoas, circulação de mercadorias, circulação do ar, etc.” (Foucault, 2008, p. 39). Enquanto a lei proíbe e a disciplina prescreve, a segurança, “sem proibir nem prescrever, mas dando-se evidentemente alguns instrumentos de proibição e de prescrição, (...) tem essencialmente por função responder a uma realidade de maneira que essa resposta anule essa realidade a que ela responde-anule, ou limite, ou freie, ou regule. Essa regulação no elemento da realidade é que é, creio eu, fundamental nos dispositivos da segurança.” (ibid., p.61). A proposta dos autores é, portanto, entender a governamentalidade algorítmica como uma nova ruptura no interior do modelo dos dispositivos de segurança proposto por Foucault (ver nota 13 do supracitado artigo).

o governo algorítmico “gera ‘alertas’ ou ‘estímulos’ para obter reflexos e não reflexividade” (2020, p. 18). Trata-se, assim, de um governo “relativamente subliminar” (ibid.) que opera, como exploraremos ao longo deste trabalho, primordialmente ao nível *não-consciente* (Hayles, 2017; Chun, 2016; Cesarino, 2021). Contudo, como bem pontuam Bruno, Bentes e Faltay (2019, p. 9), “esta condução algorítmica de condutas pode não levar em conta sujeitos específicos, mas não deixa de mirar alvos”.

Marcada por três tempos, a saber, o *dataveillance* (a coleta e conservação automatizada de quantidades massivas de dados não classificado), o *datamining* (o uso de saberes estatísticos para estabelecer correlações) e o que se poderia chamar de *perfilização* (a antecipação de comportamentos através de perfis), a governamentalidade algorítmica designa, para Rouvroy e Berns,

globalmente um certo tipo de racionalidade (a)normativa ou (a)política que repousa sobre a coleta, agregação e análise automatizada de dados em quantidade massiva de modo a modelizar, antecipar e *afetar, por antecipação, os comportamentos possíveis*. (Rouvroy, Berns, 2015, p. 42, grifo nosso)

Em consonância com a leitura dos autores franceses, Bruno observa que “é precisamente sobre a *ação possível dos indivíduos* que incide a atenção e o interesse dos diversos ramos que se dedicam ao conhecimento e ao controle de condutas nos ambientes digitais.” (Bruno, 2020, p. 263, grifo nosso). Num sentido similar, Bucher pontua como as dimensões mais poderosas dos algoritmos não estão necessariamente localizadas nos algoritmos (embora possam), mas têm a ver sobretudo “com as maneiras pelas quais esses sistemas *governam o campo possível de ação dos outros* e como essas possibilidades se tornam mais ou menos disponíveis ou indisponíveis para determinados atores em contextos específicos” (Bucher, 2019, grifo nosso). Portanto, propomos que é principalmente aquilo que entendemos aqui como o *campo do possível* que a racionalidade algorítmica visa conhecer e sobre o qual visa agir.

Essa regulação do possível que é o alvo privilegiado da racionalidade algorítmica remete ao próprio modo como Foucault define o exercício do poder. Para o autor, este exercício consiste essencialmente em “um conjunto de ações sobre ações possíveis” que “opera sobre o campo de possibilidades onde se inscreve o comportamento dos sujeitos ativos” (1995, p. 243). Com base nessa concepção, entendemos que é justamente na maneira como este governo, para utilizar expressões do próprio autor, conduz as condutas, ordena probabilidades, age sobre as ações, facilitado-as ou dificultando-as, ampliando-as ou limitando-as, tornando-as mais ou

menos prováveis, que reside o próprio da governamentalidade algorítmica e as implicações mais cruciais da racionalidade que lhe orienta. Dito de modo conciso, sua especificidade reside no modo como *governa o possível*.

Uma vez que, como vimos na Introdução, o possível aqui se refere à parcela de potencialidade e virtualidade do curso das ações, eventos, comportamentos e processos de individuação no tempo, podemos dizer que há uma dimensão temporal intrínseca a esse modelo de racionalidade e governo da realidade. Ainda que seus efeitos apontem para o presente, como pontuam Rouvroy e Berns, “*o campo de ação deste ‘poder’ não está situado no presente, mas no futuro*”. Essa forma de governo trata essencialmente daquilo que poderia acontecer, das propensões, mais do que das ações realizadas” (Rouvroy, Berns, op. cit., p. 48, grifo nosso).

Por esta razão, proponho que os algoritmos preditivos da Inteligência Artificial atuam contemporaneamente enquanto *tecnologias de governo do tempo* que operam uma *modulação algorítmica do possível*²⁰. Essa modulação do possível visa controlar a agência do acaso e da incerteza e tornar o curso das ações, eventos e comportamentos dos sujeitos mais previsíveis. Como veremos, isso é feito, essencialmente, através do uso de saberes estatísticos baseados em correlações e no reconhecimento de padrões.

A *modulação algorítmica do possível* a que nos referimos se manifesta concretamente no modo como tais algoritmos regulam, por exemplo, o que é visível e acionável nas plataformas digitais: a que ‘mundo visível’ cada usuário tem ou não acesso, que ações, escolhas e comportamentos lhe são permitidos, proibidos, induzidos e/ou evitados num determinado contexto, que tipos de perfis lhes são atribuídos para a oferta de conteúdos personalizados, etc. Atuando numa camada invisível e infraestrutural das mediações digitais (aquilo que os profissionais da ‘indústria *tech*’ chamam de *backend*), o algoritmo se converte nesse contexto num componente histórico crucial daquilo que Bruno (2015, p. 15) entende como um *regime de visibilidade*: “não tanto [o] que é visto, mas [o] que torna possível o que se vê”.

Para além da regulação do campo perceptivo nas plataformas digitais, a modulação do possível aqui também diz respeito ao modo como tais algoritmos, integrados a Sistemas de

²⁰ O uso do termo modulação aqui não é fortuito. Ele se relaciona, por um lado, à proposição de Deleuze (1992) de que o modelo de poder do controle age como uma *modulação* em oposição aos *moldes* que caracterizavam o modelo disciplinar. Por outro, a modulação também se aproxima da própria lógica operativa dos algoritmos de aprendizado de máquinas baseados na indução e na capacidade de aperfeiçoar recursivamente seu desempenho.

Decisão Automatizada (*Automated Decision Making Systems*, em inglês) que envolvem em geral *cálculos de risco*, intervêm sobre decisões em áreas cujas implicações biopolíticas (Foucault, 2008b) e efeitos sobre populações marginalizadas são enormes. Seja justiça penal, na contratação de trabalhadores, na saúde, na educação, na segurança pública, na gestão de políticas políticas ou na concessão de crédito financeiro, esses sistemas vêm estabelecendo, “o que é mais ou menos correto, mais ou menos merecedor de direitos, o que é o bom desempenho e o improdutivo, o mais ou menos ‘suspeito’”. (Moraes, no prelo).

Já num plano epistemológico, essa regulação do possível se relaciona, entre outras coisas, às “políticas da classificação” (Crawford, 2021) que condicionam o modo como os dados são capturados, classificados e nomeados e o que é reconhecido (ou não) enquanto padrão. Em suma, os sistemas de classificação através dos quais a realidade passa a ser simultaneamente categorizada e produzida. Isso porque a racionalidade algorítmica orientada por dados implica um amplo e intenso processo de (re)classificação e (re)segmentação do mundo, dos problemas, fenômenos, sujeitos, identidades, etc. Esses processos se baseiam em pressupostos que costumam permanecer tácitos, mas cujas implicações vêm afetando fortemente o modo como tais ‘objetos’ são concebidos e, no limite, criados uma vez que seus métodos simultaneamente ‘descrevem’ e criam as realidades visam apreender (Chun, 2018), conforme discutiremos em mais detalhes no tópico 2.1.

Para Crawford (op. cit), essas políticas da classificação atravessam toda a produção de conhecimento algorítmica e orientam desde o modo como a IA é produzida (dos centros de pesquisa à ‘indústria *tech*’), como os bancos de dados de treinamento são estruturados e utilizados, como esquemas específicos de ordenamento social naturalizam hierarquias e ampliam desigualdades, como modelos taxonômicos que nada possuem de neutro ‘esquadrinham’ o mundo, e assim por diante.

As dimensões acima descritas materializam como a crescente utilização de sistemas algorítmicos baseados em Inteligência Artificial nos mais diversos contextos (individuais e coletivos, públicos e privados) vêm reconfigurando nossas *políticas do possível* em múltiplos níveis. A maneira como tais tecnologias ordenam, interpretam e intervêm no mundo vem impactando “questões epistemológicas, princípios de justiça, organização social, expressão política, cultura, o entendimento dos corpos, subjetividades e identidades”; em suma, “o que somos e o que podemos ser” (Crawford, op. cit. p. 19).

*

Nos tópicos que compõem este capítulo e o próximo nos deteremos sobre diversos aspectos do modo como o que viemos descrevendo como uma racionalidade algorítmica, simultaneamente, produz conhecimentos e intervém sobre determinados problemas, fenômenos, sujeitos e realidades. Para utilizar outro termo foucaultiano, se trata de uma incursão sobre a “ordem das coisas” (Foucault, 2000, p. XV) no regime epistemológico materializado no uso de algoritmos preditivos sobre conjuntos massivos de dados.

Ao descrever o *a priori* que definiria de antemão o espaço das semelhanças e das diferenças para cada estrato histórico, o autor define tal ordem como

uma definição dos segmentos sobre os quais poderão aparecer as semelhanças e as diferenças, os tipos de variação de que esses segmentos poderão ser afetados, o limiar, enfim, acima do qual haverá diferença e abaixo do qual haverá similitude. (Foucault, 2000, p. XV)

No contexto desta investigação, trata-se do modo como o regime epistemológico centrado em dados e algoritmos captura, interpreta, classifica, codifica e simultaneamente produz realidades. Isso porque, como veremos, essas novas formas de perceber e conhecer o mundo são, ao mesmo tempo, *modelos de mundo e práticas criadoras de mundos* (Cf. Amoore et al, 2024; Barad, 2020; Haraway, 2008, 2016).

Essa ‘ordem algorítmica das coisas’ é o resultado de uma multiplicidade de fatores que incluem desde questões diretamente ligadas aos fenômenos da digitalização e do desenvolvimento técnico da IA (incluindo as disputas entre suas distintas correntes), mas também deslocamentos epistêmicos históricos mais amplos e profundos, sobretudo aqueles ligados ao que poderíamos chamar de uma *episteme e ontologia do pós-guerra*, que exploraremos sobretudo no próximo tópico.

Indagar pelo modo como a racionalidade algorítmica instaura seus códigos ordenadores e constitui seu regime de saber requer também uma atenção especial aos modos de existência (Simondon, 2008) específicos dos algoritmos: suas lógicas programáticas, linguagens, operações internas, modos de raciocínio e de cognição, etc. Para compreender esses *modos de existência algorítmicos*, é necessário realizar um ‘mergulho’ na dimensão programática dos aparelhos (Flusser, 2017), deslocando-nos do antropocentrismo e do excepcionalismo humano para acessar uma espécie de ‘perspectivismo maquínico’. Isso implica uma postura teórica que não ignore que a realidade materializada nas máquinas é “realidade humana, gesto

humano fixado e cristalizado em estruturas que funcionam” (Simondon, 2009, p. 34), mas que também reconheça aí modos *outros-que-humanos* de existência, agência, racionalidade e cognição²¹.

1.2 Racionalidade algorítmica e a crise da razão reflexiva

O modelo de racionalidade que tem os algoritmos como centro ordenador da realidade pode ser visto como um dos efeitos de um deslocamento epistêmico mais amplo e profundo, consolidado entre o pós-segunda guerra e a Guerra Fria, e que “marca a passagem do modelo iluminista de razão fundamentado na reflexividade crítica para um modelo de racionalidade baseado em regras algorítmicas.”²² (Bruno, 2021, p. 156). Naquele contexto histórico, o espectro da guerra global e da destruição nuclear levaram a uma crítica à tomada de decisões humanas (Cf. Halpern, 2022; Daston, 2022), fomentando a proposição de um modelo formal, repetível e algorítmico de tomada de decisões. Assim, através de uma composição disciplinar entre computação, a ainda incipiente inteligência artificial, teoria dos jogos e da decisão, psicologia social experimental, estratégia militar e novas teorias econômicas, a concepção iluminista da razão baseada na verdade e na faculdade de julgamento foi sendo reconceitualizada em termos de padrões e a razão passou a ser concebida em termos de procedimentos de decisão com a tarefa de calcular probabilidades (Daston, op. cit.). Como observa Daston em suas pesquisas sobre o tema, nesse período, “o sonho de reduzir inteligência, tomada de decisões, planejamento estratégico e a própria razão a regras algorítmicas havia se espalhado como fogo pela psicologia, economia, teoria política, sociologia e até mesmo filosofia”²³.

Poucas disciplinas materializaram tão bem a consolidação desse novo modelo de racionalidade do pós-guerra quanto a *cibernética*, uma das grandes matrizes epistemológicas, senão ‘a grande matriz’ da racionalidade algorítmica. Racionalidade algorítmica e racionalidade cibernética são, de muitos modos, inseparáveis, seja em relação aos modos de

²¹ Essa espécie de dupla natureza ou ambivalência parece estar inclusive em Simondon quando diz que o objeto técnico é ‘realidade humana cristalizada’ ao mesmo tempo que reconhece neles um modo particular de existência.

²² Como observa Bruno (op. cit.), não se trata de uma passagem abrupta e completa de um modelo a outro. Na prática, há convivências e sobreposições entre ambos.

²³ No original: “the dream of reducing intelligence, decision-making, strategic planning, and reason itself to algorithmic rules had spread like wildfire to psychology, economics, political theory, sociology, and even philosophy”. O trecho integra a descrição do projeto de Daston “How Reason Became Rationality”: https://www.mpiwg-berlin.mpg.de/research/projects/DeptII_Daston_Reason

conceber a realidade ou de governá-la. Base do funcionamento de máquinas que vão desde smartphones até robôs e tecnologia espacial, a disciplina é definida pelo matemático estadunidense Norbert Wiener, seu fundador e teórico mais célebre, como uma ciência que trata da *comunicação e do controle entre animais, máquinas e humanos* (Wiener, 2017). Nascida das trincheiras da Segunda Guerra como uma ciência que poderia ser útil na previsão de comportamentos humanos e maquínicos, a principal proposição da cibernética consiste em afirmar que os sistemas biológicos, técnicos e sociais compartilham de um mesmo plano ontológico (Rodríguez, 2019) e podem, portanto, ser controlados através dos mesmos princípios, a saber, a *informação* — entendida como a medida da organização de um sistema — e o *feedback* — uma causalidade recursiva através da qual estabilidade dos sistemas é regulada (Wiener, op. cit; Hui, 2019). Segundo o coletivo Tiqqun (2010), o modelo de controle transversal da cibernética concebe os comportamentos biológicos, físicos e sociais como integralmente programados e re-programáveis. Para eles, a cibernética é um elemento central tanto das técnicas de individuação quanto de totalização da atualidade.

A matriz epistemológica da cibernética também foi essencial para a consolidação de uma nova concepção de equivalência entre os processos cognitivos humanos e maquínicos. Como afirma Flusser, para quem a cibernética seria a melhor aproximação de uma cosmovisão do pós-guerra por ser “uma visão do aparelho” (2017, p. 287)²⁴, “a cibernética liquida a coisa pensante, porque demonstra que o pensamento no significado cartesiano não é uma coisa que duvida, mas uma coisa que calcula” (ibid. p. 263-4). Por isso Heidegger associou a ascensão da cibernética e o “triunfo do equipamento controlável de um mundo técnico-científico e da ordem social que lhe corresponde” (1979, p. 271) ao *fim da Filosofia* e à sua dissolução nas “ciências tecnicizadas” (ibid.). Nesse novo paradigma, segundo ele, “[p]assa a imperar o elemento racional e os modelos próprios do pensamento que apenas representa e calcula” (ibid.). Segundo Parisi (2017), ao substituir o julgamento baseado na suposição de categorias por estados de verdade realizados por máquinas, o raciocínio instrumental da cibernética absorveu a metafísica ocidental. Assim, ao paralelizar o entendimento de mentes humanas e lógicas computacionais, a cibernética também tornou-se o alicerce para as emergentes ciências da comunicação e da informação, especialmente para o ainda incipiente campo da Inteligência Artificial.

²⁴ Para ele, o aparelho é justamente “um sistema no sentido cibernético” (ibid. p. 287).

Nessa transição de modelos de racionalidade que estamos tematizando, paulatinamente ao longo das últimas décadas, a *decisão* se converteu numa operação epistemológica central. Suas implicações passaram a estar intimamente relacionadas ao *risco* e o processo de tomada de decisão, seja pelo humano ou pela máquina, foi se tornando um objeto privilegiado das práticas e técnicas de saber. Por um lado, alavancado pela popularização de áreas como a Economia Comportamental, cresce a descrença no humano como tomador de decisões racional (Ariely, 2008; Thaler, Sunstein, 2019)²⁵. Uma vez que nossas decisões já não são confiáveis como outrora, o mais recomendável é que sejam assistidas, seja por algoritmos e/ou técnicas comportamentais que intervenham sobre nossas “arquiteturas de escolha” (Zuboff, 2021; Thaler, Sunstein, op. cit.). Além disso, como aponta Berardi, “quanto mais complexa e veloz é a informação que circula, menor se torna o tempo disponível para examinar a informação necessária para realizar uma decisão consciente” (2019, p. 131). Como resultado desse deslocamento, as decisões, sejam no âmbito privado ou público, passam a ser consideradas mais objetivas, neutras, confiáveis e ágeis quanto mais automatizadas forem.

A delegação dos processos interpretativos e decisórios para máquinas computacionais por meio de Sistemas de Decisão Automatizada, sobretudo em contextos públicos, atesta também uma esquiva da responsabilização que é (ou deveria ser) uma prerrogativa da política. Para Rouvroy (2021), nesse modelo de governo trata-se de “fazer com que governem os algoritmos para não ter que assumir a carga de governar, ou seja, de tomar decisões que devem sustentar e pelas quais devem ser responsáveis”²⁶. Vendido como uma otimização da eficiência que reduz os custos do Estado, esse modelo de gestão da esfera pública que, como veremos, se relaciona a um *governo dos efeitos* no lugar das *causas* (Morozov, 2018), também torna explícito por que a adoção destes sistemas é tão conveniente ao neoliberalismo com seus programas de austeridade e esvaziamento do comum.

O que a passagem de um modelo de racionalidade fundado na reflexividade crítica iluminista para um modelo de racionalidade fundado em regras algorítmicas que viemos descrevendo ao longo deste tópico atesta é que, como afirma Mbembe refletindo sobre o futuro da razão num

²⁵ Segundo tais perspectivas, ao contrário do que propunha a economia clássica, que assumia os indivíduos enquanto agentes racionais guiados pela maximização da utilidade, nossas decisões do dia-a-dia são “*previsivelmente irracionais*” (Ariely, op. cit.).

²⁶ No original: “Hacer que gobiernen los algoritmos para no tener que asumir la carga de gobernar, es decir, de tomar decisiones que deben sostener y de las que han de rendir cuentas.”

contexto de hegemonia computacional, “[a] razão pode ter atingido seus limites. Ou, em qualquer caso, (...) a razão está sob julgamento”²⁷ (2021b, p. 26). Nessa ruptura, a razão, outrora aquilo que nos distinguiu enquanto espécie e que orientou a “tarefa das Luzes” (Foucault, 2006), é cada vez mais substituída por um modelo de racionalidade que se desdobra das máquinas, computadores e algoritmos. Assim, a razão não é mais vista como atributo do excepcionalismo humano, mas uma faculdade compartilhada com atores técnicos, agora dotados de uma *razão sintética* e considerados, inclusive, mais confiáveis que nós. Conforme veremos a seguir, esse modelo de racionalidade baseado numa *razão sintética* que se afasta do modelo de subjetividade reflexiva moderna e, portanto, da consciência, opera, primordialmente, por meio de *processos cognitivos não conscientes* (Hayles, 2017).

1.3 Algoritmos enquanto cognoscentes e o não-consciente cognitivo

No discurso popular, tornou-se comum a referência ao algoritmo como ente único — ‘o algoritmo’ — seguido da menção ao ‘proprietário’ deste artefato sociotécnico ao qual atribui-se poderes simultaneamente mágicos e persecutórios (Faltay, 2020). Assim, em geral, ouve-se expressões como ‘o algoritmo do Google’, ‘o algoritmo do Facebook’, ‘o algoritmo do Tik Tok’, ‘o algoritmo do Instagram’, ‘o algoritmo da Netflix’, e assim por diante. Em tais contextos, o algoritmo encarna uma espécie de ‘pequeno deus’ dotado de poderes oraculares capaz de ‘adivinhar’ desejos e propensões, sobretudo aqueles não explicitamente enunciados dos sujeitos. Tais afetos e imaginários são bem sintetizados na máxima de que o algoritmo é capaz de *nos conhecer melhor do que nós mesmos* (Bruno, Bentes, Faltay, 2019; Strecker, 2022). A radical ‘entrada em cena’ da IA Generativa nos últimos anos, principalmente com o ChatGPT, acentuou ainda mais a projeção de poderes oraculares e mágicos a essas máquinas, ainda que aqui se trate menos de um saber capaz de desvelar a verdade do sujeito do que um suposto ‘*dr. sabe-tudo*’.

Apesar da mistificação, esses imaginários algorítmicos (Bucher, 2018) são manifestações do reconhecimento por um público amplo sobre a centralidade que os algoritmos ocupam no atual modelo de negócios das plataformas digitais e dos saberes e técnicas de influência que sustentam tal modelo, como a perfilização algorítmica e a exploração de vulnerabilidades psicológicas, assim como a fantasia de conhecimento pleno que move a razão digital (Mbembe, 2021). Além disso, o entendimento do algoritmo enquanto “peça de mágica

²⁷ No original: “reason may well have reached its limits. Or, in any case, it is a time when reason is on trial”.

cotidiana”²⁸ (Finn, 2017, p. 16) não deixa de ser uma sobrevivência das relações históricas entre técnica, religião e magia. Para Simondon (2008), esse entrelaçamento está na gênese da tecnicidade, o modo de relação do homem com o mundo estabelecido através dos objetos técnicos.

Já no contexto técnico, sobretudo para os ‘fazedores de algoritmos’ da ciência da computação, o algoritmo é considerado um *processo* ou *procedimento* articulado a um *objetivo*: “uma receita, uma série de instruções, uma sequência de tarefas para alcançar um cálculo ou resultado particular”²⁹ (Finn, p. 17); uma “série lógica de etapas para organizar e agir sobre um corpo de dados para alcançar rapidamente o resultado desejado.”³⁰ (Gillespie, 2016, p. 19); “uma sequência de instruções que diz a um computador o que fazer”³¹ (Domingos, 2017, p.35); ou, simplesmente, “um método para resolver um problema”³² (Finn, op. cit., p. 18). Tais concepções são tributárias da formalização da noção de *computabilidade* e da Máquina de Turing em 1936, modelo fundamental da computação moderna. Isso resultou no entendimento de que um algoritmo é um conjunto de instruções sequenciais, contáveis, separáveis e finitas executáveis por uma Máquina de Turing (tese Church-Turing).

Para que tais instruções sejam operacionalizadas, é necessário que se especifique regras segundo as quais funções do código devem (ou não) ser executadas. Um dos principais mecanismos que regula a aplicação dessas regras é a estrutura de seleção decisional ‘*se, então, senão*’. Ela permite que um programa execute uma determinada ação baseada em uma condição específica³³. Todos os meios computacionais empregam uma estrutura lógica ‘*se, então, senão*’ explícita ou implícita que conecta escolhas ou seleções com consequências. A presença de uma sintaxe de comando e controle que possibilita que os algoritmos ‘façam coisas’ guarda íntima relação com a *natureza performativa* destes artefatos, assunto que será desenvolvido no tópico 2.1.

²⁸ No original: “pieces of quotidian technical magic”.

²⁹ No original: “a recipe, an instruction set, a sequence of tasks to achieve a particular calculation or result”.

³⁰ No original: “the logical series of steps for organizing and acting on a body of data to quickly achieve a desired outcome”.

³¹ No original: “a sequence of instructions telling a computer what to do”.

³² No original: “a method to solve a problem”.

³³ Se a condição for verdadeira, a ação especificada no ‘então’ será executada; caso contrário, a ação especificada no ‘senão’ será executada.

No entanto, ainda que todas essas concepções e definições sejam extremamente válidas para entender diversas dimensões da cultura algorítmica (Seyfert, Roberge, 2016) na qual estamos atualmente imersos, assim como os princípios fundamentais de seu funcionamento técnico, elas parecem não captar os deslocamentos epistemológicos em curso com a crescente *automação cognitiva* e interpenetração entre cognição técnica e humana nos modos de saber e de ser contemporâneos. Além disso, a lógica de funcionamento dos algoritmos connexionistas que, como veremos no tópico 1.4, rompem com modelos de algoritmos baseados em regras explícitas, complexificando o entendimento do ‘algoritmo como uma receita’.

Nesse ponto, o trabalho recente de Katherine Hayles (2017), alinhado a um materialismo pós-humanista, nos fornece um modelo teórico que consideramos produtivo para compreender os deslocamentos em jogo na nova “ecologia cognitiva planetária” (ibid., p. 3), assim como para um entendimento não antropocêntrico acerca dos algoritmos e seus processos epistemológicos, trazendo novas nuances ao modelo de racionalidade que é alvo de nosso interesse.

Partindo do pressuposto de que a cognição é um processo muito mais amplo do que a consciência e o pensamento, e que se trata de um *processo* e não de um *atributo*, Hayles defende que ela não se limita ao humano, mas está presente em algum grau em todas as formas biológicas e sistemas técnicos complexos, de plantas a algoritmos. Embora essa cognição que excede a consciência receba vários nomes, a autora, alinhada à descobertas recentes da neurociência e da biologia cognitiva, propõe designá-la *cognição não-consciente*. A instância do *cognitivo não-consciente* opera num nível de processamento inacessível à consciência, mas desempenha funções essenciais para esta, alimentando-a com seus *outputs* por meio de retroalimentações recursivas entre ambas instâncias³⁴.

Apesar da cognição técnica das máquinas computacionais ter sido comumente comparada às operações da consciência (como o próprio termo Inteligência Artificial sugere), segundo essa perspectiva ela estaria muito mais próxima dos processos performados pelo cognitivo

³⁴ Optamos por traduzir *unconscious* como *não-consciente* ao invés de *inconsciente* no intuito de diferenciá-lo do inconsciente psicanalítico. A própria autora estabelece essa distinção, ainda que utilize o mesmo termo: “O novo inconsciente difere do inconsciente psicanalítico de Freud e Lacan por estar em comunicação contínua e fácil com a consciência. Nessa visão, o inconsciente psicanalítico pode ser considerado um subconjunto do novo inconsciente, formado quando algum tipo de trauma intervém para interromper a comunicação e isolar essa parte da psique do acesso consciente direto. No entanto, o inconsciente psicanalítico ainda se expressa à consciência por meio de sintomas e sonhos suscetíveis de interpretação psicanalítica.” (ibid., p. 27).

não-consciente humano. Esta instância exerce funções semelhantes em sistemas humanos e sistemas técnicos complexos, compartilhando certas similaridades estruturais e funcionais, como o processamento rápido de informações, o reconhecimento de padrões, a produção de inferências para antecipação de eventos futuros.

No entendimento da cognição proposto por Hayles, a cognição é um *espectro* e não um ponto único, é um *processo* e não uma entidade. Por isso, é inerentemente dinâmica. Nessa concepção, *interpretação* e *escolha* exercem papel fundamental. Segundo a autora, a cognição é um processo dinâmico “que interpreta informações em contextos que as conectam com significado”³⁵ (ibid., p. 22). Assim, é a capacidade de realizar (ou não) esse processo que diferencia aquilo que chama *cognoscentes* e *não-cognoscentes*; uma distinção mais equilibrada para a autora do que o binômio *humano/não-humano* que por muito tempo dominou o debate nas humanidades e ciências sociais. *Cognoscentes* seriam, portanto, atores, biológicos ou técnicos, capazes de *interpretação e escolha*³⁶. Além desta distinção, Hayles, baseando-se na Teoria Ator-Rede de Latour (2012) mas desenvolvendo suas próprias contribuições, diferencia *agentes*³⁷ de *atores*. Enquanto agentes são forças materiais que têm a capacidade de agir, atores são *cognoscentes* que exercem escolhas e tomam decisões. Portanto, da perspectiva de Hayles, a questão fundamental não é simplesmente a possibilidade de agência, mas de escolha que cada ator pode realizar num determinado agenciamento.

Hayles enfatiza como os meios computacionais são um tipo especial de tecnologia uma vez que são dotados de um potencial evolutivo mais forte do qualquer outra em virtude de suas capacidades de simular outros sistemas. Podemos afirmar que isso inclui a capacidade de simular processos cognitivos humanos que implicam a consciência, como o uso de linguagem verbal e raciocínio abstrato, ainda que através de processos não-conscientes. De fato, é exatamente isso que está ocorrendo com os grandes modelos de linguagem nesse momento do desenvolvimento da IA.

³⁵ No original: “that interprets information within contexts that connect it with meaning”.

³⁶ Hayles esclarece que a interpretação pressupõe a possibilidade de escolha. Deve haver mais de uma opção para uma interpretação ser possível. Nos meios computacionais essa escolha pode corresponder ao simples binário *um* ou *zero* ou ao constructo *se, então, senão* anteriormente mencionado.

³⁷ Aqui optamos por traduzir “agents” como “agentes”, mas há traduções que dão preferência ao termo “actantes”.

Atualmente, cognição humana e técnica, sobretudo por meio das tecnologias computacionais, estariam numa relação tão simbiótica que Hayles se refere a uma complexa interpenetração entre ambas, resultando naquilo que chama de *agenciamentos cognitivos*, fortemente dotados de efeitos emergentes e habilidades de transformação ao longo do tempo. À diferença das concepções de agenciamento de Latour e Deleuze e Guattari, a noção de agenciamento cognitivo proposta por Hayles pretende enfatizar o “fluxo de informações através de um sistema e as escolhas e decisões que criam, modificam, e interpretam esse fluxo”³⁸ (ibid., p. 116).

Híbridos por natureza, [os agenciamentos cognitivos] levantam questões sobre como a agência é distribuída entre os cognoscentes, como e de que maneiras os atores contribuem para a dinâmica sistêmica e, conseqüentemente, como as responsabilidades — técnicas, sociais, legais, éticas — devem ser distribuídas.³⁹ (ibid., p. 119)

Ainda que obviamente exista uma profunda diferença entre a responsabilidade ética dos atores humanos em relação aos técnicos, uma vez que são aqueles que desenham, implementam e supervisionam os sistemas cognitivos complexos, assim como definem o grau de autonomia técnica dado aos cognoscentes não humanos, para Hayles, ambos são *atores éticos* no sentido de que estão aptos a executar ações que possuem implicações éticas.

Inspirando-nos nas proposições de Hayles, entenderemos os algoritmos como *cognoscentes técnicos* que atuam, junto a outros atores, humanos e técnicos, em agenciamentos cognitivos imersos em fluxos de informação performando *processos cognitivos não-conscientes*. Os algoritmos baseados em aprendizado de máquina ou conexionistas, em especial, são cognoscentes técnicos capazes de *reconhecer padrões e fazer inferências* a partir destes padrões, como veremos em mais detalhes no tópico seguinte.

É importante enfatizar que tais algoritmos não apenas agem por meio de uma cognição não-consciente, mas intervém sobre os processos cognitivos humanos nessa mesma esfera, sobretudo em contextos plataformizados. Autoras como Cesarino (2022) e Chun (2016) vêm destacando como é justamente no plano dos *processos cognitivos primários* (Bateson, 1972) — aqueles relacionados principalmente à formação de hábitos e memórias que escapam à consciência reflexiva — onde ocorre a conexão ou sincronização entre cognição humana e

³⁸ No original: “the flow of information through a system and the choices and decisions that create, modify, and interpret the flow”.

³⁹ No original: “Hybrid by nature, they raise questions about how agency is distributed among cognizers, how and in what ways actors contribute to systemic dynamics, and consequently how responsibilities—technical, social, legal, ethical—should be apportioned.”

maquínica nas redes digitais e onde os efeitos sobre a atenção e os comportamentos dos usuários visados pelas grandes plataformas são produzidos. Esse aspecto também se manifesta naquilo que Bentes (2022, p. 136) chama de “modelo tecnobehaviorista” utilizado pelas grandes plataformas digitais em suas técnicas de influência digital e estratégias de produção de engajamento. Tal modelo, destaca a autora, privilegia táticas que ‘contornam’ a consciência.

Essas transversalidades e ressonâncias cognitivas demonstram que enquanto as narrativas da *Singularidade Tecnológica* (Cf. Vinge, 1993, Kurzweil, 2005) especulam sobre o momento em que as máquinas se tornarão autoconscientes e superarão a capacidade cognitiva conjunta da massa cinzenta do planeta, as dinâmicas ‘mundanas’ da digitalização e da algoritmização da vida tornam explícito que as relações entre agência humana e maquínica existem e continuarão a existir por meio de sobreposições e hibridismos. Atentar para a dimensão inerentemente híbrida da atual ecologia cognitiva planetária não significa, no entanto, abrir mão de uma leitura atenta às assimetrias, tensões e conflitos que atravessam e dão forma a tais agenciamentos. Afinal, sabe-se que é justamente através de uma assimetria brutal entre usuários e plataformas que o capitalismo de vigilância opera e produz os efeitos visados sobre os usuários (Zuboff, 2021). A falta de atenção às assimetrias de poder no interior destes agenciamentos cognitivos é uma legítima crítica que se pode fazer à proposição de Hayles, inclusive. De todo modo, questionar o que se torna a cognição, ou mesmo a inteligência, em contextos de alta interpenetração e recursividade entre agentes técnicos e humanos, pode ser uma ocasião para recolocar tais questões a partir de outra perspectiva que não a da mera oposição ou competição (Cf. Bruno, Pereira, Faltay, 2023).

Ao deslocar a centralidade de processos cognitivos conscientes para o entendimento das transformações na atual paisagem midiática, ainda que não trate a questão nesses termos, as proposições de Hayles ajudam a reconhecer os contornos de um modelo de racionalidade ancorado em processos cognitivos não-conscientes que estão na base do regime de saber algorítmico e para o declínio do excepcionalismo humano. Nesse sentido, talvez uma das perguntas com as quais poderíamos permanecer é quais as implicações da crescente delegação e automação cognitiva a um tipo de racionalidade que não se funda na consciência.

1.4 Inteligência Artificial: definições e paradigmas

1.4.1 O que é e como raciocina uma IA?

Nos últimos anos, o termo Inteligência Artificial se tornou onipresente nas mídias e no debate público sobre tecnologia. Apesar de ser utilizado como sinônimo de todo o ramo da ciência da computação que se concentra na criação de sistemas capazes de realizar funções cognitivas similares aos humanos, como aprendizado, resolução de problemas e compreensão de linguagem verbal, na prática, o que vem se chamando popularmente de IA é apenas a subárea deste campo concentrada no uso de técnicas de aprendizado de máquina. O aprendizado de máquina pode ser descrito como a capacidade que os algoritmos que utilizam tais métodos, especialmente as redes neurais, têm de *reconhecer padrões* através de um cálculo indutivo sobre grandes conjuntos de dados (Cardon, Cointet, Mazieres, 2018; Pasquinelli, 2017). É essa capacidade que confere aos algoritmos ditos inteligentes habilidades cognitivas não-conscientes (Hayles, op. cit) que lhes permitem realizar tarefas de *interpretação, classificação e predição* (Pasquinelli, Joler, 2020). Quando se diz que um algoritmo inteligente faz uma predição, o que ele faz, na prática, é descobrir e aprender padrões a partir de um conjunto de exemplos (*conjunto de dados de treinamento*) e, a partir disso, produzir um modelo (*modelo estatístico*) capaz de gerar inferências por meio de generalizações sobre novos conjuntos de dados, calculando a probabilidade de determinados resultados ocorrerem no futuro com base nos padrões disponíveis.

Num exercício imaginativo, se um algoritmo de *machine learning* ‘falasse’, suas sentenças seriam algo do tipo: ‘é provável que você goste deste filme’, ‘é provável que você cometa um crime’⁴⁰, ‘é provável que você seja homossexual’⁴¹, ‘é provável que você seja um bom pagador’, ‘é provável que você compre este produto’, ‘é provável que você esteja grávida’⁴², ‘é provável que isso seja uma maçã, um spam, um rosto, uma fraude’, etc. Ainda que no ‘âmbito humano’ e ético exista uma diferença radical entre as sentenças hipotéticas

⁴⁰ Esse tipo de cálculo é realizado por sistemas como o PredPol, de policiamento preditivo, e o COMPAS, utilizado pelos tribunais dos EUA para avaliar a probabilidade de um réu se tornar um reincidente. À respeito, ver, por exemplo Dressel, Farid (2018).

⁴¹ Michal Kosinski, um dos criadores do teste de personalidade que esteve na base do caso Cambridge Analytica, e Yilun Wang, publicaram pesquisa afirmando que redes neurais são capazes de prever orientação sexual a partir de imagens faciais. Ver Kosinski, Wang (2018).

⁴² Fazemos menção ao conhecido caso da rede Target, de 2012, que ‘descobriu’ a gravidez de uma jovem antes mesmo de seus pais através da análise de dados visando ações direcionadas de marketing.

enunciadas acima, no âmbito do aparelho (Flusser, 2017) todas são produzidas pela mesma lógica de raciocínio.

As diversas aplicações de aprendizado de máquinas geralmente são divididas em duas grandes modalidades: *classificação* e *previsão* (ou regressão) (Cf. Pasquinelli, Joler, 2020). A classificação é normalmente empregada para reconhecer um sinal ou um objeto e atribuí-lo uma categoria de acordo com uma taxonomia pré-definida (um rótulo ou classe, p.e.). Importantes áreas da IA como visão computacional e reconhecimento facial podem ser agrupadas nesta categoria. Já a modalidade da previsão é utilizada para projetar tendências futuras através da extrapolação de dados do passado. Sistemas de recomendação algorítmica, sistemas baseados na previsão de riscos e toda a vertente da *IA Generativa*, ou seja, modelos algorítmicos capazes de produzir conteúdos sintéticos em texto, imagem ou áudio, podem ser incluídos na vertente preditiva.

Na base de ambas as modalidades estão as operações cognitivas chave dos algoritmos de aprendizado de máquina: o *reconhecimento* e a *geração de padrões* (Pasquinelli, Joler, 2020)⁴³ viabilizados pelo cálculo de correlações entre pontos de dados. Correlações medem, basicamente, a relação entre uma variável de entrada e uma variável de saída. Variáveis altamente correlacionadas são consideradas preditoras umas das outras (no meio técnico, diz-se que elas são *'proxies'*). Isso significa que ao rastrear uma, é possível inferir ou prever a variável oculta ou latente⁴⁴. Essas relações são então computadas numa função que pode ser aplicada a novos *inputs* para prever futuros *outputs* (Pasquinelli, 2019).

Como se vê, a interpretação e produção de padrões ocupa um lugar central no regime de saber algorítmico, uma vez que são as operações que delimitam os segmentos a partir dos quais se distinguem as regularidades e as irregularidades, as similitudes e diferenças, as normalidades e as anomalias, o visível e o invisível nos mundos codificados por algoritmos e dados massivos. Por este motivo, Apprigh (2018) propõe entender a discriminação de padrões num duplo sentido: simultaneamente um axioma fundante das culturas computacionais utilizado para descrever a imposição de identidade a um conjunto de dados com o objetivo de

⁴³ De modo simplificado, enquanto a classificação depende apenas do reconhecimento de padrões, a predição opera numa dupla camada, reconhecendo padrões e gerando padrões.

⁴⁴ A principal técnica para medir correlações entre variáveis é o Coeficiente de Pearson, utilizado desde as origens da estatística moderna. O coeficiente varia de -1 a 1 onde: 1 indica correlação positiva (à medida que uma variável aumenta, a outra também aumenta de maneira proporcional); -1 indica correlação negativa (à medida que uma variável aumenta, a outra diminui de maneira proporcional) e 0 indica que não há correlação linear entre as variáveis.

filtrar (ou seja, discriminar) informações e a dimensão intrinsecamente política que tais práticas materializam, com seus critérios de inclusão e exclusão, suas práticas de segregação e políticas de identidade, que frequentemente amplificam padrões históricos discriminatórios e excludentes.

Uma das principais implicações da centralidade dos padrões nos “modelos de mundo” (Amoore et al, 2024) algorítmicos é a tendência à *homogeneização em larga escala* e à *repetição recursiva dos passados datificados* nos conjuntos de treinamento que este modelo epistêmico produz, problema que discutiremos em distintos pontos desta pesquisa. Uma vez que os algoritmos preditivos reconhecem padrões “apenas dentro dos limites lógicos definidos pelos dados de treinamento” (Pasquinelli, Joler, 2020), suas operações terminam por derivar num comportamento que gira “paranoicamente em torno de padrões estabelecidos ao invés de reconhecer novas correlações” (Pasquinelli, 2017, p. 8). Essas limitações da IA são descritas por Pasquineli e Joler (2020) como o problema da “não detecção do novo”, a incapacidade dos modelos de aprendizado de máquina de reconhecer uma anomalia que aparece uma única vez, e da “regeneração do antigo”, “a aplicação de uma visão espaço-temporal homogênea que restringe a possibilidade de um novo evento histórico”. Como questiona Beiguelman (2021), ao refletir sobre os impactos da IA nos novos modos de ver, perceber e figurar a realidade: “e o que fica fora do padrão? Que lugar social poderá ocupar [nesse regime de saber]?” (p. 136).

Um dos aspectos cruciais que distingue os algoritmos aprendizes de seus predecessores é que eles são capazes de aprender a partir de exemplos e de aperfeiçoar recursivamente seu desempenho ao longo do tempo com base em novos dados. Essa habilidade é especialmente presente nas redes neurais profundas (*deep neural networks*, em inglês, base do chamado *deep learning*) utilizadas em aplicações como reconhecimento de imagens, processamento de linguagem natural, veículos autônomos e ferramentas de tradução. Enquanto os chamados *algoritmos baseados em regras* seguem um conjunto pré-definido de instruções para resolver problemas, os algoritmos aprendizes inferem as regras a partir de um grande número de exemplos de treinamento sem a necessidade de serem explicitamente programados para tanto⁴⁵ (Bucher, 2018). Por isso, são também chamados de algoritmos não baseados em regras (*non-ruled-based algorithms*). Como descreve um material publicitário da IBM sobre as

⁴⁵ Quando um modelo é treinado com dados ditos rotulados, ou seja, dados identificados com rótulos que designam características, classes ou propriedades, diz-se que ele é do tipo *supervisionado*. Quando emprega conjuntos de dados não rotulados, diz-se que é do tipo *não-supervisionado*.

vantagens do uso de aprendizado de máquina, nesse contexto, “você não inicia o processo de aprendizado decidindo com antecedência qual será a resposta para o problema”⁴⁶ (IBM, 2018, p. 34). O aprendizado de máquina consiste, portanto, essencialmente, no desenvolvimento de *algoritmos capazes de executar tarefas a partir de exemplos*⁴⁷.

Isso explica tanto o *extrativismo de dados* que sustenta o modelo hegemônico da IA quanto o domínio das *big techs* nesta indústria: as práticas da captura e acumulação massiva de dados por parte das grandes plataformas digitais resultaram na geração de gigantescos bancos de dados que passaram a ser utilizados para treinamento algorítmico e foram uma das condições para o desenvolvimento do aprendizado de máquina (Celis Bueno, Schultz, 2021; Crawford, 2021). Além disso, quanto maior o acesso quantitativo e qualitativo a dados, maior o *potencial preditivo* dos modelos, uma vez que mais padrões são incorporados, e maior a complexidade e sofisticação das tarefas que este é capaz de executar. Se a consolidação do capitalismo de vigilância teve como uma de suas principais facetas na última década a extração de dados comportamentais para direcionamento de publicidade personalizada nas plataformas (Zuboff, op. cit.), a ‘corrida da IA’ atualmente em curso, focada em grandes modelos, direciona esse extrativismo sobretudo para a consolidação de grandes bancos de dados para treinamento algorítmico, ainda que, na prática, tais finalidades tendam a se sobrepor.

O que a capacidade dos algoritmos da IA de aprender a partir de exemplos massivos evidencia, como diversos autores vêm ressaltando, é que eles operam através de uma *lógica de raciocínio indutiva* e não de uma *lógica dedutiva*⁴⁸ (Cf. Amoore, 2019; Rouvroy, Stiegler,

⁴⁶ No original: “you don’t start the learning process by deciding in advance what the answer to the problem will be”.

⁴⁷ Apesar das capacidades de aprender automaticamente partir de enormes bases de dados que estes sistemas possuem, é fundamental lembrar que grande parte das exaltadas habilidades autônomas e inteligentes só são viabilizadas por muito trabalho humano precarizado, fragmentado e intencionalmente invisibilizado que formam o ‘chão de fábrica da IA’. Tratam-se de trabalhadores terceirizados, massivamente do Sul Global, que realizam uma série de microtarefas repetitivas e mal remuneradas como classificar e rotular imagens e vídeos, escutar e transcrever áudios, moderar conteúdo potencialmente nocivo. Estes trabalhadores atestam a formação de um verdadeiro *precarizado algorítmico* que silenciosamente otimiza as capacidades técnicas desses sistemas e é omitido dos discursos sobre as virtudes das tecnologias que aperfeiçoam (Jones, 2021; Pereira, 2019). Para uma investigação sobre a questão no contexto brasileiro, ver a série especial do Intercept Brasil “O chão de fábrica da IA”: <https://www.intercept.com.br/especiais/chao-fabrica-inteligencia-artificial/>.

⁴⁸ A lógica clássica concebe duas classes fundamentais de inferência: a *dedução* e a *indução*. Enquanto a dedução é um tipo de inferência que extrai uma conclusão já contida nas premissas, a indução é um tipo de inferência que, após considerar um número suficiente de casos particulares, conclui uma verdade ou lei generalizável. A indução pode também ser entendida como uma crença cuja probabilidade é calculável (Flusser, 2017). Para Peirce, nem dedução nem indução são capazes de originar novas ideias, apenas de repetir fatos quantitativos. O tipo de raciocínio capaz de originar novas ideias, para o filósofo pragmático, é a *abdução* (ou inferência hipotética): “A abdução consiste em estudar os fatos e elaborar uma teoria para explicá-los” (Peirce

2016; Cardon, Cointet, Mazieres, 2018; Pasquinelli, 2017; Parisi, 2017). Essa característica é o resultado de uma história cheia de controvérsias, disputas e alternâncias entre os dois principais paradigmas que marcaram a história da IA ao longo das últimas décadas: o *paradigma simbólico* e o *paradigma conexionista*. De modo muito sintético, enquanto o primeiro se baseia na *dedução simbólica*, o segundo se baseia na *indução estatística*. No tópico a seguir, exploramos essas duas vertentes através de um brevíssimo recorrido pela história da IA.

1.4.2 Simbólicos e conexionistas: duas abordagens da IA

A partir de meados de 2010, as técnicas algorítmicas preditivas baseadas em aprendizado de máquina, sobretudo as redes neurais profundas, passaram a obter um desempenho inédito em campos como o reconhecimento de imagens e a tradução automática, tornando-se, ao longo da última década o paradigma dominante da IA. Bruno, Pereira e Faltay (2023) descrevem esse modelo epistêmico e de negócios como *preditivo-aceleracionista*⁴⁹, uma vez que está centrado na produção de previsões e numa otimização tecnológica ininterrupta e acelerada.

Apesar do paradigma conexionista, cujas origens remontam ao início das ciências da computação e da cibernética, ter alçado alguma importância dos anos 40 aos 60, ele esteve dos anos 60 até meados dos anos 90 marginalizado em relação à vertente simbólica (hoje chamado de GOFAI - *Good Old-Fashioned Artificial Intelligence*, em inglês). Entrecortando este período, houve também momentos de uma diminuição significativa no interesse, financiamento e progresso na pesquisa e desenvolvimento da Inteligência Artificial, conhecidos como ‘invernos da IA’⁵⁰. É somente após o *boom* das redes neurais profundas nos

apud Pasquinelli, 2017, p. 8). Pasquinelli propõe que a indução estatística performada por redes neurais se aproxima de uma forma de abdução leve, mas que está longe “de romper planos semióticos que não estavam conectados ou concebíveis de antemão, como em descobertas científicas ou na criação de metáforas” (ibid., p. 9, tradução) e que corresponderiam à abdução forte.

⁴⁹ É importante pontuar que quando utilizamos o significante *aceleracionista* para descrever o modelo atualmente hegemônico da IA não nos referimos ao *aceleracionismo* como corrente filosófica de esquerda que eclodiu no começo dos anos 2000, liderada por nomes como Nick Land, Alex Williams e Nick Srnicek e que propunha, *grosso modo*, que a melhor resposta política radical ao capitalismo não seria resistir através de suas próprias contradições, mas acelerar suas próprias tendências ao colapso (para um compilado sobre o tema, ver, p.e., Avanesian, 2017). Ao empregarmos o termo, no entanto, nos referimos ao processo mais geral de aceleração vertiginosa que marca o capitalismo recente ligado às indústrias de tecnologias digitais e que tem como ideologia de fundo visões aceleracionistas do capitalismo, mas sem os componentes supostamente niilistas e contestadores da supracitada corrente.

⁵⁰ Costuma-se relatar dois destes invernos. O primeiro durou de meados dos anos 70 até o começo dos anos 80, quando tanto os projetos simbólicos quanto conexionistas permaneceram ‘congelados’. O segundo, mais breve, ocorreu na primeira metade dos anos 90 (C.f. Pasquinelli, 2017; Cardon, Cointet e Mazieres, op. cit.).

anos 2010, tracionado pelo acúmulo do Big Data, pelo aumento vertiginoso da capacidade de processamento computacional, mas também por forças sociais, econômicas e políticas sobre as quais comentaremos brevemente ao final deste tópico, que o paradigma conexionista se torna realmente dominante e o próprio termo Inteligência Artificial se populariza.

Cardon, Cointet e Mazieres (op. cit.) relatam a tumultuada disputa entre estes dois modelos de IA como a transição de um paradigma baseado em *máquinas hipotéticas dedutivas* para *máquinas indutivas* ou *conexionistas*. Esta disputa reflete também deslocamentos no próprio entendimento da cognição humana e modo como ela é ‘transferida’ para as máquinas computacionais. No cerne do debate entre as duas vertentes, segundo Bridle (2019, p. 229), estava “não apenas uma discussão sobre o que significa ser inteligente, mas o que é *inteligível* a respeito da inteligência”. Para Parisi (2017), a passagem da hegemonia simbólica para a conexionista marcou uma mudança de concepção da IA como um *modo de validação* (paradigma simbólico) para um *modo de descoberta* (paradigma conexionista). As duas abordagens de IA também correspondem também a dois ‘modelos de mente’, em teoria, transpostos para processos cognitivos algorítmicos. Para a abordagem simbólica, alinhada ao cognitivismo clássico, o pensamento “consiste no cálculo de símbolos que possuem tanto uma realidade material quanto um valor de representação semântica”⁵¹ (Cardon, Cointet e Mazieres, op. cit., p. 4). Ou seja, para os simbolistas, a inteligência é fruto da manipulação de símbolos. Em termos simples, descreve Bridle (op. cit., p. 229), “exigia-se algum conhecimento de mundo para raciocinar significativamente sobre ele”. As máquinas inteligentes simbólicas fazem uso de regras explícitas definidas de antemão por seus programadores. Por isso, operam segundo um raciocínio dedutivo. Um dos maiores expoentes da vertente simbólica foi Marvin Minsky, desenvolvedor da primeira rede neural e fundador do Laboratório de IA do MIT com John McCarthy, que cunhou o termo Inteligência Artificial em 1956 na conferência de Dartmouth, considerada o marco inicial da Inteligência Artificial como campo de pesquisa e estudo.

Em contraste, para a abordagem conexionista, dispensando princípios de representação interna ou intenção, o pensamento se assemelha ao cálculo paralelo e maciço de funções elementares, distribuídas por uma rede neural, “cujo comportamento significativo só aparece no nível coletivo como um efeito emergente das interações produzidas por essas operações

⁵¹ No original: “thinking consists of calculating symbols that have both a material reality and a semantic representation value.”

elementares”⁵² (ibid, p. 4). Para dizê-lo de modo sintético, enquanto a IA simbólica opera a partir de operações do tipo *top-down*, a IA conexionista opera a partir de operações do tipo *bottom-up* (Pasquinelli, 2017). A origem da abordagem conexionista remete ao seminal ensaio de Alan Turing sobre IA de 1950 no qual sugere: “em vez de tentar produzir um programa para simular a mente de um adulto, por que não tentar produzir um que simule a de uma criança?” (Turing, 1950, p. 18). Para os conexionistas, portanto, a inteligência é uma propriedade emergente das conexões entre neurônios e, ao transpor este comportamento para as máquinas, elas podem ser induzidas a pensar (Bridle, 2019).

Máquinas conexionistas são sistemas indutivos e adaptativos capazes de descobrir e gerar novas conexões a partir de contínuos processos de treinamento. Conforme esclarece Amoore (2019, p. 2), “diferentemente das formas dedutivas de raciocínio, em que uma regra ou hipótese é formulada e testada empiricamente, esses algoritmos geram indutivamente atributos potenciais a partir de padrões em um corpus de dados”⁵³. Para dar um exemplo prático, quando uma rede neural é treinada para distinguir dados de cachorros, não é necessário que se defina de antemão características a serem reconhecidas nas imagens. Ao invés disso, mostra-se muitos exemplos para a rede do que é um gato e do que é um cachorro até que ela aprenda distinguí-los, generalizando sobre novas imagens a partir destes exemplos particulares.

As redes neurais atuais são a versão mais avançada e sofisticada da primeira máquina conexionista, o paradigmático *Mark I Perceptron*, desenvolvida em 1957 pelo psicólogo e cientista da computação Frank Rosenblatt no Cornell Aeronautical Laboratory, inspirado no modelo neural proposto por McCulloch e Walter Pitts⁵⁴ (Figura 1). Construído com o propósito de reconhecimento de imagens, o *Perceptron* se tornou o emblema de uma nova máquina computacional de aprendizagem dotada de comportamento considerado inteligente. Na publicação em que descreve o mecanismo e as expectativas de que dispositivos desse tipo

⁵² No original: “which only appears on the collective level as an emerging effect of the interactions produced by these elementary operations.”

⁵³ No original: “Unlike deductive forms of reasoning, where a rule or hypothesis is formulated and tested empirically, these algorithms are inductively generating potential attributes from the patterns within a corpus of data.”

⁵⁴ De inspiração biológica, o modelo neural formal proposto por McCulloch e Pitts em 1943 se tornou um objeto muito relevante para os cibernéticos da época e a pedra angular da calculadora das primeiras máquinas ditas inteligentes (Cardon, Cointet e Mazieres, op. cit.).

seriam capazes de formação de conceitos, tradução de idiomas, inteligência militar e solução de problemas por meio de lógica indutiva, Rosenblatt afirma:

Um requisito primário de um sistema deste tipo é que ele deve ser capaz de reconhecer padrões complexos de informação que são fenomenalmente similares [...] um processo que corresponde aos fenômenos psicológicos de "associação" e "generalização de estímulos". O sistema deve reconhecer o "mesmo" objeto em diferentes orientações, tamanhos, cores ou transformações e em uma variedade de fundos diferentes. Deve ser viável construir um sistema eletrônico ou eletromecânico que aprenda a reconhecer semelhanças ou identidades entre padrões de informação óptica, elétrica ou tonal, de uma maneira que pode ser muito análoga aos processos perceptivos de um cérebro biológico. *O sistema proposto se baseia em princípios probabilísticos ao invés de determinísticos para seu funcionamento, e adquire sua confiabilidade a partir das propriedades de medidas estatísticas obtidas de grandes grupos de elementos*⁵⁵. (Rosenblatt apud Pasquinelli, 2017, p. 4, grifo nosso)

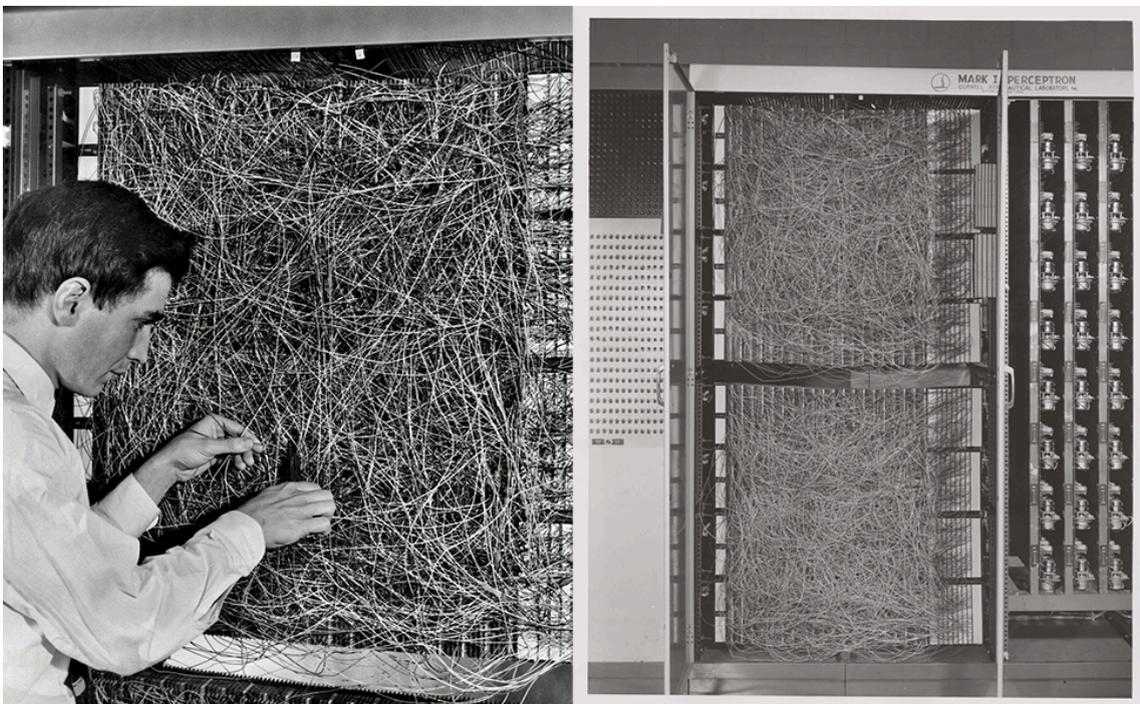


Figura 1: Rosenblatt manipulando o Perceptron (esquerda) e parte da máquina conexionista (direita).

⁵⁵ No original: “A primary requirement of such a system is that it must be able to recognize complex patterns of information which are phenomenally similar [...] a process which corresponds to the psychological phenomena of “association” and “stimulus generalization.” The system must recognize the “same” object in different orientations, sizes, colors, or transformations, and against a variety of different backgrounds. [It] should be feasible to construct an electronic or electromechanical system which will learn to recognize similarities or identities between patterns of optical, electrical, or tonal information, in a manner which may be closely analogous to the perceptual processes of a biological brain. The proposed system depends on probabilistic rather than deterministic principles for its operation, and gains its reliability from the properties of statistical measurements obtained from large populations of elements.”

Como assinala Pasquinelli (2019), desde o *Perceptron*, o objetivo principal do aprendizado de máquinas é armazenar um modelo estatístico que sintetiza a multiplicidade de um conjunto de dados de treinamento num único arquivo, ao invés de ‘memorizar’ milhares de imagens de um mesmo objeto de distintos ângulos. Além de materializar o *reconhecimento de padrões* e a *probabilização de inferências* como operações elementares das máquinas conexionistas, o *Perceptron* estabeleceu a arquitetura básica de uma rede neural, vigente até hoje. Esta arquitetura é composta de três tipos de camadas de neurônios: camada de entrada (*input layer*), camadas ocultas (*hidden layers*) — que são cada vez mais numerosas, daí o termo *deep learning* — e camada de saída (*output layer*) (Figura 2). Cada neurônio da rede funciona como um nó e as relações entre estes são reguladas por *feedback loops* que, por meio de tentativas e erros, fazem com que o modelo ‘aprenda’ determinado padrão. No caso do reconhecimento de imagens, por exemplo, cada camada ‘extrai’ um tipo de informação sobre a imagem: brilho dos pixels, formas e bordas, partes de objetos, objetos inteiros, etc, até que a camada de saída seja capaz de tomar uma decisão sobre o que consiste (ou não) a imagem. A complexidade de uma rede neural é determinada pela quantidade de camadas ocultas e de neurônios que essas camadas possuem, possibilitando o cálculo paralelo de muitos parâmetros.

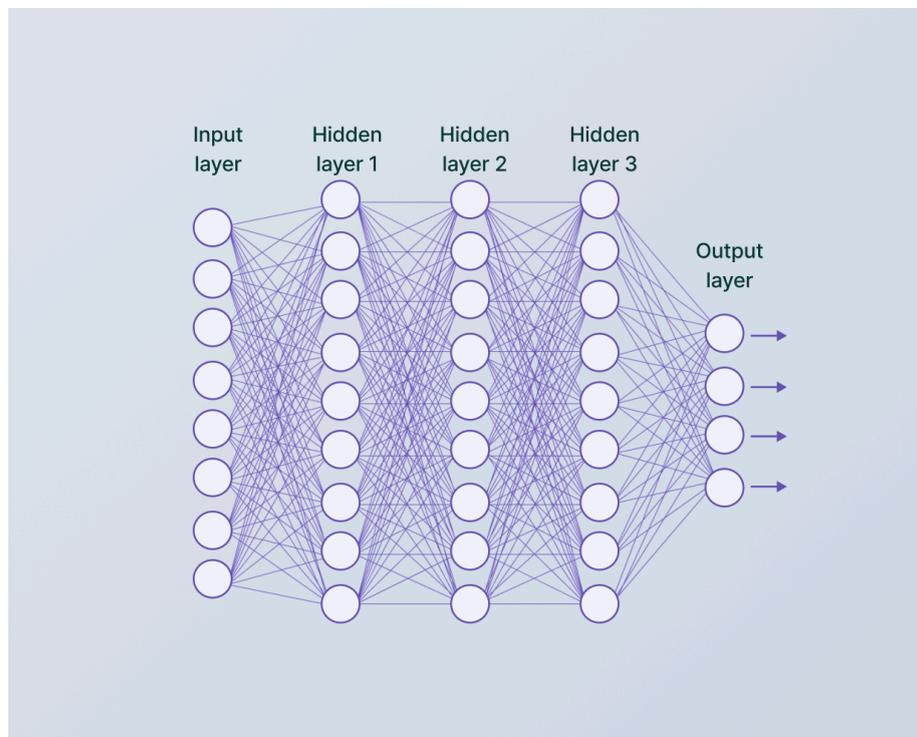


Figura 2: Arquitetura de camadas de uma rede neural. Fonte: V7 labs⁵⁶

Propondo um modelo analítico para conceitualizar comparativamente as abordagens simbólica e conexionista, Cardon, Cointet e Mazieres (op. cit.) propõem que uma máquina dita inteligente articula um *mundo*, uma *calculadora* e um *alvo*. O *mundo* corresponde ao meio, aos dados, à base de conhecimento, em suma, aos *inputs* que alimentam o sistema; a *calculadora* é o programa, o cálculo, o modelo algorítmico; o *alvo* são os objetivos, os resultados, os *outputs*. Com base na articulação desses elementos, os autores apontam que enquanto as máquinas simbólicas inserem na calculadora tanto o mundo (os dados) quanto o alvo (os objetivos), as atuais máquinas conexionistas ‘esvaziam’ a calculadora para que o mundo adote seu próprio alvo. Nesse sentido, é possível afirmar que existe uma dimensão ‘caixa-preta’ intrínseca a toda máquina conexionista, uma vez que sua calculadora, ou seja, seu modelo algorítmico, não é definido de antemão por seus programadores, mas se torna uma função dos dados com que é treinado. É comum que os agentes humanos envolvidos no desenvolvimento destes modelos não consigam compreender totalmente como os resultados foram produzidos (Celis, Schultz, op. cit). Esse traço é mais radical nas chamadas redes neurais profundas, uma vez que suas operações cognitivas envolvem camadas ocultas e bilhões ou até trilhões de conexões (como nos grandes modelos de linguagem, por exemplo). Por este motivo, a *interpretabilidade*, a possibilidade de compreensão por humanos das decisões realizadas por modelos algorítmicos, e a *explicabilidade*, a capacidade de um modelo oferecer alguma explicação sobre os resultados que apresenta, tornaram-se nos últimos anos áreas de pesquisa extremamente visadas, integrando a chamada *IA Explicável* (Vieira, 2020).

Em termos simondonianos (2008) podemos dizer que as máquinas conexionistas são dotadas de um alto grau de *tecnicidade*, uma vez que preservam certa “margem de indeterminação” (p. 33) que lhes permite ser sensíveis a uma informação exterior e se aperfeiçoar ao longo do tempo. Ao refletir sobre o modo de existência dos objetos técnicos, Simondon argumenta que o simples automatismo corresponde a um grau muito baixo de aperfeiçoamento técnico. “O verdadeiro aperfeiçoamento das máquinas, aquele que *eleva o grau de tecnicidade*, corresponde não a um acréscimo de automatismo, mas, pelo contrário ao fato de que o funcionamento de uma máquina *preserve uma certa margem de indeterminação*” (2008, p. 33, grifo nosso). Como aponta Parisi (2017), um fator crucial para os ordenamentos do saber que fundamentam o regime de conhecimento destas máquinas, especialmente os algoritmos

⁵⁶ <https://www.v7labs.com/blog/convolutional-neural-networks-guide>

de aprendizagem profunda, reside na capacidade que têm de raciocinar ‘com e por meio’ do *desconhecido*, da *incerteza* e da *indeterminação*. Simultânea e paradoxalmente, um dos efeitos do uso destas tecnologias é, justamente, um ‘encolhimento’ da agência do acaso, do indeterminado, do improvável nos fluxos de automatização nos quais se inserem. Bonaldo (2023) enfatiza como, desde suas origens, o aprendizado de máquina esteve ligado ao controle da contingência e a uma elaboração conceitual do mundo segundo a atividade estatística. Ao longo deste trabalho, veremos como esse aspecto, diretamente relacionado à estrutura temporal recursiva das máquinas cibernéticas, é um aspecto crucial para compreender as cronopolíticas da IA.

A oposição entre os dois paradigmas da Inteligência Artificial, que exploramos ao longo deste tópico, nos interessa sobretudo na medida em que permite reconhecer as disputas, heterogeneidades e rupturas nas matrizes epistemológicas que dão forma ao longo do tempo àquilo que estamos chamando de racionalidade algorítmica. A passagem do predomínio de uma lógica dedutiva para uma lógica indutiva que a inflexão das máquinas simbólicas para as máquinas conexionistas sintetiza aponta para uma transformação importante nos modos como a realidade é apreendida, ordenada e produzida através dos modelos de raciocínio das máquinas computacionais e suas lógicas programáticas.

Simultaneamente, é interessante ressaltar como os dois paradigmas também materializam de modo historicamente situado distintos entendimentos do humano supostamente transpostos para os modelos cognitivos que cada corrente propõe. Nesse sentido, não se trata de entender a história da IA como a de uma crescente acurácia na artificialização de processos cognitivos humanos, mas de refletir como esses deslocamentos, entre outros aspectos, colocam em distintas co-determinações o humano e o maquínico. Um dos pontos cruciais dessa questão, alinhado às discussões que viemos mobilizando ao longo deste capítulo, é o lugar dado (ou não) à consciência em cada matriz, seja no próprio design dos sistemas algorítmicos ou nos modelos cognitivos que priorizam. Enquanto para a vertente simbólica a consciência ocupa um lugar fundamental, seja na elaboração das regras e hipóteses através das quais os sistemas funcionam ou na suposição (ainda de ecos iluministas) de que a inteligência é fruto da manipulação de símbolos por uma razão abstrata, a corrente conexionista, com seu design *bottom-up* sem regras de funcionamento definidas de antemão, aponta mais uma vez para um ocaso do sujeito reflexivo e para o privilégio das instâncias do cognitivo não-consciente (Hayles, 2017) no regime de conhecimento que subjaz à racionalidade algorítmica.

Essa topologia também se manifesta num capítulo bastante curioso da história da IA. Curiosamente, um dos grandes defensores dos princípios conexionistas foi Friedrich Hayek, tributado como o ‘pai do neoliberalismo’ (Cf. Bridle, op. cit; Pasquinelli, 2021; Halpern, 2022). Equiparando o modelo de mente conexionista ao modelo neoliberal de mercado e sociedade enquanto sistemas distribuídos, adaptativos e auto-organizados, cujo funcionamento não se submeteria a uma consciência ou controle central, as proposições de Hayek sobre o conexionismo e as ciências cognitivas foram fundamentais para uma naturalização da teoria da ‘ordem espontânea do mercado’ do neoliberalismo⁵⁷. Ler Hayek por essa lente, sugere Pasquinelli (op. cit.) ajuda a iluminar a influência da racionalidade econômica nos primeiros paradigmas da IA (vigente até hoje), a rastrear ressonâncias entre *modelos de mente, de mercados e de máquinas* no pós-guerra, e a registrar a influência de forças políticas e sociais na criação de tais modelos.

Atentando para tais forças na primeira década dos anos 2000, convém pontuar brevemente, por fim, alguns fatores externos às querelas entre simbólicos e conexionistas que ajudam a explicar a ‘virada conexionista’ que pavimentou a *era da IA* que vivemos atualmente. Além de mudanças cruciais nas infraestruturas de cálculo, como o incremento na capacidade de processamento das máquinas computacionais e a disponibilidade inédita de dados sob o paradigma do Big Data que ocorreram naquele período, frequentemente mencionados quando se narra essa história, a atual hegemonia do paradigma conexionista também pode ser interpretada como uma resposta à crescente demanda por uma *racionalidade preditiva e antecipatória* nos modos de conhecer e gerir a realidade. Essa demanda foi impulsionada, entre outros fatores, pela consolidação de um novo paradigma securitário-militar pós 11/9 voltado para a identificação antecipada de riscos e ‘desconhecidos’ (Chamayou, 2015); pelo processo de plataformação da web (Helmond, 2015, Srnicek, 2017) e a consolidação de uma economia da vigilância que visava elaborar novos modos de lucrar com o extrativismo de dados (Zuboff, 2021); e pelas múltiplas convergências entre as dinâmicas da neoliberalização e da digitalização (Cesarino, 2022). Por suas capacidades de reconhecer padrões e fazer previsões, as máquinas conexionistas respondiam de modo mais adequado a tais demandas, mostrando-se tecnologias alinhadas às profundas mudanças sociais, econômicas e políticas do início do século XXI.

⁵⁷ Essas ideias encontram-se especialmente em *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*, publicado em 1952.

Capítulo 2. Regime de saber algorítmico: três deslocamentos na ‘ordem das coisas’

A que lei obedecem essas mutações que, de repente, decidem que as coisas já não são percebidas, descritas, expressas, caracterizadas, classificadas e conhecidas da mesma maneira?

Michel Foucault

Ao longo deste capítulo, nos aprofundaremos sobre alguns aspectos do regime de saber algorítmico, tratando de pensá-los a partir de três deslocamentos que a racionalidade algorítmica e o modelo de conhecimento que lhe sustenta produz na “ordem das coisas” (Foucault, 2000). Para tanto, me debruço sobre aquilo que chamo de *três prevalências* do regime de saber algorítmico que materializam modos específicos de apreender, interpretar e produzir realidades, problemas e subjetividades, e de governar o campo do possível:

- 1) a prevalência da *performatividade* sobre a *representação*;
- 2) a prevalência da *correlação* sobre a *causalidade*;
- 3) a prevalência do *provável* sobre o *possível*.

Tais prevalências são, sobretudo, um modelo analítico que nos permitirá pôr em evidência alguns deslocamentos que o modelo de racionalidade que tem os algoritmos como centro ordenador dos modos de conhecer e governar a realidade implica — por isso, se tratam de prevalências e não necessariamente de substituições totais.

Cabe enfatizar que esses deslocamentos estão vinculados a processos históricos mais amplos e profundos (aos quais fazemos menção em alguns pontos do texto, inclusive) que ultrapassam o problema específico que estamos abordando. Além disso, é claro que eles não esgotam as características disso que chamo de regime de saber algorítmico. Tratam-se, assim, de uma escolha deliberada e direcionada sobre aspectos que me parecem significativos e relevantes para entender como essas mutações reconfiguram os modos de apreender e agir sobre o mundo, os fenômenos e os sujeitos.

2.1 Performatividade x representação

O primeiro aspecto do regime de saber algorítmico que exploraremos é a *prevalência da performatividade sobre a representação* como modo de apreender, mas também de produzir realidades. Nos últimos anos, diversos autores vêm ressaltando o caráter performativo da mediação algorítmica e da epistemologia orientada por dados (Cf. Mackenzie, 2005; Introna, 2016, 2016a; Seyfert, Roberge, 2016; Faltay, 2020; Bruno, 2021; Cesarino, 2021, 2022;). Esse caráter manifesta-se em múltiplas dimensões, algumas das quais exploraremos a seguir.

A primeira delas pode ser identificada no próprio código, a linguagem na qual a computação é realizada, como uma prática performativa (Cf. Hayles, 2005). A relação performativa com a linguagem no regime computacional encontra-se na própria matriz da Máquina de Turing e no entendimento do algoritmo como uma série de instruções. Como propõe Goffey (2008, p. 17), “algoritmos fazem coisas, e sua sintaxe incorpora uma estrutura de comando e controle para permitir que isso aconteça”⁵⁸. Conforme vimos no tópico 1.3, um dos principais dispositivos para tanto é a estrutura de seleção decisional ‘*se, então, senão*’, que regula os fluxos de decisão algorítmicas. Mesmo em situações em que os algoritmos não operam a partir de regras deste tipo, como as redes neurais, que utilizam operações matemáticas muito mais complexas, as tarefas que executam são viabilizadas por uma linguagem que ‘faz coisas’⁵⁹.

Flusser (2017), refletindo sobre a ruptura epistêmica do pós-guerra, uma época dominada por aparelhos cibernéticos e seu modo de estruturar o mundo, identifica uma mudança no modelo de linguagem predominante que estrutura a realidade. Para ele, a modernidade seria marcada por um “pensamento discursivo” (p. 326) ou um “pensamento sujeitiforme” (p. 326) que predica sujeitos em direção a objetos. Segundo o filósofo, as línguas estruturadas a partir desse modelo estabelecem ao seu redor mundos de coisas que se relacionam entre si de um modo similar à relação sujeito-objeto-predicado. A relação entre linguagem e mundo que se desdobra a partir deste modelo epistêmico tende a um entendimento causal dos fenômenos, questão sobre a qual retornaremos no próximo tópico. Em contraponto, o momento histórico

⁵⁸ No original: “algorithms do things, and their syntax embodies a command structure to enable this to happen”.

⁵⁹ Faço menção aqui à conhecida *teoria dos atos de fala* de Austin (1962). De modo sintético, a teoria se baseia na hipótese de que palavras *fazem coisas* ao invés de apenas *significá-las*. Em contraponto à função constativa, que serve para descrever o mundo das coisas e dos acontecimentos, a função performativa faz uso da linguagem como uma forma de ação e não de simples representação.

iniciado no pós-guerra, dominado pela visão de mundo cibernética, seria aquele em que a “língua do aparelho” (p. 353) passa a ser o modelo dominante sobre a realidade. “A conversação significativa da atualidade processa-se no nível do aparelho” (p. 347), afirma ele. Segundo Flusser, a língua do aparelho teria se originado de uma acentuação do verbo inglês *to do* no discurso predicativo.

Com efeito, essa língua se distingue da nossa estruturalmente pela supressão do sujeito e do objeto. É uma língua exclusivamente predicadora, nada prática (...) As sentenças do aparelho não têm sujeito nem objeto, articulam processos em funcionamento. *O aparelho pensa exclusivamente de forma funcional.* (ibid., p. 353-354, grifo nosso)

Ou seja, a ‘língua funcional’ do aparelho não tem como função significar o mundo ou representá-lo, mas *executar procedimentos*, o que o tornaria, segundo o filósofo, absurdo para nós. Em consonância com a proposição de Flusser, Cardon (2018) afirma que as regras que guiam o comportamento dos algoritmos são *procedimentais* e não *substantivas*. Eles operam num regime semiótico que podemos entender, junto à Lazzarato (2014), como *a-significante*. Em oposição às semióticas *significantes*, tais semióticas, próprias de um mundo “*maquinocêntrico*” e não mais “*logocêntrico*” (ibid., p. 56), não remetem a um sujeito enunciador e operam ‘por fora’ dos processos de significação, representação e consciência, mobilizando relações completamente distintas entre linguagem e produção de subjetividade⁶⁰. Um dos modos mais tangíveis com que esse traço se manifesta é na conhecida prevalência da *sintaxe* sobre a *semântica* das máquinas algorítmicas uma vez que mesmo grandes modelos de linguagem como o GPT são incapazes de compreender o sentido dos textos que produzem.

Para além de questões linguísticas ou semióticas, ainda que a elas diretamente relacionadas, a prevalência da performatividade sobre a representação no regime de saber algorítmico diz respeito àquilo que Bruno (2021) entende como a principal *força epistêmica* dessa racionalidade. Para a autora, nesse regime, mais do que a descrição de uma realidade pré-existente, como num modelo de conhecimento baseado na representação, importa sobretudo a *produção de realidades e subjetividades*. Em suas palavras:

⁶⁰ Nessa proposição, os dois regimes semióticos se relacionam a dois vetores distintos de subjetivação cujo duplo investimento estaria na base da eficácia e poder capitalistas: a *sujeição social* e a *servidão maquínica*. A sujeição social se relaciona com a formação de uma *subjetividade individual*, “ao nos atribuir uma identidade, um sexo, uma profissão, uma nacionalidade e assim por diante” (Lazzarato, 2014, p.27), produzindo-nos como sujeito individuados. Já a servidão maquínica remete a técnicas não representativas, operacionais, diagramáticas que exploram e produzem subjetividades parciais, modulares e sub e supra individuais. Essa dimensão age sobre os níveis *pré-pessoais* e *suprapessoais* da subjetividade, visando e produzindo aquilo que Deleuze chamou de *dividual*. Enquanto a sujeição social mobiliza semióticas significantes, em particular a linguagem, envolvendo a consciência e representações que visam um sujeito individuado, a servidão maquínica funciona com base em semióticas a-significantes (índices de mercado de ações, moeda, equações matemáticas, diagramas, linguagens computacionais, etc) que não envolvem a consciência.

[o] tipo de conhecimento gerado por processos algorítmicos está em muitos casos em ruptura com os modelos representacionais de conhecimento, cuja força epistêmica seria a capacidade de descrever ou compreender uma realidade ou fenômeno dado, mantendo-se o mais fiel possível a um referente qualquer. *A força da racionalidade performativa dos algoritmos não está em descrever ou representar, mas sim em gerar efeitos e produzir realidades.* (ibid., p. 156, grifo nosso)

Identificando um traço similar, Introna enfatiza como o desenvolvimento de algoritmos não tem a ver simplesmente com a execução de instruções definidas por programadores, uma vez que eles “também encenam os objetos que supostamente refletem ou expressam”⁶¹ (Introna, 2016, p. 4). Dito de outro modo, algoritmos não são entidades primordialmente *descritivas*, mas sobretudo *performativas*. Conseqüentemente, eles simultaneamente ‘descrevem’ e criam as realidades que visam apreender (Chun, 2018) — ainda que isso soe paradoxal com a ‘aura’ de objetividade e neutralidade ressaltada por seus promotores para legitimar seus usos.

Tais entendimentos se alinham a perspectivas teóricas que enfatizam a dimensão performativa dos artefatos, sobretudo nos Estudos sobre Ciência, Tecnologia e Sociedade (CTS) (Cf. Latour, 2012; Callon, 1998; Pickering, 1995; Law, Singleton, 2000), como um contraponto às abordagens que os enquadram como meros depositários da agência e excepcionalidade humana. Grosso modo, segundo tais perspectivas, não é possível cindir a realidade, ou os fenômenos que a compõem, dos instrumentos (materiais e epistemológicos) utilizados para apreendê-la, observá-la, registrá-la, classificá-la. Segundo Karen Barad (2007, 2020), que tem proposto um entendimento radical da performatividade no processo de constituição da própria matéria, “medições são práticas agenciais que não são apenas revelatórias, mas performativas: elas ajudam a constituir e são uma parte constitutiva do que está sendo medido”⁶² (Barad, 2020). Por isso, “medições são criadoras de mundo: matéria e significado não preexistem, mas são co-constituídos por meio de intra-ações de medição”⁶³ (ibid.). Inspirando-nos em tais concepções, podemos afirmar que as práticas ligadas à datificação e à algoritmização são *práticas criadoras de mundos (world-making practices ou worlding practices)* (Barad, op. cit.; Haraway, 2008, 2016). Por isso, também, análises e leituras meramente indexicais dos sistemas algorítmicos (ou seja, entendendo-os como meros

⁶¹ No original: “also enact the objects they are supposed to reflect or express”.

⁶² No original: “Measurements are agential practices, which are not simply revelatory but performative: they help constitute and are a constitutive part of what is being measured.”

⁶³ No original: “Measurements are world-making: matter and meaning do not preexist, but rather are co-constituted via measurement *intra-actions*.”

reflexos passivos de seus criadores) não dão conta da complexidade desses modos de existência técnicos e tampouco de seus efeitos sociais, culturais e subjetivos.

Ainda que as IAs Generativas materializem essa performatividade de um modo praticamente literal, esse traço se manifesta de um modo transversal nos mundos criados pelas práticas e técnicas algorítmicas: nos supostos poderes oraculares atribuídos a tais tecnologias, nos problemas e sujeitos visados por tais tecnologias (Hong, 2022), nas políticas da classificação que ordenam os bancos de dados de treinamento e prescreve modos de segmentação da realidade (Crawford, 2021), e assim por diante.

Em casos como a previsão de orientação sexual a partir de imagens faciais⁶⁴ (Kosinski, Wang, 2018), a previsão de risco de reincidência criminal por sistemas como o COMPAS⁶⁵ (Dressel, Farid, 2018), o paradigmático caso da Cambridge Analytica e a inferência de personalidades a partir de perfis psicométricos (Hern, 2018), a previsão do potencial de candidatos à vagas de emprego a partir da análise de expressões faciais e entonação de voz em plataformas como o HireVue⁶⁶, ou mesmo em práticas aparentemente banais como a segmentação de personalidade segundo preferências musicais em sistemas recomendação como o Spotify (Anderson et al, 2020), o fato das previsões serem verdadeiras ou falsas importa menos do que os efeitos que desencadeiam sobre os contextos em que atuam.

Por isso, o fato de muitas das classificações, predições e recomendações feitas por essa ‘maquinaria adivinhatória’ serem falhas ou espúrias, para utilizar um termo da ciência de dados, não significa que não gerem *efeitos que ajudam a produzir as condições de sua própria verdade*. Com isso, não quero dizer que é indiferente o fato de uma predição ser mais ou menos acurada, sobretudo em campos como a segurança pública ou a justiça, uma vez

⁶⁴ Desenvolvida por Michal Kosinski, um dos criadores do teste de personalidade que esteve na base do escândalo da Cambridge Analytica, e Yilun Wang, a pesquisa fez uso de redes neurais para extrair informações de 35.326 imagens faciais e construir um modelo preditivo classificador de orientação sexual. Segundo os dados divulgados, o classificador teria uma acurácia de 81% para homens (gays e heterossexuais) e 74% para mulheres com apenas uma foto, chegando a 91% e 83% para homens e mulheres, respectivamente, com cinco imagens faciais por pessoa. Alinhando-se a teorias eugênicas como a teoria hormonal pré-natal, os pesquisadores afirmam que “homens e mulheres gays tendem a ter morfologia facial, expressão e estilos de aparência atípicos de gênero.” (Kosinski, Wang, 2018).

⁶⁵ A ferramenta, cuja sigla significa Perfil de Gestão Correccional de Infratores para Sanções Alternativas (*Correctional Offender Management Profiling for Alternative Sanctions*, em inglês) desenvolvida em 1998 pela empresa Northpointe (hoje *equivant*) é utilizada pela corte estadunidense para avaliar a probabilidade de um réu se tornar reincidente. Segundo Dressel e Farid (2018), o componente de previsão de reincidência do COMPAS está em uso desde 2000. O software prevê o risco de um réu cometer crimes num prazo de dois anos a partir da avaliação de 137 características sobre o indivíduo e seus comportamentos passados.

⁶⁶ <https://www.hirevue.com/>

que já é amplamente sabido que esses sistemas atualizam algorítmicamente violências históricas exercidas sobre corpos racializados intensificando injustiças históricas⁶⁷. Falhas, sobretudo em contextos em que a dimensão biopolítica (Foucault, 2008b) da algoritmização é mais acentuada, são espaços de disputa e devem ser evidenciadas, problematizadas e politizadas. Afinal, como nos lembra Chun (2018), a defasagem entre predição e realidade é um espaço crucial para a agência e ação política. O que estou argumentando, no entanto, é que a força do saber-poder algorítmico não reside propriamente na acurácia das predições que lhe servem de instrumento, mas no modo como tais predições criam as realidades que dizem interpretar e legitimam ações antecipatórias no presente (Hong, 2022).

O simples fato dos idealizadores, desenvolvedores e promotores desses sistemas ‘enquadrarem’ o comportamento humano não somente como previsível mas cuja previsibilidade é melhor reconhecida e interpretada por dados e algoritmos do que por outros métodos, ajuda a produzir, performativamente, as subjetividades supostamente capturadas por estas tecnologias, corroborando suas virtudes oraculares. Identificando uma dinâmica similar ao descrever as convergências entre o populismo, a digitalização e o neoliberalismo, que exploram em suas estratégias efeitos essencialmente performativos, Cesarino (2021, p. 86) afirma que “[s]ão linguagens que não descrevem uma realidade preexistente, mas geram efeitos sobre, e coproduzem, os sujeitos que comunicam”.

Pelas razões que viemos expondo, como vários autores têm argumentado, predições algorítmicas e seus modos de interpretar o mundo se aproximam daquilo que o sociólogo Robert Merton (1968) designou “profecias autorrealizadoras”: um prognóstico que, ao se tornar uma crença, provoca sua própria concretização (Cf. Healy, 2015; Chun, 2018; Faltay, 2020; Birhane, 2021; Perdomo et al, 2021; Hong, 2022). Essa dinâmica também se reflete nos próprios problemas para os quais essas tecnologias se anunciam como solução. Como afirmam Roberge e Seyfert,

[e]nquanto a ciência da computação define algoritmos como procedimentos ou receitas para resolver problemas, abordagens como a sociologia cultural enfatizam seus efeitos performativos, suas funções recursivas pelas quais as práticas

⁶⁷ Uma investigação da ProPublica em 2016 sobre o sistema de avaliação de risco COMPAS, por exemplo, revelou que negros têm quase o dobro de chances de serem classificados como de alto risco em comparação com brancos, mas, prática, não reincide. Por outro lado, brancos são mais frequentemente classificados como de baixo risco, mas continuam cometendo crimes. Além disso, a investigação mostrou que apenas 20% das pessoas previstas para cometer crimes violentos realmente o fizeram (Angwin et al, 2016).

algorítmicas não apenas criam novos problemas, mas também *criam os problemas para os quais são, em última análise, a resposta*.⁶⁸ (2016, p. 20, grifo nosso)

Num sentido similar, Hong (2022), afirma que o uso de técnicas preditivas orientadas por dados como solução para uma ampla variedade de problemas sociais, da criminalidade à produtividade, produz problemas tangíveis em torno dos quais provas aparentemente concretas de automação e inteligência artificial podem ser produzidas. Segundo ele, o uso do Big Data e IA aplicados a problemas sociais tende a invisibilizar dados que não podem ser facilmente capturados e julgamentos que não podem ser racionalizados pelos dados disponíveis. “A datificação nunca conta apenas o que já está lá, como nos mitos do positivismo ingênuo; em vez disso, *ela cria certos mundos e extingue outros*. Criminalidade, produtividade ou emoção são identificadas como uma coisa discreta a ser medida e, subsequentemente, otimizada.”⁶⁹ (Hong, 2022, p. 375-376, grifo nosso).

Um dos problemas elementares da performatividade das previsões algorítmicas é o fato de que, com frequência, elas desencadeiam mudanças na distribuição da variável que os sistemas objetivam prever (Perdomo et al, 2021) de modo recursivo. Assim, previsões de trânsito influenciam os padrões de trânsito, previsões de crimes influenciam as alocações da polícia, recomendações algorítmicas influenciam padrões de consumo, previsões de preço das ações modificam o mercado, e assim por diante.

No caso de sistemas baseados na previsão de riscos como o policiamento preditivo ou reincidência criminal, os efeitos performativos tornam-se ainda mais radicais uma vez que a ideia é que não se ‘pague para ver’ se as previsões são confiáveis, mas que elas disparem ações antecipatórias que evitem que aquele futuro previsto ocorra, instaurando uma “gestão de possíveis” (Bruno, 2012, p. 59) que remete a ficções como *Minority Report* (2002).

Um aspecto crucial do caráter performativo dos algoritmos e de seus efeitos reside na lógica operacional e estrutura temporal da *recursividade*, sobre a qual nos deteremos no Capítulo 3 ao explorarmos a matriz cronopolítica da cibernética. Ao introduzir uma estrutura temporal

⁶⁸ No original: “Where computer science defines algorithms as procedures or recipes for solving problems, approaches such as cultural sociology emphasize their performative effects, their recursive functions by which algorithmic practices not only create new problems, but also create the problems for which they are ultimately the answer.”

⁶⁹ No original: “Datafication never just counts what is already there, as in the myths of naive positivism; rather, it creates certain worlds and extinguishes others. Criminality, productivity, or emotion is identified as a discrete thing to be measured and, subsequently, to be optimized”.

não mais baseada numa forma linear, mas numa espiral (Hui, 2019), a recursividade é o dispositivo central que produz os efeitos sistêmicos e amplificadores da mediação algorítmica nos ecossistemas digitais, das ‘câmaras de eco’ à repetição dos padrões contidos nos bancos de dados de treinamento dos modelos algorítmicos. A hermenêutica recursiva também supõe uma nova relação com o *erro* materializada no mecanismo do *feedback*, ou seja, a retroalimentação dos outputs de erro no aprendizado do sistema (Morozov, 2018). Uma vez que as falhas e erros são recursivamente incorporados para o refinamento dos modelos algorítmicos, a noção de verdade (ou falsidade) não é mais uma categoria propriamente adequada aqui. Nesse paradigma, “cada erro é produtivo”⁷⁰ (Luhmann apud Hui, p. 20) para o aprendizado das máquinas. Dito de outro modo, todos os *outputs* são ‘verdadeiros’ uma vez que do ponto de vista estritamente informacional modulam o aperfeiçoamento dos sistemas.

O predomínio da performatividade sobre a representação que estamos ressaltando aqui pode também ser interpretado como um dos efeitos de uma crise da representação num contexto mais amplo e profundo. Para Rouvroy, o fenômeno da digitalização e da algoritmização se relacionam a uma crise radical da representação em si mesma (Rouvroy, Stiegler, 2016). Por isso, para a autora, a ascensão da governamentalidade algorítmica não coincidiria com a emergência de outro regime de verdade, mas numa crise nos próprios regimes da verdade de forma mais ampla. Para a autora, o conceito de verdade está cada vez mais sendo substituído pela “realidade pura ou atualidade pura, a ponto de, eventualmente, as coisas parecerem *falar por si mesmas*.”⁷¹ (ibid., grifo nosso, p. 7). Essa crise resultaria numa busca pela objetividade absoluta, encontrada, por exemplo, na autoevidência dos dados que “falam por si mesmos” (Anderson, 2018).

A ruptura apontada por Rouvroy pode parecer contraditória com o argumento que viemos desenvolvendo ao longo deste tópico. Mas é justamente na ‘dobra’ entre um modelo de saber que opera por meio de efeitos performativos e funções recursivas, e um suposto ‘realismo algorítmico’ que se anuncia como capaz de captar o real ‘nele mesmo’, dispensando representações (questão sobre a qual nos aprofundaremos a seguir), que a racionalidade algorítmica se situa e exerce seu saber-poder. Ou seja, quanto mais se declara como capaz de

⁷⁰ No original: “every error is productive”.

⁷¹ No original: “pure reality or pure actuality, to the extent that eventually things seem to be speaking by themselves.”

produzir conhecimentos ‘diretamente’ no mundo, sem a mediação de representações, mais sua performatividade tem poder e legitimidade para produzir mundos.

2.2 Correlação x causalidade

O segundo aspecto do regime de saber algorítmico que exploraremos é a predominância da *correlação* sobre a *causalidade*⁷² como modo de construção de evidências e entendimento dos fenômenos. Esse aspecto está diretamente ligado à importância da ciência de dados no caldo epistemológico do regime de saber algorítmico e à hegemonia dos modelos indutivos sobre os modelos dedutivos de raciocínio no campo da IA que vimos no Capítulo 1. Essa prevalência foi enfatizada por várias autoras (Rouvroy, Stiegler, 2016; Zuboff, 2021; Chun, 2021; Mcquillan, 2016; Birhane, 2021), mas talvez ainda seja o paradigmático texto de Anderson (2008) sobre o “fim da teoria” mencionado na Introdução a melhor síntese da substituição de um modelo de conhecimento dedutivo (baseado em hipóteses e modelos a priori) pela mera indução dos fatos que ‘emergem’ dos dados através das correlações:

Os petabytes permitem-nos dizer: “A correlação é o suficiente.” Podemos parar de procurar modelos. Podemos analisar os dados sem hipóteses sobre o que eles podem mostrar. Podemos jogar os números nos maiores clusters de computação que o mundo já viu e permitir que algoritmos estatísticos encontrem padrões onde a ciência não consegue⁷³. (Anderson, 2008, grifo nosso)

A diferenciação entre correlação e causalidade é um princípio fundamental da ciência de dados e pode ser resumida no seguinte mantra: *correlação não implica causalidade*. Como consta em um material comercial da empresa SAP sobre a Análise Preditiva:

Os algoritmos usam o aprendizado de máquina para identificar e medir quais preditores estão altamente correlacionados com os alvos que você está tentando prever. Em outras palavras, eles tentam identificar preditores que parecem ter um forte relacionamento ou conexão com os resultados. Isso não é o mesmo que identificar preditores como sendo a causa de resultados específicos⁷⁴. (SAP, 2017, p. 10)

⁷² Uma observação importante é que a causalidade que abordamos neste tópico se refere à causalidade linear, uma vez que, como veremos no tópico 3.1, os sistemas algorítmicos são regidos por causalidades circulares ou recursivas.

⁷³ No original: “Petabytes allow us to say: ‘Correlation is enough.’ We can stop looking for models. We can analyze the data without hypotheses about what it might show. We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot.”

⁷⁴ No original: “Algorithms use machine learning to identify and measure which predictors are *highly correlated* with the targets that you are trying to predict. In other words, they try to identify predictors that appear to have a strong relationship or connection with outcomes. This is not the same as identifying predictors as actually being the cause of specific outcomes.”

Algoritmos de aprendizado de máquina são ótimos para estabelecer a dependência entre duas variáveis, mas não para explicar por que essa dependência existe. A relação causa e efeito, no entanto, não é necessária para prever a probabilidade de ocorrência dessa correlação. Desde que ela seja estável, ou seja, expresse um padrão, pode ser utilizada para fazer previsões estatísticas mesmo que não seja possível compreender por que os elementos se comportam enquanto tal. É por isso que Anderson afirma que “[a] correlação substitui a causalidade e a ciência pode avançar mesmo sem modelos coerentes, teorias unificadas, ou realmente qualquer mecanismo de explicação”⁷⁵ (op. cit., grifo nosso).

Desse modo, mais do que a *impossibilidade* de acessar as causas, talvez seja a *indiferença* a essa dimensão uma das marcas deste regime de saber, uma vez que este conhecimento não é considerado relevante para conhecer e intervir sobre realidades, eventos ou comportamentos dos sujeitos. Recorrendo novamente a Anderson:

É um mundo no qual as quantidades massivas de dados e as matemáticas aplicadas substituem todas as outras ferramentas que poderiam ser utilizadas. Adeus a todas as teorias sobre os comportamentos humanos, da linguística à sociologia. Esqueçam a taxonomia, a ontologia e a psicologia. *Quem pode saber por que as pessoas fazem o que fazem? O fato é que o fazem e que nós podemos traçá-lo e medi-lo com uma fidelidade sem precedente. Com dados suficientes, os números falam por si só.*⁷⁶ (Anderson, 2008, grifo nosso).

Ou seja, nesse modelo de conhecimento não importa saber a razão pela qual os sujeitos fazem algo, apenas que o fazem e quais correlações são possíveis de mensurar, extrair e posteriormente monetizar sobre tais ações, comportamentos e fenômenos. Isso porque um dos princípios fundamentais do capitalismo de dados é que “o valor de cada pedaço de dado não está contido em si mesmo, mas é essencialmente relacional” (Rouvroy, Athanasiadou, Klumbyté, 2022, p. 131). Assim, são as correlações descobertas entre os dados que lhes dão sentido, utilidade e valor.

O modo mais evidente com que essa “indiferença radical” (Zuboff, 2021) às causas se manifesta cotidianamente é no conhecido fato de que, sob a convergência da plataformação com a economia da atenção, pouco importa se o que levou um usuário a se engajar com determinado conteúdo nas redes foi amor ou ódio; o que importa são as métricas e

⁷⁵ No original: “Correlation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all.”

⁷⁶ No original: “This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves.”

performances (Cf. Cesarino, 2022). Ainda assim (ou, justo por isso), sabe-se que hoje o ódio ‘engaja’ bem mais do que amor.

Mas é claro que o simplista ‘fim da teoria’ anunciado efusivamente por Anderson corresponde apenas à obsolescência nesse regime do método científico baseado em causalidades, uma vez que é evidente que há todo um novo aparato teórico que sustenta esse novo modo de conhecer e agir sobre o mundo. Afinal, enunciar-se como ‘não teórico’ é apenas mais um modo de promover as virtudes de objetividade e neutralidade desse paradigma. Como afirma Mbembe (2021, p. 93), “por trás de cada estatística e de cada dado e de cada algoritmo, existe, de fato, implícita ou explicitamente, uma hipótese, uma teoria que não diz seu nome”. A ideia de que não é mais relevante saber ‘por que as pessoas fazem o que fazem’, por exemplo, guarda estreita relação com o entendimento da Economia Comportamental de que o *homo sapiens* é um processador falho de informações que muitas vezes age em desacordo com seus próprios interesses (Cf. Ariely, 2020). Diante dessa previsível irracionalidade, o preditor mais confiável do comportamento futuro, portanto, não são as motivações dos sujeitos, mas seu comportamento pregresso.

Mas longe de ser um fenômeno eminentemente recente, o deslocamento que estamos tematizando aqui pode ser visto como mais um desdobramento das rupturas epistemológicas do pós-guerra. Para Flusser (2017, 2019), diferente dos pensamentos *finalistas* e *causalistas*, que dominaram o pensamento moderno e histórico, o pensamento e a sociedade dominados pela cibernética, seriam *programáticos*. Segundo ele, a centralidade da noção de *programa* nesse novo modelo de interpretação e ação sobre o mundo, expressa no raciocínio *funcional* dos aparelhos cibernéticos, poria em crise os modelos causalistas e finalistas de interpretação dos fenômenos. Fazendo uma síntese dessa transição, ele afirma:

A noção da programação da existência humana e do mundo é relativamente nova. Numerosos aspectos implícitos em tal noção ainda não foram conscientizados. A nossa herança mítica nos habituou à noção de uma existência e de um mundo regidos pelo destino, as ciências da natureza despertaram em nós a noção de uma existência e de um mundo regidos pela causalidade. A atualidade exige que repensemos tais noções, a do destino, a da causalidade e a do programa. (Flusser, 2019, p. 25)

Diante dessa mutação, para ele, a ciência se veria “obrigada a abrir mão das explicações causais em numerosos campos, e de substituí-las pelas *formais*” (ibid., p. 47). Isso se manifestaria de modo especialmente agudo no campo da psicologia, no qual “[p]erguntas do tipo ‘por que os homens se comportam assim e não diferentemente?’ vão sendo substituídas por perguntas do tipo: ‘como se comportam os homens?’” (ibid.).

Possivelmente nenhuma vertente da psicologia sintetiza tão bem a transição descrita por Flusser quanto o behaviorismo, matriz fundamental para compreender tanto a indiferença às causas quanto a prevalência de interpretações baseadas em correlações e que, como demonstrou Zuboff (2021) está na base da arquitetura de modificação comportamental que o capitalismo de vigilância instrumentaliza. O modelo behaviorista de sujeito substitui uma “engenharia das almas” por uma “engenharia do comportamento” (Zuboff, 2021, p. 427) reduzindo a experiência humana a comportamentos observáveis e mensuráveis. De modo simplificado, o chamado behaviorismo radical de Skinner⁷⁷ está interessado em correlações entre um estímulo e suas prováveis respostas (o que na gramática algorítmica poderia traduzido para o constructo ‘se, então’): “‘se’ há a presença de um determinado estímulo ‘então provavelmente’ haverá uma certa resposta comportamental e, ainda, o reforço desta ação aumentaria a probabilidade de sua recorrência no futuro” (Bentes, 2022, p. 158). Esta abordagem da psicologia, surgida no começo do séc. XX, mostra-se indiferente às motivações ou interioridades do sujeito uma vez que considera “apenas aqueles fatos que podem ser observados de modo objetivo no comportamento de uma pessoa em relação à sua história comportamental progressa” (Skinner apud Zuboff, op. cit., p. 414). Diferentemente de abordagens da psicologia que visam um objeto causa, como a psicanálise, o que importa aqui é saber como o comportamento — entendido como qualquer modificação de um organismo, humano ou não-humano, em relação ao seu ambiente — responde a determinados estímulos e pode ser condicionado e previsto por meio de uma série de “programas de reforço” (Zuboff, op. cit., p. 412). O ser humano é concebido aqui como um mero “organismo entre organismos” (ibid., p. 414), pura alteridade exteriorizada. Como esclarece o próprio Skinner, trata-se aqui do “estudo da vida do Outro — mas da vida apenas quanto sua significação social, e não quanto sua significação para si mesmo” (apud Zuboff, op. cit., p. 414). Por isso, para a ciência do “comportamento operante” (ibid., p. 417), instrumentalizado e atualizado pelas plataformas digitais para produzir técnicas de influência, persuasão e condicionamento de hábitos dos usuários, derivando no que Bentes (2022) chama de *modelo tecnobehaviorista*, aquilo que não pode ser observado e mensurado, é irrelevante.

⁷⁷ Tomando como objeto de estudo *o comportamento*, o behaviorismo se estabeleceu ao longo do séc. XX como um conjunto de abordagens que se propõe a torná-lo observável, mensurável e modificável. O maior expoente desta controversa vertente da psicologia é B. F. Skinner, que propôs as premissas e técnicas do chamado behaviorismo radical. Partindo do modelo do behaviorismo clássico de Watson e Pavlov, ancorado no paradigma do *estímulo e resposta*, mas somando-lhe o princípio do *reforço*, Skinner elabora as bases de uma tecnologia de condicionamento dos comportamentos visando a aplicação de métodos de modificação comportamental a nível individual e populacional.

Rouvroy também encontrará ressonâncias do behaviorismo na “nova forma de produzir conhecimento sobre preferências, atitudes, comportamentos ou eventos futuros sem considerar as motivações psicológicas, discursos ou narrativas do sujeito, mas sim baseando-se em dados”⁷⁸ (2013, p. 1), chamando-o *behaviorismo de dados*. Neste modelo, “o conhecimento não é mais produzido *sobre* o mundo, mas *a partir* do mundo digital” (ibid. p. 4, grifo da autora).

A centralidade da correlação e dos métodos indutivos no regime de saber algorítmico implica também numa outra *normatividade*, como a autora aponta em sua teorização posterior sobre a governamentalidade algorítmica com Berns (Rouvroy, Berns, 2015). Uma vez que com o Big Data e as máquinas conexionistas não é mais necessário elaborar hipóteses para serem validadas, os autores sustentam que esse novo modelo de produção de conhecimento, em contraponto às práticas estatísticas modernas ancoradas sobre a *média* ou o *normal*, é um modelo “*a-normativo*” (ibid., p. 37). Segundo eles, trata-se de “um sistema de relações eminentemente evolutivas, entre diversas medidas, irreduzíveis a qualquer média.” (ibid., p. 36) que substitui formas *a priori* por *emergências* (de padrões, métricas, perfis, etc) detectadas e refinadas em tempo real. Para os autores, esta seria mais uma razão pela qual a governamentalidade algorítmica se afastaria de intencionalidades subjetivas, delegando à semiose dos dados e algoritmos a geração e interpretação automática de conhecimento sem deixar-se intervir pela percepção, imaginação e entendimento humanos⁷⁹.

Um dos pressupostos que guia a adoção desse novo método de conhecer, interpretar e prever fenômenos é bem sintetizado por Bruno (2021) ao refletir sobre a dimensão laboratorial das plataformas: “porque devo elaborar uma teoria sobre o mundo (ou sobre qualquer fenômeno) e testá-la em condições controladas, se posso testar diretamente sobre o mundo?”. Quando tanta coisa pode ser testada, observada e transformada em métricas, “os debates teóricos e filosóficos exaustivos só atrapalham” (Morozov, 2018, p. 108). Outro fator elementar é, claro, econômico: não há motivos para ‘perder tempo e dinheiro’ elaborando, validando e/ou refutando hipóteses se é possível simplesmente ‘deixar que as máquinas trabalhem’ e esperar os valiosos *insights* surgirem, em sintonia com todas as metodologias de desenvolvimento

⁷⁸ No original: “new way of producing knowledge about future preferences attitudes, behaviours or events without considering the subject’s psychological motivations, speeches or narratives, but rather relying on data”.

⁷⁹ Ainda que abordem um traço importante da governamentalidade algorítmica, é legítimo também observar os autores não problematizam de modo suficientemente crítico a questão da objetividade, uma vez que é evidente que existe toda uma falácia da objetividade, neutralidade e não-mediação que sustenta esse novo paradigma e que nutre o que chamamos de sua força epistêmica performativa.

ágil e o imperativo da otimização que move o Vale do Silício. Como notam Mayer-Schönberger e Cukier (2013), a busca pela causalidade vem se tornando uma ‘velha questão’ numa era mais interessada no ‘como, quando, dependendo do quê’ e na qual não há tempo para descobrir as causas ‘por trás’ das correlações.

Uma das grandes vantagens desse modelo é a possibilidade de prever atributos latentes sem a necessidade de ter acesso direto aos dados visados. Segundo essa nova ciência de dados, os ‘melhores preditores’ não são aqueles explicitamente relacionados ao traço que se deseja mapear (como ‘curtir’ uma página diretamente relacionada a esse atributo), mas inferências feitas a partir de correlações pouco óbvias, que a ‘velha ciência’ ou nossa ‘vã percepção’ jamais detectariam, como entre *curtir Hello Kitty* e a *propensão a ser democrata*, ou entre *ficar confuso depois de acordar de cochilos* e a *propensão à homossexualidade masculina*. Essas correlações integraram um estudo sobre perfis psicométricos a partir da análise de ‘curtidas’ no Facebook que foi base para o caso Cambridge Analytica⁸⁰ (Kosinski, Stillwell, Graepel, 2013).

O pressuposto que guia esse desconcertante oráculo é de que os dados rastreados a partir de ações que escapam à consciência são mais confiáveis e reveladores sobre os comportamentos, inclinações e vulnerabilidades dos sujeitos em suma, sobre sua verdade, do que os conscientes; ideia que em si mesma não é nova e manifesta ressonâncias inclusive do conceito de inconsciente psicanalítico, mas que ganha outros contornos e implicações no contexto da exploração de uma Economia Psíquica dos Algoritmos (Bruno, 2018; Bruno, Bentes, Faltay, 2019) por corporações de tecnologia e grupos políticos marcadamente antidemocráticos para previsão e modificação de comportamentos.

Simultaneamente, essa abordagem dá margem para que uma série de *correlações espúrias* — relações que não passam de meras ocorrências aleatórias entre duas variáveis e cuja associação é puramente arbitrária⁸¹ — sejam consideradas preditivas, principalmente sobre

⁸⁰ Essas correlações apareceram no estudo conduzido por Kosinski, Stillwell, Graepel (2013) que foi uma das bases para os experimentos do paradigmático caso da *Cambridge Analytica*. O estudo afirmava ser possível inferir atributos como orientação sexual, etnia, visões religiosas e políticas, traços de personalidade, inteligência, felicidade, uso de substâncias viciantes, separação parental, idade e gênero a partir de curtidas do Facebook.

⁸¹ Uma correlação espúria, em estatística e ciência de dados, é uma relação aparente entre duas ou mais variáveis que surge devido a um fator externo ou ao acaso. Apesar de ser relativamente comum encontrar a definição de que correlações espúrias são aquelas em que se estabelece uma *falsa relação causal* entre variáveis, como bem pontua Hildebrandt (2008), correlações não espúrias não necessariamente estabelecem uma relação causal entre variáveis (afinal, correlação não implica causalidade). Muitas vezes uma correlação espúria deriva de uma variável oculta que influencia as variáveis correlacionadas.

personalidades e comportamentos. Em meio ao “dilúvio dos dados” (Anderson, op. cit.) e à infinidade de correlações possíveis em bancos de dados cada vez mais volumosos, não só torna-se difícil distinguir quais correlação são realmente consistentes, como existem interesses comerciais e políticos nessa confusão de fronteiras. Muitas das aplicações das ciências de dados e seus métodos de previsão de personalidade “apresentam perigosas tendências em elaborar sentido identificando padrões de regularidade onde há apenas ruído.” (Faltay, op. cit., p. 124). Além disso, como pontuamos no tópico anterior, o fato de algumas (ou mesmo a maior parte) das correlações serem falhas ou espúrias não significa de modo algum que não produzam as condições de suas próprias verdade inclusive porque, assim como a maior parte das predições algorítmicas, sua factualidade permanece quase sempre inverificável. Talvez o caso seja justamente o contrário. Uma vez que a lucratividade dos dados depende de fatores como encontrar correlações irreconhecíveis nos esquemas interpretativos existentes e recombina dados livremente para novas análise (Hong, 2022), quanto mais ‘inusitadas’ (e provavelmente espúrias) forem essas correlações, mais lucrativas tendem a ser. O projeto *Spurious Correlations*, mantido por Tyler Vigen, faz uma sátira bem humorada sobre correlações espúrias ao mesmo tempo em que explica como elas são construídas (Figura 3).

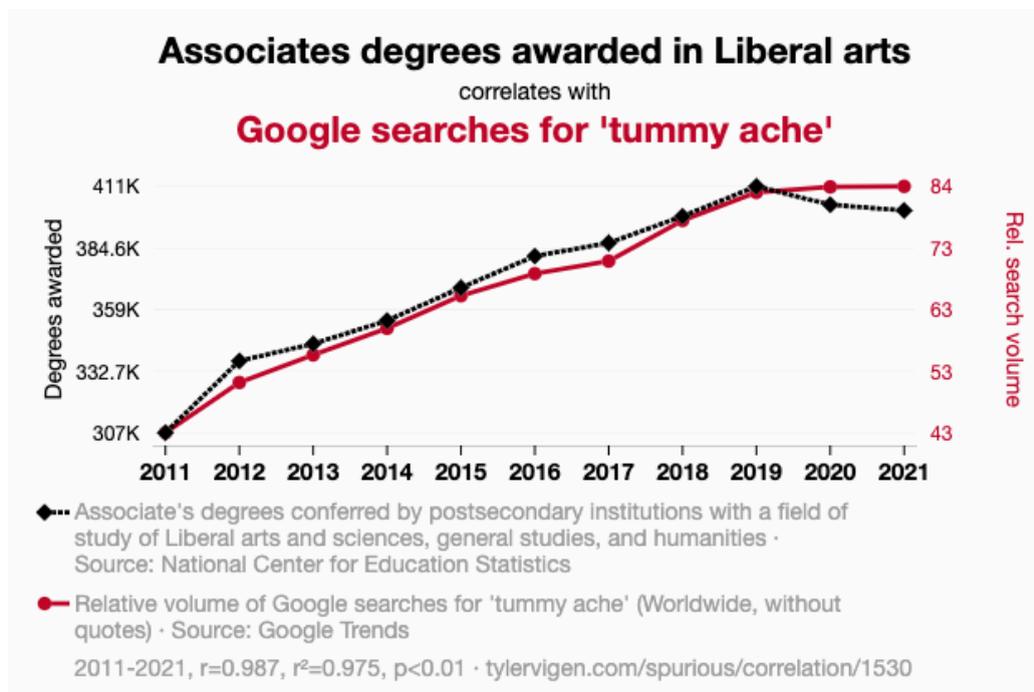


Figura 3: Gráfico da correlação espúria entre diplomas concedidos em artes liberais x pesquisas no Google por 'dor de barriga'. Fonte: Tyler Vigen⁸².

A abordagem da experimentação ‘direta’ sobre o mundo que os métodos indutivos incitam e que o economista-chefe da Google Hal Varian (2014) vislumbrou como uma oportunidade de experimentação contínua, revela também como “a racionalidade algorítmica implica a transformação do mundo num imenso laboratório” (Bruno, 2021). Capitanado pelas grandes plataformas digitais, esse modelo simultaneamente econômico, epistemológico e de controle comportamental, que Bruno, Bentes e Faltay propõem designar “laboratório de plataforma” (2019, p. 14), transforma o espaço digital num poderoso laboratório que direciona imensos volumes de dados para estratégias de previsão e indução de comportamento no qual os usuários são submetidos a contínuos testes e intervenções como “cobaias involuntárias” (Morozov 2018, p. 105). Os experimentos realizados neste *laboratório-mundo* (Bruno, 2018) vão dos simples, mas massivamente aplicados, testes A/B⁸³ até experimentos como o do do ‘contágio emocional’⁸⁴ e o do comparecimento às urnas realizados no Facebook⁸⁵.

82

https://www.tylervigen.com/spurious/correlation/1530_associates-degrees-awarded-in-liberal-arts_correlates-with-google-searches-for-tummy-ache.

⁸³ Um teste A/B é uma técnica de experimentação utilizada para comparar duas versões de uma variável (A e B) com o objetivo de determinar qual é mais eficaz em alcançar uma meta específica. Isso envolve criar as duas versões, dividir o público-alvo em dois grupos aleatórios, coletar dados sobre o desempenho de cada versão, e analisar esses dados para decidir qual versão tem melhor desempenho. Essa prática é amplamente utilizada em áreas como marketing digital, desenvolvimento de produtos e design de interface de usuário.

⁸⁴ O experimento realizado na rede social Facebook em janeiro 2012 manipulou o feed de 689 mil usuários, sem o conhecimento ou consentimento deles, com o propósito de analisar se o humor ou estado emocional dos usuários seria ‘contaminado’ pelo conteúdo visualizado no feed. No estudo, os usuários foram divididos em dois grupos: um foi exposto a postagens de conteúdo emocionalmente positivo e o outro a conteúdos emocionalmente negativos ao longo de uma semana. De acordo com artigo publicado pelo estudo, o resultado teria confirmado a hipótese de contágio emocional. Monitorando os conteúdos dos posts dos usuários com palavras emocionalmente positivas ou negativas, os pesquisadores concluíram que eles tinham maior propensão a mudarem de humor de acordo com os conteúdos que visualizavam, caracterizando um “contágio emocional”.

⁸⁵ No dia da eleição de 2010 para o Congresso dos Estados Unidos, o Facebook testou um novo recurso em que o usuário poderia compartilhar se havia ido votar em seu feed de notícias. A ferramenta, entretanto, não funcionava para todos da mesma forma, como apontou uma pesquisa publicada pela revista Nature em 2012. Para 1% dos usuários aparecia uma mensagem "informativa" no topo da timeline, estimulando o voto, informando locais de votação e um botão escrito "eu votei" para clicar. Para 98% das pessoas (aproximadamente 60 milhões) havia uma mensagem similar a essa, mas que incluía também a foto de até seis amigos que afirmavam terem votado. De acordo com a pesquisa, o experimento resultou num aumento no comparecimento de aproximadamente 340 mil pessoas, que teriam sido estimuladas a votar influenciados pela imagem dos amigos que afirmaram terem ido às urnas. O Facebook manteve o recurso nas eleições seguintes (2012, 2014, 2016 e 2018).

Esse laboratório-mundo substitui “os muros dos tradicionais e purificados laboratórios científicos” (ibid., p. 14) — que tinha como uma de suas finalidades a validação e/ou refutação de hipóteses e modelos previamente elaborados segundo métodos consolidados — por “caixas pretas digitais bastante opacas” (ibid.) completamente alheia a tais protocolos científicos e éticos. Elas parecem ser a realização perfeita da proposição behaviorista de Skinner: “[n]ão é uma questão de trazer o mundo para dentro dos laboratórios, mas de estender as práticas de uma ciência experimental *para o mundo como um todo*” (apud Zuboff, op. cit., p. 412, grifo da autora).

Ainda que não seja nosso foco na presente investigação, consideramos que o predomínio da correlação sobre a causalidade como modelo interpretativo do mundo ajuda a explicar alguns modos contemporâneos de construção de sentido de um modo mais amplo. Pensemos por exemplo no negacionismo climático, para Latour (2020), a raiz de todos os escapismos e mesmo da ascensão dos populismos de extrema direita nos últimos anos. Para além das questões eminentemente políticas, ideológicas e econômicas implicadas, em termos puramente lógicos, não seria (também) o negacionismo climático uma incapacidade de interpretação dedutiva e causal dos fenômenos que originam o problema? A dissociação das causas sistêmicas e dos efeitos desencadeados e a produção de correlações espúrias (como entre a ocorrência de temperaturas extremas baixas e a inexistência do aquecimento global) parecem ser uma das marcas epistemológicas desse negacionismo. É claro que não se está sugerindo que a racionalidade algorítmica está na origem de fenômenos contemporâneos como o negacionismo climático. O que cogitamos é a hipótese de que ambos se inscrevem ou são o efeito de deslocamentos epistêmicos históricos que põem em crise a causalidade, ao menos aquela da ciência moderna ancorada na validação de hipóteses e na descoberta dos ‘porquês’, como modelo explicativo da realidade.

Em seus estudos sobre a digitalização da política no Brasil com ênfase nas redes bolsonaristas, Cesarino identifica o recurso ao que chama de “elos causais ocultos” (2021, p. 81) pelas narrativas conspiratórias. Nesse contexto, afirma, os sujeitos não buscam verificar a verdade por meio de controles e procedimentos fixados por estruturas estabilizadoras modernas (como a ciência, a imprensa ou as instituições do estado democrático de direito), mas conectam “evidências esparsas por meio de uma lógica aditiva” (ibid., p. 82) nutrida pela experiência pessoal e imediata. Esses elos causais ocultos remeteriam a padrões epistêmicos descritos pela antropologia clássica em práticas como a bruxaria que hoje reapareceriam sob a forma de causalidades encantadas (Comaroff, Comaroff, 2000) num contexto de

convergência entre neoliberalismo, digitalização e conspiracionismo. Para Cesarino, o recurso a essas causalidades escondidas está diretamente relacionado ao colapso das causalidades sistêmicas e lineares na “arquitetura digital do neoliberalismo” (op. cit., p. 75), que opera através de causalidades recursivas.

Para além do contexto específico do negacionismo climático, poderíamos dizer que a estrutura epistêmica de gramáticas conspiratórias que caracteriza o fenômeno da pós-verdade (Cesarino, op. cit.) dentre outras estratégias, produz e é produzida por um dilúvio de correlações espúrias que constroem os elos ocultos de que fala Cesarino (entre a implantação de antenas 5G e o início da pandemia de Coronavírus, por exemplo) e raciocínios condicionais que muito se assemelham ao constructo *se, então* (‘se as pessoas seguem morrendo, a vacina não funciona; ‘se não vemos a curvatura do horizonte, a Terra é plana’, etc) e visam construir sentido e narrativas em contextos de alta entropia informacional. Ainda que em alguns momentos a lógica que guia tais raciocínios seja encadeada por um raciocínio que se aproxima da dedução, eles nunca parecem ser capazes de acessar as causalidades sistêmicas. Como aponta Faltay (2020), ao investigar práticas conspiratórias nas plataformas digitais, existem similitudes estruturais e operacionais entre modos de percepção conspiracionistas e algorítmicos. Ambos “detectam uma estrutura onde outros vêem caos, encontram significados onde outros não os percebem” (ibid., p. 123).

Apesar de que não é nosso objetivo aqui nos aprofundarmos sobre essas gramáticas conspiratórias, elas nos interessam na medida em que evidenciam certas ressonâncias epistemológicas e cognitivas entre a ‘camada técnica’ e a ‘camada humana’ nos agenciamentos cognitivos que constituem os ambientes digitais contemporâneos (Hayles, 2017; Cesarino, 2022; Faltay, 2020). Uma delas diz respeito, como viemos demonstrando, ao um *declínio do raciocínio dedutivo* baseado na validação de premissas e hipóteses e a *ascensão de raciocínios indutivos* que operam por meio de correlações, reconhecimento de padrões e sintaxes condicionais do tipo ‘se, então’. Outra correspondência entre cognição técnica e humana nesses ambientes, está na substituição de formas *a priori* por *emergências* na produção de conhecimento e no modo como as contingências são recursivamente incorporadas em novos sistemas adaptativos. De modo semelhante aos algoritmos aprendizes, a conduta dos sujeitos nas redes conspiracionistas parece ter a capacidade de sempre incorporar novas informações, incluindo os erros e falhas de suas próprias interpretações, inferindo novas ‘regras’ sem a necessidade de manter-se fiel a nenhum pressuposto (que

poderia ser validado ou refutado), produzindo e sendo produzida por uma espécie de ‘sistema de crenças adaptativo e indutivo’⁸⁶.

A prevalência da correlação sobre a causalidade enquanto modelo epistemológico é também própria de um *governo dos efeitos e não das causas* enquanto modelo político. Sobre essa questão, Morozov (2018) recupera uma palestra de Agamben (2013) na qual este propõe que uma das transformações que o paradigma governamental baseado no dispositivo de segurança proposto por Foucault (2008) desencadeou foi uma modificação da tradicional relação hierárquica entre causas e efeitos. “Já que governar as causas é difícil e caro”, afirma Agamben, “é mais seguro e útil tentar governar os efeitos” (op. cit.). Para o filósofo, esse princípio passou a orientar transversalmente os mais variados domínios — da economia à ecologia, das políticas externas e militares às medidas internas de polícia — e ajuda a compreender, por exemplo, por que a liberalização da economia coexiste com a proliferação do controle. “Se o governo aponta aos efeitos e não às causas, será obrigado a estender e a multiplicar o controle. As causas exigem ser conhecidas, enquanto os efeitos apenas podem ser verificados e controlados” (Agamben, 2013).

Segundo Morozov, o modelo algorítmico de governo do mundo social é a concretização de um *programa político de governo dos efeitos em detrimento às causas em formato tecnológico*. Essa lógica se tornaria explícita sobretudo quando aplicada às redes de proteção do Estado de bem-estar social através, por exemplo, da responsabilização dos indivíduos por sua própria saúde. Nesse sentido, Bruno (2021) aponta a estreita relação entre a fragilização de um projeto coletivo de sociedade, que o neoliberalismo promove e produz, e o modelos de assistência algorítmica centrados na individualidade e no ‘empoderamento’ do usuário para autogerir o máximo de esferas possíveis de sua vida. Orientados por esses ideais, somos incitados, por exemplo, a automonitorar nossa saúde física ou mental por meio de aplicativos que promovem o ideal marcadamente neoliberal do “tudo por conta própria” (Cf. Bruno et al, 2020).

Sustentada pelo extrativismo de dados tão caro aos capitalistas de vigilância, a obsessão pelo monitoramento dos efeitos anda a par com a indiferença pelas causas sistêmicas dos problemas sociais. A mesma racionalidade orienta a adoção indiscriminada de ‘tecnologias

⁸⁶ Faltay (op. cit.) observa como o método de organização da realidade que as teorias da conspiração elaboram transformam os possíveis equívocos e falhas identificados em suas explicações em “elementos indiciários que reforçam a aplicabilidade de suas perspectivas” (p. 123) geralmente apontando para “complôs ocultos que agem para desmascarar a suposta verdade por detrás de um evento” (ibid.).

inteligentes’ e sistemas de decisão automatizada no setor público, em geral envolvendo a detecção antecipada dos chamados ‘perfis de risco’, seja para o acompanhamento de políticas públicas ou para decisões automatizadas em áreas como educação, saúde pública ou trabalho, adoção quase sempre viabilizada pelo tecnoassistencialismo das *big techs*.

Realização da “boa e velha utopia tecnocrática da política apolítica” (Morozov, op. cit. p. 92), tais iniciativas visam sincronizar o Estado à regulação algorítmica e monitorar continuamente o comportamento dos cidadãos, atendendo às demandas neoliberais por austeridade. Ao mesmo tempo, mantém invisíveis os fatores sociais, políticos e econômicos que estão na origem dos problemas cujas iniciativas visam supostamente solucionar, uma vez que as causas, sobretudo das múltiplas formas de precarização da vida sob o neoliberalismo, não são computáveis pela governamentalidade algorítmica. Se, como afirma Mbembe, nesse projeto “o que não é computável, não existe” (2019b), a proliferação do uso desses sistemas no setor público e a concepção dos problemas sociais somente enquanto objetos de cálculo (Hong, 2022) tende a tornar tais causas, progressivamente, ainda mais invisíveis e intangíveis. Como lembra Morozov: “O diabo não usa dados. É muito mais difícil monitorar as injustiças sociais do que a vida cotidiana dos indivíduos submetidos a elas” (op. cit., p. 91).

2.3 Provável x possível

O último aspecto que exploraremos sobre os deslocamentos que a racionalidade algorítmica e o regime de saber que lhe sustenta produz na “ordem das coisas” (Foucault, 2000) se refere àquilo que entenderemos como a *prevalência do provável sobre o possível* como técnica de governo do tempo e do futuro. Este aspecto se difere de alguns modos dos que analisamos até então. Em primeiro lugar por se tratar, mais do que exatamente de uma prevalência, de uma *contenção ou redução do possível aos limites do provável*. Ele também adquire uma relevância superior aos demais, uma vez que está no centro de uma das principais hipóteses que nos orientam. Como já enunciada, a de que um dos principais objetivos da racionalidade algorítmica é conhecer, interpretar e controlar o desdobramento dos comportamentos, fenômenos, eventos e processos de individuação ao longo do tempo reduzindo o *campo do possível* ao *campo do provável*. Ainda que o foco deste tópico continue sendo determinado ordenamento epistemológico que é condição de possibilidade para a ação do saber-poder algorítmico, neste ponto já estamos eminentemente numa incursão na temporalidade deste modelo de governo e de racionalidade. Poderíamos dizer que refletir sobre essa dimensão do

regime de saber algorítmico abre a possibilidade de investigar uma espécie de *ontologia temporal* sobre a qual se fundaria sua racionalidade.

As questões principais que conduzem a reflexão deste tópico podem ser enunciadas da seguinte maneira: o que significa que as ações sobre as ações possíveis que, segundo Foucault (1995) todo exercício do poder circunscreve, esteja, no caso dos sistemas baseados em IA, delimitado pela primazia do provável? Quais as implicações do *governo do possível* se transformar num *governo do provável*? Quais as características da ‘nova’ estatística que está na base do funcionamento desses dispositivos e como ela difere da ‘velha estatística’ que servia de instrumento biopolítico de governo?

Conforme vimos, um modelo algorítmico baseado em aprendizado de máquina raciocina, basicamente, por meio de um cálculo de probabilidades que, reconhecendo estatisticamente padrões e correlações num determinado conjunto de dados prediz, por meio de generalizações, quais os *futuros prováveis*. Uma vez que a produção de conhecimento dessas tecnologias de governo do tempo está intrinsecamente relacionada ao uso de saberes estatísticos poderíamos afirmar que a racionalidade algorítmica se baseia numa normatividade estatística que visa probabilizar a totalidade da realidade, esquadrinhando-a em termos puramente numéricos e a-significantes (Lazzarato, 2014, Rouvroy, Berns, 2015) e traçando um ‘mapa’ de suas tendências, propensões, inclinações e potencialidades sobre as quais visa agir (seja modificando-as ou comercializando-as).

Como observa Bonaldo (2023, p. 16), “a redução dessas operações a técnicas aparentemente inofensivas, saudadas como inovadoras, pode esconder a evidência de que o aprendizado de máquina elabora conceitualmente o mundo como uma extensão da atividade estatística”. Por esta razão, segundo Rouvroy e Berns, “[o] governo algorítmico parece (...) assinar a conclusão de um processo de dissipação das condições espaciais, temporais e linguística da subjetivação e da individuação em benefício de uma regulação objetiva, operacional das *condutas possíveis*” (op. cit., p. 44, grifo nosso). Essa governamentalidade abre a “possibilidade de governar os comportamentos sem se ocupar diretamente com os indivíduos para se contentar em governar a partir de uma expressão estatística da realidade” (ibid., p. 52). Assim, uma das principais ações que o governo algorítmico estabelece é uma *modulação do possível através das ‘malhas’ do provável*.

Diante das “zonas de indeterminação”⁸⁷ (Bergson, 2011) que cercam tradicionalmente o campo de ação possível dos viventes no tempo, somado a um contexto sociohistórico de múltiplas incertezas e intensos fluxos informacionais como o atual, a racionalidade algorítmica instaura uma *probabilização do possível* que visa acelerar os fluxos decisórios, reduzir a reflexividade e a incerteza e tornar as ações, comportamentos e eventos, em suma, a realidade, mais previsível (e lucrativa) para quem detém esse saber. Essa probabilização reduz radicalmente a potência e a indeterminação do possível, do virtual, do futuro (como se queira chamá-lo, afinal).

Nesse regime, simultaneamente de racionalidade, de temporalidade e de individuação, “se captura o possível, reduzindo-o à mera probabilidade” enquanto “o provável, por sua vez, nos é imposto como necessário” (Berardi, 2019b, p. 29-30). Berardi (op. cit), recuperando um termo do artista Warren Neidich, chama esse dispositivo de *statisticon*: um redutor de acontecimentos futuros à probabilidade e à previsibilidade. Situando-se no cruzamento entre a máquina técnica e a máquina social, o *statisticon* atuaria por meio de automatismos tecno-informáticos que capturam os fluxos de atividade social sob a forma de dados visando adaptar as articulações da máquina global às expectativas do organismo social e, numa via simétrica, adaptar às expectativas do organismo social à máquina global. Entre os efeitos mais graves do reducionismo que estamos descrevendo está o que Amoore chama de uma “dupla exclusão política do algoritmo”: “a condensação de múltiplos potenciais em uma única saída” [a mais provável] e o “fechamento preventivo de reivindicações políticas baseadas em atributos de dados que buscam reconhecimento antecipado” (Amore, 2020, p. 20-21).

Em sistemas de decisão automatizada essa exclusão apontada por Amoore tem efeitos graves, uma vez que seu *output* é, de fato, uma *predição-decisão*. Mesmo em casos em que não se trata de “uma única saída”, como nos grandes modelos de linguagem, como veremos no estudo de caso do ChatGPT, esse reducionismo produz uma homogeneização e empobrecimento da diversidade informacional dos ecossistemas digitais ao produzir textos baseados unicamente em correlações probabilísticas. Contudo, é claro que há usos e contextos em que essa redução das possibilidades às probabilidades pode ser muito

⁸⁷ Tendo como centro nosso próprio corpo e como condutor o processo sensório-motor, tais zonas de indeterminação que cercam a atividade dos organismos, são aquilo que permite que entre os estímulos que recebemos e as reações que a eles encadeamos se instaure a *possibilidade de escolha* (dentre as ações possíveis). Quanto mais incerta for a reação, mais hesitação deve haver entre estímulo e resposta. Bergson propõe que há uma relação entre a *indeterminação* e o *grau de liberdade* dos viventes, uma vez que ações verdadeiramente livres dependem da existência e cultivo de tais zonas.

bem-vinda, como na medicina diagnóstica, área em que os usos da IA para previsão de doenças, por exemplo, têm obtido muito êxito.

Os modelos preditivos de conhecimento e controle da realidade ancorados na probabilidade operam não só por meio da inferência de predições sobre o futuro com base no reconhecimento de padrões do passado, mas na intervenção *antecipada* no curso das ações e eventos visando aumentar ou diminuir a probabilidade de que determinadas predições se realizem ou não (conforme se tratem de riscos ou oportunidades). A antecipação é o complemento da predição estatística e uma tática muito eficaz para tornar comportamentos e eventos mais previsíveis: sabida a probabilidade de que algo vai acontecer, atua-se ativamente para incitar ou impedir tal evento, seja a tendência de um usuário ler uma mensagem em determinado horário ou de um determinado comportamento do inimigo em contexto militar. Em sua pesquisa sobre a influência da matriz comportamental sobre as tecnologias de influência das plataformas digitais, Bentes destaca como as plataformas “criaram uma arquitetura de modificação de comportamentos (Zuboff, 2020) ampla, que permite o controle fino de estímulos sensoriais conduzidos por inferências algorítmicas que pretendem *tornar determinadas escolhas e ações mais ou menos prováveis*” (Bentes, 2022, p. 238, grifo nosso). No Capítulo 3, retornaremos à temporalidade antecipatória ou preemptiva, sobretudo ligada ao contexto militar-securitário, como uma das matrizes da conopolítica algorítmica.

Uma questão fundamental acerca do modo como o provável é gerido nesse modelo de racionalidade diz respeito ao fato de que, apesar de se tratar de um cálculo de probabilidades, projeta-se sobre ele efeitos performativos de *certeza*. A crença de que ‘é provável que o futuro seja similar ao passado’, ‘é provável que os padrões se repitam ciclicamente’, em si mesma, não tem nada de radicalmente novo, como o demonstram a história da astronomia e da agricultura, por exemplo. O que parece configurar uma novidade é que essa crença, tomada antecipadamente como uma certeza, passe a estruturar um modelo de tomada de decisões automatizado que incide muitas vezes de modo irreversível no curso de atualização dos eventos, sobretudo em contextos de grande complexidade como ações securitárias-militares, financeirização, campanhas eleitorais, políticas públicas em áreas como saúde, educação e trabalho. Para dar um exemplo da área de recursos humanos, se um modelo de recrutamento e seleção de candidatos como o *HireVue*⁸⁸ avalia alguém com 75% de

⁸⁸ <https://www.hirevue.com/>

probabilidade de ser um mau funcionário, é improvável que a empresa queira contratá-lo caso tenham candidatos com pontuações de risco mais baixa. No caso do uso de reconhecimento facial aplicado a segurança pública as consequências desse ‘provável com efeitos de certeza’ são ainda mais nefastas, uma vez que são tecnologias comprovadamente falhas na identificação de corpos negros comparado a brancos e que têm sido utilizadas como instrumentos necropolíticos que reproduzem “práticas coloniais de apagamento, classificação e dominação (Silva, 2021, p. 88).

O governo do possível que a racionalidade algorítmica engendra também se manifesta nas formas de subjetivação que produz e privilegia. Rodríguez (2019), ao refletir sobre as tecnologias de subjetivação nas sociedades de controle, aborda as implicações do deslocamento da centralidade do *indivíduo* para o *dividual* identificado por Deleuze em seu conhecido ensaio (Deleuze, 1992) e atualizado por diversos autores nos últimos anos (Cf. Lazzarato, 2014; Raunig, 2016; Appadurai, 2016; Bruno, Rodríguez, 2021). Nesse sentido, o autor aponta para a captura de um “devir possível” (op. cit., p. 454) ao nível das subjetivações nesse regime. Assim, a contínua recombinação de fragmentos infra-individuais (dados) para produzir modelos supra-individuais (padrões, cálculos de risco, perfis) que conforma o modelo de subjetivação do dividual “não se referem a um indivíduo considerado em sua totalidade presente em sua individualidade atual, (...), mas mais precisamente a suas potencialidades, ao que pode ser em função do que é e o que pode ser em função do que já está fazendo” (ibid., p. 454, grifo nosso). O dividual apontaria assim para um “duplo projetado, não atual, possível”⁸⁹ (ibid.) que parte de uma materialidade (os rastros digitais) mas não aponta, necessariamente, para uma materialização. Acompanhando, mas complementando Rodríguez diríamos que esse duplo projetado que o dividual inscreve é, essencialmente, um *duplo provável* que busca captar ‘quem provavelmente o sujeito é’ e, principalmente, ‘o que provavelmente fará’, uma vez que nesse modelo a atividade se sobrepõe à identidade, ou ainda, a identidade é subsumida da atividade. O modelo de individuação (Simondon, 2009) privilegiado do dividual no regime algorítmico é o *perfil*, segundo Bruno (2013, p. 161) “uma categoria que corresponde à probabilidade de manifestação de um fator (comportamento, interesse, traço psicológico) num quadro de variáveis”. O perfil materializa como, sob a racionalidade algorítmica, a produção de saber sobre os sujeitos a partir de dados e algoritmos têm muito menos a ver com um indivíduo pré-constituído captado em sua singularidade do que com correlações interpessoais que

⁸⁹ No original: “duplo proyectado, no actual, posible”.

permitam estabelecer regularidades entre similares e extrair indutivamente categorias (tipos, classes, etc) a partir da identificação e generalização de padrões relacionais (Lury, Day, 2019; Bentes, 2022) — ainda que os discursos sobre a personalização algorítmica e seus efeitos performativos advenham justamente de sua suposta capacidade de hiperindividualização, de captar as singularidades dos sujeitos sobre as quais nem eles mesmos tinham acesso.

Lury e Day (2019), refletindo sobre a personalização algorítmica como um modo específico de individuação, ressaltam como o suposto indivíduo visado por estas técnicas é sempre “mais e menos que um”, produzindo o que Cesarino chama de “uma ferradura entre *hiperindividualização* e *hiper-relacionalismo*” na qual “os sujeitos são ao mesmo tempo ‘um e muitos’” (Cesarino, 2022, p. 113). Como sintetiza Bentes sobre essa dinâmica: “o modo de individuação da personalização algorítmica está sujeito a uma divisibilidade recursiva na qual um indivíduo é refratado em múltiplas ordens parciais que permitem inúmeras formas de comparação continuamente atualizadas.” (Bentes, op. cit., p. 107). Essa refração e metaestabilidade (Simondon, 2009) manifesta-se, inclusive, no interior dos próprios perfis ‘individuais’, atualizados quase em tempo real a cada ação que o usuário executa nos ambientes digitais que o monitora. Isso faz com que, por exemplo, um usuário pode ser “92% masculino às 9:30 pm, mas oito horas depois, às 5:30 am, após uma noite de sono sem visitar novos sites, esse usuário pode ser agora 88% masculino” (Cheney-Lippold, 2017, p. 71). Ou ainda, que não haja qualquer coerência de identidade entre meus perfis como eleitora, cidadã, consumidora, profissional, etc. Nesse contexto, o sujeito se converte numa “série de combinados parciais no meio de um campo refratário a toda e qualquer unificação” (Mbembe, 2021, p. 97), apontando para um *cronotopo fractal* que atravessa tanto a camada técnica quanto a humana das arquiteturas digitais contemporâneas e sobre o qual nos debruçaremos sobretudo no estudo de caso sobre o TikTok (tópico 3.2.1).

Do ponto de vista da intersecção entre os dispositivos da segurança e do controle (Foucault, 2008; Deleuze, 1992), o que essas tecnologias de governo do tempo voltadas para a predição visam, em última instância, é introduzir uma “racionalização do acaso” (Foucault, 2008, p. 76) ou, mais precisamente, uma *gestão calculada da agência do acaso*. Como argumenta Bonaldo (2023, p. 6), “[n]ão é a eliminação da contingência, mas as técnicas de seu controle, que estão em jogo”. Rouvroy e Berns, aos quais novamente retornamos porque parecem ser os autores que captaram mais precisamente a relação entre a governamentalidade algorítmica e a captura das potências do porvir (p. 53), apontam como ela manifesta o sonho de um acaso sistematizado.

Mais ativamente, o governo algorítmico não apenas percebe o possível no presente, produzindo uma ‘realidade aumentada’, uma atualidade dotada de uma ‘memória do futuro’, *mas também dá consistência ao sonho de um ‘acaso’ sistematizado*: nosso real teria se tornado o possível, nossas normas querem antecipar, corretamente e de maneira imanente, o possível, o melhor meio sendo certamente apresentar-nos um possível que nos corresponda e para o qual os sujeitos só precisam deslizar. (2015, p. 48, destaque meu)

Nesse sentido, essas tecnologias e sistemas algorítmicos preditivos são indiscutíveis herdeiros das técnicas e práticas que regulam a agência do acaso, entre as quais a *estatística* é o instrumento por excelência (Hacking, 1990, 2006). A ideia de fundo que guiou o desenvolvimento histórico de tais ferramentas é a de que o *acaso pode ser domesticado e a incerteza pode ser medida* (Hacking, 2006; Berstein, 2011). No entanto, apesar de partilharem tal premissa, existem diferenças substanciais entre o uso de saberes estatísticos enquanto instrumento biopolítico de governo sobre uma população — que integra aquilo que Foucault vai entender como o *dispositivo da segurança* (Foucault, 2008) — e o “novo poder estatístico” (Rouvroy, Berns, 2010) que emerge com o uso de IA e Big Data por parte de grandes corporações de tecnologia, conforme argumentamos a seguir.

Desenvolvida em meados do século XVIII como uma disciplina que se ocuparia dos fatos notáveis sobre o Estado (Rodríguez, 2019), a estatística se consolida no século XIX como um dos saberes privilegiados para a constituição dos Estados-nação europeus por sua capacidade de fornecer uma compreensão da população em sua totalidade. As próprias noções de *população* e de *meio* são tributárias do desenvolvimento da estatística (ibid.). Ian Hacking (2006) em seus estudos sobre a história da probabilidade, relata como a “avalanche de números impressos” que inunda a Europa ocidental no século XIX — quando taxas anuais de crimes, suicídios, encarceramento de loucos, prostituição, doenças, produção, comércio etc, antes secretos, são tornadas públicas por agências governamentais — possibilitou a identificação de certas *regularidades* no aparecimento e desaparecimento dos fenômenos sociais. Essa mudança resultou num entendimento probabilístico do mundo, em detrimento de um determinístico, segundo o qual o acaso é simultaneamente algo indeterminado e está sujeito a leis gerais sociais ou físicas (Caliman, Almeida, 2009). Paulatinamente, o reconhecimento das (ir)regularidades dos fenômenos, na medida em que viabilizou o *cálculo da média*, formalizou uma nova concepção do *normal* e do *anormal*⁹⁰.

⁹⁰ Em *Segurança, Território e População*, Foucault (2008) descreve como o dispositivo da disciplina e da segurança tratam de forma completamente distinta a normalização. A disciplina parte de um modelo prescritivo, uma norma, e a partir disso distingue o normal (quem é capaz de se conformar à norma) e o anormal (quem não é). Foucault chama a esse modelo, que vai da norma à demarcação do normal e do anormal, de *normação*. Já no dispositivo da segurança, ao agir por meio de uma “racionalização do acaso e das probabilidades” (p. 76),

A caracterização que Foucault (2008) faz do dispositivo da segurança é compatível com a emergência de um outro entendimento dos fenômenos no tempo a partir da popularização da probabilidade apontada por Hacking (op. cit.). Seja na “gestão das séries abertas” (ibid., p. 27) — que insere os fenômenos numa “série de acontecimentos prováveis” (ibid., p. 9) controlados por uma estimativa de probabilidades — ou no papel fundamental que o *futuro* passa a ter nessa nova tecnologia de governo (ibid., p. 26), é possível reconhecer a consolidação de uma forma de controle sobre o tempo, e conseqüentemente sobre o espaço, exercida por meio da *probabilização do futuro*. A importância da dimensão temporal no regime descrito por Foucault fica bastante explícita quando o autor declara que “[o] espaço próprio da segurança remete portanto a uma *série de acontecimentos possíveis, remete ao temporal e ao aleatório*, um temporal e um aleatório que vai ser necessário inscrever num espaço dado” (ibid., p. 26, grifo nosso). Se bem o diagnóstico de Foucault coincide em muitos aspectos com o regime de conhecimento e de temporalidade do governo baseado em dados massivos e algoritmos que viemos descrevendo e segue sendo valioso para pensar as matrizes do caráter preditivo e antecipatório desse modelo, existem diversas rupturas entre aquele uso do saber estatístico e o que estamos investigando, conforme exploraremos.

Mas além da ‘matriz biopolítica’ da estatística enquanto instrumento de governo populacional nascida no séc. XVIII, é fundamental lembrar como o deslocamento de um entendimento determinístico para um probabilístico do mundo se espalhou para vários domínios durante os séculos seguintes, notadamente para a física do séc. XX, influenciando fortemente a cibernética. Finn (2017) destaca como, ao nível filosófico, a visão de Wiener sobre a cibernética também refletia essa transição da certeza para a probabilidade. Ao definir a nova ciência da comunicação e do controle, ele chega a afirmar: “[s]e eu fosse escolher um santo patrono para a cibernética, elegeria Leibniz” (Wiener, 2017, p. 35). Os avanços da relatividade einsteiniana e a mecânica quântica sugeriram que a incerteza, ou a indeterminação, era constituinte do universo, marcando também nessa área a crise do modelo causalista de entendimento do universo, “a noção de que o universo operava por leis simples e todo-poderosas que poderiam ser descobertas e dominadas”⁹¹ (Finn, op. cit., p. 27) e a emergência de um modelo baseado na complexidade, no relativismo e na probabilidade. Derivado para a cibernética principalmente via teoria da informação e mecânica estatística, o

primeiro identifica-se o normal e o anormal através das curvas de ocorrência dos fenômenos, para daí deduzir-se as normas. Para ele, aqui já não se trataria de uma *normação*, como na disciplina, mas de uma *normalização*.

⁹¹ No original: “the notion that the universe operated by simple, all-powerful laws that could be discovered and mastered”.

paradigma probabilista se baseava na “noção de que a probabilidade se aplica tanto à informação quanto à realidade material”⁹² (ibid.).

Essa transição corresponde muito bem às diferenciações feitas por Flusser entre os modelos causalistas, finalistas e programáticos que expusemos no tópico anterior. Conforme descreve Petronio (2019) sobre a interpretação do filósofo, “a principal transformação trazida pelo paradigma do programa é o novo estatuto conferido ao acaso” (p. 185). Para o causalismo, as *leis* que regem a realidade podem ser inferidas por dedução. Para o finalismo, a natureza obedece a um *telos* que coordena a totalidade dos eventos. Por sua vez, “os programas lidam com o acaso e com a redução das possibilidades em probabilidades como modelo de controlar pragmaticamente o acaso” (ibid.).

Mas retornemos novamente à ‘matriz biopolítica’ para refletir sobre algumas diferenças entre as práticas modernas da estatística e as contemporâneas ligadas às práticas digitais e algorítmicas. Uma primeira delas é que a gestão desse saber-poder passa das mãos do Estado para as das grandes corporações de tecnologia; ainda que, na prática, o que testemunhamos é uma colaboração íntima entre ambos que materializa uma das facetas do colonialismo de dados (Mejías, Couldry, 2019; Cassino, Souza, da Silveira, 2021). Como apontam Pasquinelli e Joler (2020) ao abordarem a IA enquanto instrumento de conhecimento, “a norma institucional tornou-se computacional”: a classificação dos sujeitos, dos corpos e comportamentos parece não ser mais um assunto para os registros do Estado sobre sua população, mas para os algoritmos e centros de dados privados. Esse fato, por si só, tem implicações radicais.

Davis (2017), ao descrever a crescente perda de importância da estatística, coletada e compilada por especialistas técnicos frente aos dados que acumulam e são analisados por padrões no contexto digital, aponta como a maioria de nós é totalmente alheia ao que todos esses dados dizem a nosso respeito, individual ou coletivamente. Sabe-se que o caráter opaco e ininteligível dos processos algorítmicos é um dos motivos pelos quais as assimetrias de poder desse regime de saber se manifesta e sua contestação por meio de ações coletivas é dificultada. Ou seja, não há nada que ancore esse novo modelo de produção de saber a interesses públicos e coletivos. Pouquíssimas descobertas sociais desse tipo de análise de dados acabam no domínio público (Davis, op. cit.). Em geral, as publicações que descrevem tais ‘descobertas’ visam quase que exclusivamente alardear as virtudes do ‘oráculo dos

⁹² No original: “is the notion that probability applies to information as much as to material reality”.

dados’ e como somos previsíveis (caso se tenha acesso irrestrito a nossos rastros digitais, é claro). Isso significa que pouco (ou nada) contribuem para construir debates políticos fundados numa *realidade compartilhada*. Para Davis (op.cit.), “uma sociedade pós-estatística é uma proposição potencialmente assustadora, não porque careceria de quaisquer formas de verdade ou conhecimento, mas porque as privatizaria drasticamente.”⁹³

Outra ruptura entre o ‘velho’ e o ‘novo’ poder estatístico (que já mencionamos), se refere ao fato de que o modelo algorítmico de análise da realidade dispensa médias, categorias, regras ou hipóteses prévias, uma vez que, como vimos, opera por meio da indução de correlações. Por isso, Rouvroy e Berns (op. cit.) o chamam de *a-normativo*. Enquanto a estatística moderna dependia de um sistema de escalas e categorias fixas de análise, dentro das quais os sujeitos e fenômenos eram classificados e fixados, os vastíssimos e velozes conjuntos de dados contemporâneos são explorados por algoritmos em busca de padrões, correlações e tendências emergentes de forma recursiva e potencialmente em tempo real sem categorias pré-existentes, criando segmentações do tipo *bottom-up* segundo a lógica conexionista. Sob a ‘nova estatística’, primeiro capturam-se os dados, depois formulam-se as questões de pesquisa (Davis, op. cit.). Esse novo paradigma trabalha com uma granularidade muito maior de dados e sistemas classificatórios, o que permite criar uma infinidade de categorizações atomísticas e microsegmentadas dos fenômenos e indivíduos (Cardon, 2019; Chun, 2021; John-Matheus, Cardon, Balagué, 2022), muitas vezes incongruentes entre si, inclusive. Essa fragmentação virtualmente infinita do sujeito em categorias estatísticas cada vez mais granulares também implica, segundo Rouvroy, num abandono da experiência e da política do comum:

fragmentado, o sujeito se transforma numa forma de uma miríade de dados que o ligam a uma multidão de perfis (como um consumidor, um fraudador em potencial, um funcionário mais ou menos confiável e produtivo e assim por diante). Todos eles estão relacionados a ele ou ela sem inscrevê-lo em qualquer contexto coletivo (diferentemente dos modos ‘clássicos’ de categorização, como o perfil étnico, que eram ajustados a categorizações socialmente comprovadas e, portanto, suscetíveis de dar origem a ações coletivas), o indivíduo também é dispensado de prestar contas a alguém, tornando-se infinitamente calculável, comparável, indexável, intercambiável e em concorrência⁹⁴. (Rouvroy, 2016, p. 34)

⁹³ No original: “A post-statistical society is a potentially frightening proposition, not because it would lack any forms of truth or expertise altogether, but because it would drastically privatise them.”

⁹⁴ No original: “fragmented, the subject comes in the form of a myriad of data that link him or her to a multitude of profiles (as a consumer, a potential fraudster, a more or less trustable and productive employee and so on). All of them are related to him or her without inscribing him or her in any collective context (differently from the ‘classical’ modes of categorisation, such as the ethnic profiling, that were adjusted on socially proved categorisations and therefore susceptible to give rise to collective actions), the individual is also dispensed from

A própria noção de *população*, central para os dispositivos da segurança e o modelo biopolítico de governo (Foucault, 2008, 2008b), se torna obsoleta no contexto digital. Se a estatística tradicional visava fornecer uma compreensão da população em sua totalidade, ancorada em um território geográfico específico, o *statisticon* digital (Berardi, 2019b) está predominantemente centrado na construção de perfis e personalizações para fins comerciais. Wendy Chun (2021) descreve essa mudança como uma transformação do predomínio da *nação* para o da *vizinhança* através do axioma da *homofilia*, o “amor pelo similar”. Esse axioma, que segundo a autora governa as arquiteturas digitais atualmente, presume que “similaridade gera conexão”. É a “*conexão do igual com igual*”. A relação entre indivíduos e populações ainda importa, diz ela, mas o grupo relevante agora é a ‘vizinhança de rede’ ou o ‘*cluster* homofílico’, agrupamentos baseados em atributos similares e não em noções de igualdade. O título de um post do TikTok resume com precisão essa inflexão: “As subculturas são a nova demografia”.⁹⁵

A centralidade que o princípio da agrupação por da similaridade assume na estruturação das redes e nas mediações algorítmicas, para Chun, estão na base da segregação, polarização e proliferação dos filtros-bolha ou câmaras de eco que passaram a dominar as redes e as identidades, sociabilidades e subjetivações nelas produzidas nos últimos anos. Para a autora, tais efeitos não seriam, portanto, acidentais, mas desdobramentos lógicos dos axiomas segregacionistas que orientam a arquitetura e o funcionamento das redes na contemporaneidade.

Ainda que o foco de análise de Chun (2021) ao analisar o dispositivo da homofilia esteja em seus efeitos sobre sujeitos, esse princípio orienta a ordenação de todos os dados geridos sob a mediação algorítmica no contexto da plataformização; inclusive porque, do ponto de vista técnico, a homofilia pode ser considerada uma modalidade de reconhecimento de padrões e correlações. No atual estágio das redes, podemos afirmar que a homofilia está mais centrada na conexão de conteúdos semelhantes do que de indivíduos semelhantes. Mais precisamente, ela atua antes segmentando conteúdos para, em seguida, segmentar os sujeitos em torno desses conteúdos, ainda que essas duas camadas se retroalimentem de forma recursiva. Essa ‘homofilia de conteúdos’ fica evidente tanto na arquitetura de construção das IAs Generativas, especialmente nos LLMs que operam a partir de agrupamentos semânticos

giving an accounting to someone, becomes infinitely calculable, comparable, indexable, interchangeable and in concurrence.”

⁹⁵ https://www.tiktok.com/business/en-US/blog/subcultures-are-the-new-demographics?ab_version=control

similares, quanto no modelo predominante hoje nas redes sociais, liderado pelo TikTok, que desvincula a personalização algorítmica da conexão social entre indivíduos, conforme veremos nos estudos de caso.

Bruno (2020), retomando os estudos de Chun sobre a homofilia e refletindo sobre as implicações da personalização do campo perceptivo nas plataformas digitais, explora dois efeitos produzidos pelas arquiteturas estruturadas a partir do princípio da homofilia/similaridade. Esses efeitos dialogam com a captura do possível em detrimento do provável nos modos algorítmicos de governo do tempo sobre os quais viemos refletindo ao longo deste tópico. O primeiro deles é aquilo que a autora entende como “*o confisco do comum e das diferenças*” que ele implica (p. 260). Ainda que essas mediações se alimentam e dependam de nossa ininterrupta interação e socialização, incitados e apreendidos sob as métricas do engajamento, o campo do comum que emerge nesse contexto é regulado por algoritmos que produzem conexões que “tendem a confinar as pessoas em mundos pouco permeáveis a contradições, ambiguidades, diversidade e diferenças” (ibid.). Ao favorecer a socialização com base na similaridade, os algoritmos expropriam “uma riqueza coletiva e relacional que potencialmente poderia dar lugar a múltiplos modos de fazer comum, inclusive aqueles gestados em meio a desacordos e diferenças” (ibid.). Somado à privatização da produção de saberes pelas *big techs* descrita anteriormente, o confisco do comum diagnosticado por Bruno aponta para um abandono radical de toda perspectiva de coletivização em favor de uma (suposta) hiperindividualização (Rouvroy, 2016) e da *reprodução recursiva e massiva do similar* (de identidades, de conteúdos, etc), geralmente vendidas sob a cifra da autenticidade e da otimização. Ou pior, a estruturação das redes e, por conseguinte, de nossos modos de individuação e socialização nelas produzidos a partir do princípio homofílico alimenta a perversa crença de que *o comum é o similar* quando, na verdade, “o comum necessita e pressupõe [justamente] a não-coincidência” (Rouvroy, Berns, op. cit., p. 55).

A propensão das arquiteturas algorítmicas a tornarem pouco provável a emergência de diferenças e disparidades entre entidades heterogêneas envolve, para Bruno, um segundo efeito, que ela descreve como “*o sequestro do futuro nas plataformas digitais*” (op. cit., p. 262, grifo nosso) e que ressoa vários aspectos que viemos abordando e ainda abordaremos nesta investigação. O “sequestro do futuro” descrito por Bruno se refere à capacidade de agir sobre os comportamentos em tempo real que marca o modelo de negócios hegemônico das plataformas digitais e que depende da contínua antecipação de nossos comportamentos

envolvendo, simultaneamente, os *mercados de comportamentos futuros* descritos por Zuboff (2021) e a capacidade de modificar comportamentos em tempo real intervindo sobre nossas arquiteturas de escolha nesses ambientes. Essa intervenção contínua e calculada sobre o ‘próximo passo’ — cliques, curtidas, visualizações e interações com este ou aquele conteúdo, compartilhamentos, etc. — implica para a autora “um sequestro, no nível cotidiano, do nosso campo de ação possível, colocado a serviço de mais e mais engajamento” (Bruno, op. cit., p. 263). Ao conduzir, sutil mas permanentemente, nossas condutas em tempo real e regular nosso campo perceptivo nos ecossistemas digitais, esse modelo de mediação aumenta a probabilidade de que atuemos conforme as previsões sobre nós mesmos e torna, recursivamente, nossos *comportamentos ainda mais previsíveis*. Nesse contexto, “o futuro e a ação possível, como reserva aberta de possibilidades, de encontros e de inesperados, são sequestrados nessas microantecipações cotidianas nos ambientes e plataformas online” (ibid., p. 264).

Os apontamentos de Bruno sobre o confisco do comum e o sequestro do futuro pelas/nas plataformas digitais reiteram a hipótese que viemos desenvolvendo ao longo do presente texto: que as dimensões mais poderosas e as implicações mais drásticas da racionalidade algorítmica e do modelo de conhecimento que lhe sustenta — modelo este, como vimos, ancorado na predição, na indução, na performatividade, na correlação e na probabilização da realidade — concernem ao modo como regula o *campo do possível*. Reiteramos assim como o alvo privilegiado do modelo de conhecimento e gestão da realidade baseado em dados massivos e algoritmos ditos inteligentes e de sua *cronopolítica* é, precisamente,

“o que poderia advir”, a parte de incerteza, de virtualidade, de potencialidade radical que faz dos seres humanos processos livres para se projetar, relatar-se, tornar-se sujeitos, individuar-se, seguindo trajetórias relativamente e relacionalmente abertas. (Rouvroy, Berns op. cit., p. 54, grifo nosso).

Uma vez que não se trata somente de novas formas de conhecer o sujeito e o mundo, mas de como eles se produzem no tempo, refletir a fundo sobre as implicações da agência algorítmica para a constituição do campo do possível e defender, como sugere Zuboff, o “direito ao tempo futuro” (2021, p. 376) — o que depende, entre outras coisas da preservação de zonas de indeterminação nos fluxos do devir (Bergson, 2011) e de uma experiência e política do comum que não sejam tributárias da similaridade, mas da diferença — parecem se tornar tarefas urgentes para produzir uma bifurcação nas culturas algorítmicas e nos mundos que suas práticas produzem.

EIXO CRONOPOLÍTICO

Capítulo 3: Regimes de temporalidade algorítmicos

Uma das hipóteses centrais que orienta nossa pesquisa é que, em virtude da centralidade da predição para a racionalidade algorítmica, os algoritmos da inteligência artificial são *tecnologias de governo do tempo* cujas maiores implicações residem no modo como modulam o curso das ações, eventos e comportamentos possíveis. Conforme vimos, esse conhecimento e intervenção sobre a dimensão temporal, ou esse governo algorítmico do tempo, se dá principalmente através do cálculo dos futuros prováveis, visando tornar o curso das ações e eventos mais previsíveis. Como aponta Bucher, os algoritmos baseados em aprendizado de máquina ou conexionalistas “são antecipatórios na sua própria ‘lógica operativa’, o que significa que a *antecipação está inscrita nos próprios mecanismos e regras do sistema*”⁹⁶ (Bucher, 2012, p. 14, grifo nosso). Esse modelo de conhecimento e gestão preditivos e antecipatórios da realidade, que exploramos em profundidade nos dois primeiros capítulos, constitui o cerne de um modelo econômico que, mais do que centrado naquilo que os padrões e correlações entre dados ‘dizem’ sobre o que os fenômenos e os próprios sujeitos *são*, está interessado naquilo que *podem ou tendem vir a ser*; um modelo que “opera no futuro tanto quanto o cálculo de probabilidades coincide com a extração preditiva de *mais-valia*”⁹⁷ (Parisi, Dixon-Román, 2020, p. 117).

Deste modo, indagar sobre os problemas específicos da temporalidade na época digital (Hui, 2021) e, mais especificamente, sobre *quais temporalidades foram abertas pela algoritmização da vida e do próprio tempo* e, numa via complementar, quais seus efeitos, são questões que merecem ser melhor aprofundadas em nossa investigação. Para tanto, este e o próximo capítulo estarão centrados sobre o que chamamos de *cronopolítica algorítmica*, buscando refletir sobre o governo algorítmico do tempo para além da redução do possível ao provável — ainda que este permaneça sendo um ponto fundamental de nossa proposição.

Ao longo deste capítulo exploraremos os regimes de temporalidade algorítmicos em duas vias. Num primeiro momento, centrando-nos em algumas *matrizes da cronopolítica algorítmica*. Tais matrizes, a saber, a *cibernética*, a *militar-securitária* e a *financeira*, nos

⁹⁶ No original: “are anticipatory in their very ‘operational logic’ (Wardrip-Fruin, 2009: xi), which means that anticipation is inscribed into the very mechanics and rules of the system.”

⁹⁷ No original: “it operates in the future as much as the calculation of probabilities coincides with the predictive extraction of surplus value.”

ajudarão a dar mais visibilidade às forças históricas, sociais, econômicas e políticas que dão forma ao desenvolvimento técnico dos algoritmos ao longo do tempo. Afinal, parafraseando Deleuze (1992), os algoritmos sozinhos não explicam nada. Em termos metodológicos, mapear essa dimensão consiste em “esforçar-se por discernir os projetos estratégicos que comandam as escolhas técnicas ao mesmo tempo que são, por sua vez, determinados por elas” (Chamayou, 2015, p.24) e permite tornar explícito os *diagramas* em que os *programas* destas tecnologias se atualizam (Bruno, 2012)⁹⁸.

É preciso salientar que existe uma dimensão hipotética na própria escolha dessas matrizes. Isso porque, longe de serem as únicas matrizes possíveis disso que entendemos como cronopolítica algorítmica, são aquelas que, dadas as questões que nos inquietam de modo privilegiado, nos pareceram oferecer mais pistas para o entendimento tanto dos regimes de racionalidade quanto de temporalidade que orientam o governo algorítmico do tempo. Além disso, existe uma dimensão histórica-material-política-econômica que direciona essa escolha. Essa dimensão remete principalmente à conjuntura histórica do pós-guerra e à co-emergência das bases da cibernética, das tecnologias computacionais, do complexo industrial-militar, da doutrina neoliberal e da financeirização das estruturas econômicas (Cesarino, 2022).

Não se trata de uma xx exaustiva de tais matrizes, mas da exploração de aspectos específicos que nos pareceram mais relevantes para o entendimento dos regimes de temporalidade algorítmicos: a *recursividade* na cibernética; a *antecipação* e *gestão de riscos* na militar-securitária; a *aceleração* e a *comodificação do futuro* na financeira. Finalmente, observamos que elas não estão, por assim dizer, no mesmo plano topológico, uma vez que a matriz cibernética e sua temporalidade recursiva subjaz às outras duas, sobretudo à financeira.

A segunda parte deste capítulo se dedica a dois estudos de casos, um sobre o *TikTok* e outro sobre o *ChatGPT*. Ainda que se tratem de sistemas algorítmicos bastante distintos, seja em seus usos e finalidades ou nas especificidades técnicas de seus algoritmos, ambos são manifestações contundentes da vertente *preditivo-aceleracionista* (Bruno, Pereira, Faltay,

⁹⁸ Inspirando-se na leitura de Deleuze sobre Foucault, Bruno (2012) propõe uma diferenciação entre as noções de *programa* e *diagrama*. Ao analisar tecnologias de vigilância, a autora alerta que seus programas, “isto é, aquilo que as suas redes de sustentação promovem como suas qualidades, objetivos, necessidades” (p. 26) não resumem nem coincidem plenamente com o diagrama ao qual pertencem. Um diagrama, afirma, excede um modo de exercício do poder (atualizado no programa), sendo antes uma exposição das relações de força que constituem o poder (Deleuze, 2005). “Deste modo, a análise de um e outro implica não apenas descrever o modo como uma determinada configuração de poder se atualiza num programa qualquer, mas também estar atento às múltiplas relações de força que o tornam possível e que também o atravessam, excedendo as suas determinações e abrindo vias de ruptura e transformação” (ibid., p. 27).

2023) que vem dominando a IA nos últimos anos. Ambos também evidenciam problemas de ordem cronopolítica que simultaneamente materializam questões que viemos abordando ao longo desta investigação e abrem outras, complexificando nossos próprios argumentos e hipóteses.

No caso do TikTok, app que se transformou rapidamente num fenômeno de engajamento entre os usuários, analisaremos como o controle algorítmico da temporalidade é um elemento fundamental para compreender a eficiência e eficácia do TikTok tanto na gestão dos comportamentos dos usuários quanto na potencialização do extrativismo que sustenta sua arquitetura algorítmica. Veremos como, a aceleração e a fragmentação, que se desdobram em múltiplas estratégias e camadas do app, transformam-o no que chamamos de *ambiente digital de alta frequência*.

Já no estudo de caso do ChatGPT, cujo lançamento em novembro de 2022 desencadeou o chamado ‘hype da IA’ e modificou radicalmente os imaginários e expectativas mercadológicas sobre a IA, revisitaremos a questão da redução das possibilidades às probabilidades através de uma exploração minuciosa do funcionamento dessa *megamáquina probabilística*. Refletiremos sobre o que é, como opera e quais as implicações desse modo outro-que-humano de conhecer e agir sobre a realidade que, sob a superfície apresenta habilidades cada vez mais similares ao humano, mas é movido por um “modelo de mundo” (Amoore et al, 2024) fundamentalmente distinto do nosso.

3.1 Matrizes da cronopolítica algorítmica

3.1.1 A matriz cibernética: temporalidade recursiva e causalidade circular

Apesar de que já abordamos em diversos momentos ao longo deste trabalho a relação intrínseca entre racionalidade algorítmica e cibernética, neste tópico nos deteremos sobre a estrutura temporal da recursividade, um dos fundamentos operacionais das máquinas cibernéticas. A centralidade do *mecanismo do feedback* como forma de controle em sistemas cibernéticos e da temporalidade recursiva lançam luz sobre uma cronopolítica que é transversal a toda a infraestrutura digital e aos fenômenos que aí emergem. Esse mecanismo modula desde o modo como as camadas internas de uma rede neural aprendem um padrão; a atualização ininterrupta que caracteriza a temporalidade das redes (Chun, 2016); a ressonância de padrões cognitivos entre agência técnica e humana em contextos de alta

interpenetração entre ambos (Cesarino, 2021, 2022; Hayles, 2017); as técnicas de condicionamento e automatização comportamental utilizadas pelas plataformas digitais (Eyal, 2020, Bentes, 2022); a proliferação das ‘câmaras de eco’ e dos ‘filtros bolha’ no ciberespaço; a amplificação de padrões históricos datificados em sistemas algorítmicos, e assim por diante. A recursividade é tão estruturante e onipresente nos fenômenos digitais atuais, modulando de micro a macro objetos e eventos, que tende a tornar-se invisível. Não é exagero afirmar que a recursividade se tornou uma das estruturas temporais mais relevantes da contemporaneidade. Trata-se também de um aspecto crucial para compreender mais profundamente aquilo que viemos descrevendo como o caráter performativo da mediação algorítmica (ver especialmente tópico 2.1) e os efeitos sistêmicos dessa mediação nos ecossistemas digitais.

Modus operandi de todas as máquinas digitais, “a visão original da cibernética de fato fundou os parâmetros elementares do complexo industrial-militar a partir do qual emergiu a atual indústria *tech*” (Cesarino, 2022, p. 34). No fundacional livro *Cibernética ou controle e comunicação no animal e na máquina*, no qual apresenta os princípios de uma teoria que contemplasse o funcionamento de máquinas e organismos vivos, Wiener ([1948] 2017) ressalta a centralidade do mecanismo do *feedback* para o entendimento unificado dos fenômenos na cibernética. O próprio termo grego do qual provém cibernética — que pode ser traduzido como *timoneiro* e possui a mesma raiz da palavra *governo* — faz referência ao sistema de pilotagem de um navio como “uma das primeiras e mais bem desenvolvidas formas de mecanismos de feedback” (Wiener, op. cit., p. 34).

Na clássica metáfora utilizada para ilustrar um sistema cibernético, o timoneiro, a fim de fazer com que a embarcação navegue na direção pretendida, percebe as oscilações do ambiente e faz os ajustes necessários para conduzi-la. Nesses sistemas, o ajuste a uma finalidade pré-determinada se dá utilizando a diferença entre o objetivo e o movimento realmente efetuado como um novo *input* que regula o comportamento do sistema na direção desejada. Ou seja, os *outputs* de erro retornam como novos *inputs* e são incorporados ao aprendizado do sistema. Os ciclos de atualização materializados nos *feedback loops*, ou no que preferimos aqui chamar de recursividade, introduzem assim *causalidades circulares* que modulam a adaptação dinâmica entre o sistema e seu ambiente (Cesarino, op. cit.). Um ciclo de *feedback* é considerado *negativo* quando seu efeito regula e estabiliza um sistema

(regulando sua temperatura, velocidade ou posição, por exemplo) e *positivo*, ao contrário, quando amplifica as tendências em curso e tende a conduzir o sistema ao desequilíbrio⁹⁹.

Segundo a perspectiva cibernética, essa dinâmica subjaz ao funcionamento de todos os sistemas vivos e orienta desde o banal ato de agarrar um copo até o mecanismo da seleção natural de Darwin (Cesarino, op. cit.). Para a metaciência proposta por Wiener, máquinas, organismos vivos, indivíduos ou sistemas sociais são concebidos como funcionalmente equivalentes, uma vez que todos são modelos de organização e comunicação regulados por *feedbacks*. Por este motivo, seus idealizadores acreditavam que os princípios da cibernética poderiam orientar transversalmente campos como a economia, a antropologia, a sociologia e a psicologia. Governar, em termos cibernéticos, consiste, portanto, em utilizar a recursividade como princípio regulador da estabilidade dos sistemas.

É interessante notar também como a *previsão* integrou a matriz epistemológica e cronopolítica da cibernética desde as suas origens. O próprio Wiener trabalhou ativamente no aperfeiçoamento da precisão dos sistemas de artilharia antiaérea durante a Segunda Guerra, aplicando princípios da então incipiente ciência para a previsão da trajetória dos aviões em combate (Wiener, op. cit). Halpern observa que, considerando esse investimento originário na previsão de comportamentos humanos e maquínicos em contextos bélicos (a trajetória de projéteis, os movimentos dos exércitos inimigos, etc), não é de se surpreender que a *gestão do futuro* tenha se tornado uma preocupação central para a cibernética (Halpern, 2022).

Fazendo uma retomada da história da cibernética, Hayles (1999) observa como num momento inicial, sobretudo no período das famosas Conferências Macy¹⁰⁰, a ideia de *feedback* esteve muito vinculada a de *homeostase*, aproximando a função regulatória do mecanismo em sistemas biológicos para maquínicos. Já num segundo momento, quando surge a chamada *cibernética de segunda ordem*, o *feedback* se aproxima da ideia de *reflexividade*, uma vez que a questão-chave passa a ser como os sistemas são constituídos

⁹⁹ Apesar de fazer algumas menções sobre *feedbacks* positivos, Wiener enfatiza sobretudo a importância dos *feedbacks* negativos para o funcionamento das máquinas cibernéticas. Curiosamente, como abordamos ao longo deste tópico, a maioria das infraestruturas digitais contemporâneas é dominada por excessos de *feedbacks* positivos.

¹⁰⁰ As Conferências Macy foram uma série de encontros interdisciplinares realizados entre 1946 e 1953 em Nova Iorque, patrocinados pela Fundação Macy. Com o objetivo de reunir cientistas de distintas áreas para discutir ideias que pudessem ser aplicadas a diferentes campos do conhecimento e com um foco especial na compreensão de sistemas complexos, elas estabeleceram as bases teóricas da cibernética.

levando em consideração a agência do observador¹⁰¹. Para Hayles, a “[r]eflexividade é o movimento através do qual aquilo que foi utilizado para gerar um sistema é transformado, através de uma perspectiva alterada, em parte do sistema que ele gera”¹⁰² (ibid, p. 8). Segundo a autora, a recursividade “tende notoriamente ao regresso infinito”¹⁰³ (ibid., p. 9) e pode ter efeitos subversivos ao confundir e emaranhar as fronteiras que impomos ao mundo.

Em *Recursividade e Contingência*, Hui (2019) propõe uma abordagem dos sistemas cibernéticos a partir do conceito do *orgânico*. Considerando a cibernética como parte do paradigma das ciências que denomina de *organicismo*, o autor entende a cibernética enquanto um *mecano-organicismo* cuja matriz epistêmica seria capaz de superar os dualismos vivo/não-vivo, natureza/tecnologia, máquina/organismo, micro/macro (Hui 2019, 2020). Isso porque, como enfatiza o autor, a noção de recursividade representa uma ruptura com o modelo mecanicista que dominou os séculos XVII e XVIII, especialmente o mecanicismo cartesiano e a causalidade linear dos fenômenos que esse modelo pressupõe. Enquanto o modelo mecanicista opera por meio de uma *causalidade linear* (A-B-C), a cibernética opera por meio de uma *causalidade circular* (A-B-C-A’) que introduz um movimento reflexivo não linear.

No primeiro capítulo de seu supracitado livro, Wiener opõe o *tempo newtoniano* — baseado num movimento mecanicista, homogêneo e reversível — ao *tempo bergsoniano* — orgânico, biológico, evolutivo, criativo e irreversível (Bergson, 2011) — identificando neste último um ‘eco filosófico’ das transformações que a mecânica quântica e a mecânica estatística traziam para o entendimento dos fenômenos. Apesar de afirmar que o vitalismo havia ido longe demais em seus “desejos antropomorfizantes” (Wiener, op. cit., p. 61), fica claro que, para Wiener, o *tempo orgânico* bergsoniano (nos termos do filósofo francês a *durée*) estaria mais próximo daquele em que operam as máquinas cibernéticas do que o newtoniano. Ao final daquele capítulo, Wiener afirma que

o autômato moderno existe no mesmo tipo de tempo bergsoniano como organismo vivo; e, portanto, não há razão nas considerações de Bergson para que o modo essencial de funcionamento do organismo vivo não seja o mesmo que a do autômato

¹⁰¹ Enquanto a chamada *cibernética de primeira ordem* tratava os sistemas como entidades independentes do observador, a cibernética de segunda ordem concentrava-se não apenas nos sistemas de controle e comunicação, mas também nos observadores desses sistemas e nas interações reflexivas entre observadores e sistemas.

¹⁰² No original: “Reflexivity is the movement whereby that which has been used to generate a system is made, through a changed perspective, to become part of the system it generates.”

¹⁰³ No original: “Reflexivity tends notoriously toward infinite regress”.

deste tipo. O vitalismo triunfou em tal medida que até mecanismos correspondem à estrutura temporal do vitalismo. (Wiener, op. cit., p. 67, grifo nosso)

Contudo, apesar dessa “incorporação do orgânico” pelas máquinas cibernéticas, Hui (2019, 2020) ressalta que esse traço não implica equivalência entre organismos e máquinas. Para ele, o ponto-chave do *devir orgânico* das máquinas cibernéticas está relacionado, fundamentalmente, ao modo como o mecanismo da *recursividade* lida com a *contingência*.

[a recursividade] se caracteriza pelo movimento em *looping* de retornar a si mesmo para se determinar, enquanto todo movimento está aberto à contingência, que por sua vez determina sua singularidade. Podemos imaginar uma forma espiral, em cada movimento circular, que determina seu devir parcialmente a partir dos movimentos circulares passados.¹⁰⁴ (Hui, 2019, p. 4).

Segundo Hui, a capacidade dos sistemas cibernéticos de integrar recursivamente eventos contingentes (como ruídos e falhas) é o que os diferencia dos sistemas mecânicos e possibilita sua constante transformação. Enquanto no modo de operação mecânico, um evento contingente pode levar ao colapso do sistema, no modo de operação recursivo, a contingência enriquece o sistema e permite que ele se desenvolva. Nesse entender, tanto um organismo vivo quanto um algoritmo de aprendizado de máquina são capazes de ‘absorver a contingência’ tornando-a produtiva. Retomando as reflexões de Hui sobre a temporalidade das máquinas cibernéticas, Parisi e Dixon-Román afirmam que

a automação pode incluir tanto a contingência quanto o acaso em si, porque *a temporalidade do objeto técnico ou das máquinas cibernéticas admite precisamente que erros, incidentes, falhas fazem parte do processo causal do aprendizado de máquina*. Ou seja, as contingências não interrompem o curso, mas expõem precisamente o funcionamento da causalidade não linear¹⁰⁵. (Parisi, Dixon-Román, 2020, grifo nosso)

A contingência, nesse contexto, também está intrinsecamente ligada à ideia de *indeterminação*. Como mencionamos no primeiro capítulo, é justamente a preservação daquilo que Simondon (2008) chama de “margens de indeterminação”, uma abertura às atualizações de seu repertório ou aprendizado inicial, que confere aos algoritmos ditos inteligentes suas propriedades evolutivas. Autoras como Fazi encontrarão justamente numa “ontologia da contingência” (Fazi, 2018, p. 1) ligada à incompletude e indeterminação que é

¹⁰⁴ No original: “it is characterized by the looping movement of returning to itself in order to determine itself, while every movement is open to contingency, which in turn determines its singularity. We can imagine a spiral form, in its every circular movement, which determines its becoming partially from the past circular movements.”

¹⁰⁵ No original: “automation can include both contingency and chance within itself, because the temporality of the technical object or cibernetic machines precisely admits that errors, incident, failure, are part of the causal process of machine learning. Namely, contingencies do not interrupt the course but precisely expose the workings of non-linear causality.”

constitutiva das máquinas computacionais seu grande potencial epistemológico e estético, conforme veremos melhor no último capítulo desta tese.

Outra particularidade da estrutura temporal da recursividade relaciona-se à *cronotopologia não-linear* que causalidade circular produz. Essa cronotopologia atravessa a camada técnica e humana dos ecossistemas digitais produzindo efeitos coemergentes em um mesmo campo dinâmico de complexidades (Cesarino, 2022). É por meio dessa dinâmica que somos integrados individual e coletivamente à temporalidade recursiva das máquinas computacionais resultando num emaranhamento das infraestruturas cognitivas maquínicas e humanas — que, como vimos, ocorre sobretudo num plano não-consciente (Hayles, 2017; Cesarino, 2022; Chun, 2016). Isso torna os efeitos performativos que emergem nesses ecossistemas ainda mais complexos de serem mapeados, uma vez que os limites entre *causa* e *efeito* se tornam borrosos e muitas vezes indiscerníveis. Assim, em sistemas cibernéticos, cognoscentes técnicos e humanos (Hayles, op. cit.), hoje materializados sobretudo em algoritmos e usuários, entram numa relação recursiva de *co-determinação emergente*. Isso não significa, é claro, como já alertamos antes, que esses agenciamentos cognitivos (Hayles, op. cit.) não estejam atravessados por intensas assimetrias de saber e poder. No contexto da plataformização e do capitalismo de vigilância, é essa assimetria que garante que os “loops cada vez mais intensivos entre cognição humana e algorítmica, numa dialética intensiva de produção e de desestruturação de hábitos” extraíam “da cognição elementar dos usuários os efeitos previstos no modelo de negócios das plataformas” (Cesarino, 2021, p. 87).

A cronotopologia não-linear própria dos fenômenos cibernéticos requer, portanto, outro modo de mapear os efeitos da mediação algorítmica, tornando as leituras do tipo *causa-efeito* em grande medida obsoletas aqui. Como pontua Cesarino,

por sua própria natureza cibernética, os algoritmos não controlam os usuários da forma linear com que costumamos pensar as relações de causa e efeito. Seus efeitos sociais, além de complexos e até paradoxais, só podem ser visualizados de forma indireta. (Cesarino, 2022, p. 17)

Numa entrevista a Amoore e Piotukh (2019), Hayles sugere que comportamentos complexos requerem interações recursivas ao mesmo tempo em que arquiteturas recursivas limitam quanto se pode conhecer de tais sistemas, como no caso dos algoritmos de aprendizado profundo. Isso porque, à diferença de um sistema que opera por causalidades lineares (por exemplo, um sistema químico em que os reagentes A e B interagem para formar B e C), num sistema guiado por causalidades recursivas (nesse caso, seguindo o exemplo, os reagentes C e

D também interagem entre eles para formar A e B), a análise do comportamento do sistema se torna muito mais complexa uma vez que todas as partes do sistema interagem simultaneamente afetando-se mutuamente. Dito de outro modo, ao desafiar fechamentos formais, os sistemas recursivos também estão intrinsecamente ligados às restrições que limitam a compreensão precisa do seu próprio funcionamento. Os traços dos sistemas cibernéticos que viemos descrevendo ajudam a compreender por que, as novas mídias, sobretudo as plataformas digitais, têm desencadeado com frequência efeitos sistêmicos imprevisíveis e não intencionados pelos designers dessas tecnologias (Cesarino, op. cit.).

Um último aspecto que gostaríamos de destacar em relação à matriz cronopolítica se refere às defasagens entre a visão original da cibernética sobre o controle dos sistemas através do método do feedback e seus desdobramentos contemporâneos. Como vimos no início deste tópico, na proposição original de Wiener e seus colegas cibernéticos, os *feedbacks negativos* teriam por função ajustar continuamente o comportamento de um sistema de modo a corrigir desvios e mantê-lo em equilíbrio.

No entanto, ao observar o status atual da maioria dos sistemas algorítmicos ou a paisagem informacional das redes digitais de modo geral, o que testemunhamos é a ampla prevalência de retroalimentações positivas ou *feedbacks de autorreforço* que, entre outros efeitos, amplificam padrões históricos discriminatórios, a radicalização e a polarização políticas. No âmbito específico do treinamento de modelos algorítmicos, isso leva à intensificação de padrões (ou, no jargão da área, dos vieses) reconhecidos pela máquina nos bancos de treinamentos. Exemplificando com aplicações de computação visual, Celis descreve assim o efeito: “se uma nuvem se parece um pouco com um pássaro, a rede fará com que ela se pareça mais com um pássaro. Isto, por sua vez, fará com que a rede reconheça o pássaro ainda mais fortemente na próxima iteração e assim por diante, até que um pássaro altamente detalhado apareça, aparentemente do nada.”¹⁰⁶. De modo geral, poderíamos dizer que a maioria dos sistemas cibernéticos contemporâneos, em especial aqueles que sob a lógica econômica da plataformização, encontram-se hoje em *estado de runaway* ou feedback positivo descontrolado (Cesarino, op. cit., Chaney, 2017).

¹⁰⁶ No original: “si una nube se parece un poco a un pájaro, la red hará que se parezca más a un pájaro. Esto a su vez hará que la red reconozca al ave aún más fuertemente en la próxima iteración y así sucesivamente, hasta que aparezca un pájaro altamente detallado, aparentemente de la nada”.

Extrapolando esse ‘problema cibernético’ para a esfera social, Berardi sugere que, no contexto da infoaceleração e hiperconectividade, “a mudança está tomando a forma de feedback positivo” (2020, p. 16): “à medida que a vontade consciente e racional se torna incapaz de conferir e ajustar as tendências, as próprias tendências se reforçam a si mesmas até o ponto do colapso total” (ibid., p. 17). Hayles aborda a questão do seguinte modo ao fazer um certo balanço sobre os erros e acertos do paradigma cibernético:

[m]eio século depois, podemos ver com o benefício da retrospectiva de que maneiras o paradigma cibernético foi tanto profético quanto equivocado. Acertou ao prever que os modos de comunicação entre humanos, formas de vida não-humanas e máquinas se tornariam cada vez mais críticos para o futuro do planeta; estava errado ao pensar que os mecanismos de feedback eram a chave para controlar esse futuro¹⁰⁷. (Hayles, 2017, p. 202)

Dito isso, podemos elencar alguns motivos pelos quais os feedbacks de autorreforço passaram a dominar os ecossistemas digitais contemporâneos, sobretudo aqueles mediados por algoritmos da IA: a complexidade e escala dos sistemas cibernéticos atuais, incomparáveis aos sistemas que serviram de modelo para as proposições originais da cibernética (pense-se, por exemplo, no volume de dados processados por modelos algorítmicos de grande escala, como os grandes modelos de linguagem); o modo como tais sistemas integram agenciamentos cognitivos em que a cognição se torna tão distribuída entre os atores técnicos e humanos e as interações entre eles tão complexas e recursivas que as antigas noções de controle se tornam obsoletas; a crescente aceleração dos ciclos de feedback por meio dos quais os sistemas se atualizam (por exemplo, nos algoritmos de alta frequência do mercado financeiro, que exploraremos a seguir); a crescente autonomia desses sistemas técnicos e a ausência de revisão humana em contextos que envolvem decisões automatizadas; e, principalmente, o desinteresse político e econômico por parte das grandes plataformas digitais em corrigir tais ‘desequilíbrios informacionais’ — afinal, como testemunhamos amargamente nos últimos anos, a polarização política e a instrumentalização do ódio e do ressentimento são altamente lucrativos para seus modelos de negócio¹⁰⁸.

¹⁰⁷ No original: “Half a century later, we can see with the benefit of hindsight in what ways the cybernetic paradigm was both prophetic and misguided. It was correct in anticipating that modes of communication between humans, nonhuman life-forms, and machines would come to be increasingly critical to the future of the planet; it was wrong in thinking that feedback mechanisms were the key to controlling this future.”

¹⁰⁸ Diante da obsolescência da noção de controle sobre tal qual Wiener havia teorizado, Hayles (op. cit.) propõe que um modo mais efetivo de intervenção sobre agenciamentos humano-maquínicos complexos é identificar “pontos de inflexão” (p. 203), pequenas diferenças que, no entanto, podem desencadear efeitos sistêmicos em grande escala e alterar a forma como um sistema se desenvolve no tempo.

Contudo, ainda que a ideia do feedback como mecanismo regulador da estabilidade dos sistemas cibernéticos não tenha ‘vingado’, é inegável como a recursividade e sua causalidade circular sobre a qual viemos refletindo são estruturantes dos regimes de temporalidade algorítmicos. Os ciclos de *feedback* que atualizam ininterruptamente a ‘mega-máquina digital’ funcionam, assim, como um dos principais dispositivos que estruturam o campo do possível que a governamentalidade algorítmica regula, nutrindo a força da sua racionalidade performativa (Bruno, 2021). Para Rouvroy, a “normatividade digital circular torna-se indiscernível dos loops de feedback ou da recursividade no coração do metabolismo algorítmico” (Rouvroy, 2020, p. 17).

No último capítulo da tese, retornaremos à cronopolítica cibernética para refletir mais profundamente como as *repetições recursivas* são uma das marcas dos futuros maquínicos agenciados pelos algoritmos da IA e para explorar a tensão entre a indeterminação enquanto constituinte destas máquinas e modo como elas diminuem a indeterminação dos fluxos de automação nos quais agem.

3.1.2 A matriz militar-securitária: temporalidade antecipatória

As relações históricas entre o complexo industrial-militar e o desenvolvimento das tecnologias computacionais são amplamente conhecidas e já foram bastante abordadas por diversos autores. Como vimos com a história da cibernética, tais relações remontam ao contexto do pós Segunda Guerra e ao investimento massivo em pesquisas que pudessem resultar em vantagem militar por meio da tecnologia, viabilizadas sobretudo através da Agência de Pesquisa de Projetos Avançados do Pentágono (DARPA). É nesse contexto histórico, como se sabe, que surge a própria internet, a cibernética ganha projeção como campo de saber que poderia auxiliar na predição do comportamento humano e maquínico (Wiener, op. cit.) e são produzidos os primeiros modelos conexionistas da IA.

Sem perder de vista esse panorama mais amplo de fundo e o modo como os vínculos entre o complexo industrial-militar e a chamada indústria *tech* foram se renovando ao longo das últimas décadas, influenciando enormemente o desenvolvimento da IA, interessa-nos aqui explorar um momento específico desta história que consideramos crucial para compreender o *caráter antecipatório* da racionalidade algorítmica. Esse caráter manifesta-se principalmente em sistemas baseados em previsão de riscos, como scores de crédito financeiro, sistemas de previsão de crimes e de reincidência criminal, ou sistemas que prevêm ‘riscos sociais’

atrelados a políticas públicas, como gravidez na adolescência¹⁰⁹ e vulnerabilidade social na infância¹¹⁰, etc.

Esse momento específico é o surgimento de um *novo paradigma militar-securitário no pós 11/9* baseado na antecipação ou preempção como fundamento das políticas de segurança da Guerra ao Terror. Como veremos a seguir, a consolidação de uma doutrina militar preemptiva como resposta a um futuro que se mostrava repentinamente altamente imprevisível e ameaçador e diante do qual era necessário prevenir-se, modificou as cronopolíticas militar-securitárias e a abordagem dada ao *desconhecido* (Massumi, 2005). Essas inflexões reordenaram as práticas da vigilância e do gerenciamento de risco (Hong, Szpunar 2019). No âmbito tecnológico, entre outros efeitos, essa reconfiguração impulsionou o desenvolvimento de tecnologias de reconhecimento automatizado de imagens, especialmente no uso de IA em sistemas de videovigilância, e estabeleceu as bases para o surgimento de uma nova lógica de acumulação baseada na vigilância na indústria *tech* (Zuboff, 2021).

É importante pontuar, no entanto, que, apesar de ter se transformado num evento chave para os deslocamentos da relação com o futuro na contemporaneidade, esse ‘futuro incerto’ diante do qual é necessário antecipar-se não foi inaugurado com o 11/9. Ele pode ser visto como parte de uma mutação mais ampla da noção futuro na virada do século. XX para o XXI, questão que abordaremos no último capítulo.

Numa declaração que viria a se tornar emblemática da racionalidade que orientou a nova doutrina, o então Secretário de Defesa do governo Bush, Donald Rumsfeld, ao ser questionado sobre a falta de evidência sobre a existência de armas de destruição em massa no Iraque sob Saddam Hussein, afirmou:

[c]omo sabemos, existem conhecidos conhecidos; há coisas que sabemos que sabemos. Também sabemos que existem desconhecidos conhecidos; isto é, sabemos que há algumas coisas que não sabemos. Mas também existem os *desconhecidos desconhecidos* – aqueles que não sabemos que não sabemos¹¹¹ (DARPA, 2002)

¹⁰⁹ Ver, por exemplo, o *Projeto Horus*, que teve pilotos implementados na Argentina e no Brasil, ambos em parcerias do poder público com a Microsoft:
<https://notmy.ai/pt/noticias/gravidez-na-adolescencia-abordada-pelo-colonialismo-de-dados-de-um-sistema-que-e-patriarcal-desde-o-projeto/>

¹¹⁰ Ver, por exemplo, o *Projeto Alerta Niñez*, que teve piloto implementado no Chile:
<https://notmy.ai/pt/project-item/sistema-alerta-ninez-san-pt/>

¹¹¹ No original: “As we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say, we know there are some things we do not know. But there are also unknown unknowns – the ones we don’t know we don’t know.”.

A partir daquele momento, quando a invasão do Iraque era iminente, a implementação de táticas e técnicas preditivas para caça dos “desconhecidos-desconhecidos” de Rumsfeld se tornou a base de uma abordagem que visava identificar e eliminar a ameaça antes que se concretizasse. Essa abordagem orientou desde o paradigma militar da *Inteligência Baseada em Atividade*, que substituiu o reconhecimento de identidades nominais do inimigo pelo reconhecimento de padrões de comportamentos suspeitos¹¹² (Chamayou, 2015), até o desenvolvimento de sistemas de videovigilância automatizada para os drones. Dispositivos privilegiados da “política de eliminação profilática” (ibid., p. 44) posta em ação na Guerra ao Terror, as ferramentas de computação visual dos drones foram viabilizadas por estreitas colaborações entre empresas de tecnologia e o Departamento de Defesa dos Estados Unidos, expressas, por exemplo, no Projeto *Maven*¹¹³.

De modo mais amplo, a radical mutação nas políticas e práticas da vigilância pós 11/9 assentou as bases epistemológicas para a consolidação do modelo econômico do capitalismo de vigilância a partir da primeira década do século XX. Como aponta Zuboff (2021) sobre este período, o “excepcionalismo de vigilância” (p. 138) foi uma condição fundamental para o desenvolvimento de uma nova lógica de acumulação baseada em vigilância por parte da Google num território de franca desregulação jurídica e burocrática. Numa convergência de interesse entre agências estatais de inteligência e empresas de tecnologia, que culminaria no escândalo da NSA em 2013, o novo modelo de conhecimento e vigilância “baseado em atividades” e na coleta massiva de dados comportamentais logo viria a se espriar do domínio securitário-militar para o comercial. Esse modelo ditou o modo como, progressivamente, os comportamentos dos usuários passaram a ser monitorados e modulados em função de seus rastros de atividade. Zuboff também destaca como esse ambiente legitimou um novo imperativo de desenvolvimento que se contrapunha à temporalidade típica dos processos democráticos e que pode ser resumido na máxima “*rapidez a qualquer custo*” (Zuboff, op. cit. 138, grifo nosso).

¹¹² Tal diferenciação separava os chamados *ataques de personalidade* dos *ataques de assinatura* (Chamayou, op. cit.). Enquanto nos ataques de personalidade as identidades eram conhecidas, os ataques de assinatura visavam indivíduos com identidade desconhecida, mas cujos padrões de vida “leva a supor, indica ou *assina* o pertencimento a uma organização terrorista” (ibid., p. 57).

¹¹³ O Projeto Maven foi uma iniciativa do Departamento de Defesa dos Estados Unidos que visava incorporar IA em operações militares. A Google participou do projeto entre 2017 e 2018 fornecendo tecnologia de IA para análise automatizada de imagens capturadas por drones. O objetivo era utilizar algoritmos de aprendizado de máquina para identificar objetos e padrões em grandes volumes de dados visuais.

Numa proposição que obteve ampla adesão, Massumi (op. cit.) denomina a lógica desse novo paradigma militar-securitário de *lógica operativa da preempção*: “[a] preempção é um conceito temporal. Ela denota *agir no tempo anterior*: no tempo da ameaça, antes que ela surja como um perigo claro e presente”¹¹⁴ (ibid., p. 7, grifo nosso). A preempção *traz o futuro para o presente*, tornando tangíveis as consequências de uma eventualidade que pode ou não ocorrer e que é indiferente, de fato, a sua ocorrência efetiva. Para o autor, a preempção se converteu numa força motriz que modificou as relações de poder excedendo os limites da guerra e transformando o mundo num ambiente de ameaças generalizado. As estratégias e políticas de segurança que se seguiram ao 11/9 e dominaram o discurso público e político forjaram assim uma nova relação entre o *desconhecido* e a *ameaça*. Nessa nova concepção, o risco de emergência das ameaças deve ser permanentemente calculado e antecipado pelas tecnologias do controle e da vigilância. Uma outra declaração de Rumsfeld sintetiza bem essa racionalidade cronopolítica:

[nosso desafio neste novo século é difícil: defender nossa nação contra o desconhecido, o incerto, o invisível e o inesperado...Devemos deixar de lado formas confortáveis de pensar e planejar...para que possamos deter e derrotar adversários que ainda não surgiram para nos desafiar.¹¹⁵ (Rumsfeld apud Massumi, 2005, p. 246)

Em sua análise das mutações nas práticas securitárias pós 11/9, Amoore (2013) observa como a busca e captura do “elemento desconhecido” se tornou o imperativo de um modelo de governamentalidade antecipatória que fez convergir as esferas da economia e da segurança. Essa convergência de imaginários e técnicas de governo dos futuros entre estas duas esferas, segundo Amoore, se deu sobretudo por meio das *tecnologias do risco*.

A “política da possibilidade” (ibid.) centrada no risco aborda uma série aparentemente ilimitada de corpos, populações, espaços, transações financeiras, movimentos, etc de acordo com os graus de risco que oferecem. Como observa, muito mais que um evento isolado tangível, o risco é uma entidade performativa que produz os efeitos que nomeia. Para a autora, as tecnologias do risco carregam sempre “a promessa de administrar a incerteza e

¹¹⁴ No original: “Preemption is a time concept. It denotes acting on the time before: the time of threat, before it has emerged as a clear and present danger”.

¹¹⁵ No original: ““Our challenge in this new century is a difficult one: to defend our nation against the unknown, the uncertain, the unseen, and the unexpected... We must put aside comfortable ways of thinking and planning...so we can deter and defeat adversaries that have not yet emerged to challenge us”

tornar um futuro incognoscível e indeterminado cognoscível e calculável”¹¹⁶ (ibid., p. 7). As chamadas *sociedades do risco* (Beck, 2011) apontam, assim, sempre para a tensão entre os temores diante de um futuro desconhecido e indeterminado e a tentativa de *controlar antecipadamente essa imprevisibilidade*. No contexto financeiro, como veremos no próximo tópico, as contradições produzidas por essa relação particular com o futuro se tornam ainda mais acentuadas.

À diferença da estatística moderna, ancorada no conhecimento sobre médias e populações, Amoore destaca como as tecnologias do risco consolidadas no pós 11/9 operavam sobre “um sujeito fracionado cujos elementos de risco o dividem até dentro de si”¹¹⁷ (Amoore, op. cit., p. 8) e no qual a autora identifica ecos do *dividual* deleuziano (Deleuze, 1992). Visando agir sobre futuros incertos, o risco fraciona o sujeito em múltiplos graus que indicam quão perigoso, vulnerável, móvel, etc ele *potencialmente* é.

Massumi (2005) enfatiza que um aspecto fundamental do regime da preempção é que a incerteza do potencial porvir nunca é consumida totalmente por nenhum evento. Sempre resta um excedente não consumado de perigo que se renova a cada ação preventiva e tece essa cadeia de ações e reações. Assim, seu funcionamento opera a partir de uma causa virtual cujo potencial nenhuma atualização esgota. De modo similar, Hong e Szpunar (2019), ao caracterizarem a *razão antecipatória* vinculada à futurização da segurança, apontam que, nesse modelo cronopolítico, toda atualização (seja um acidente, uma prisão preventiva ou o tão temido ‘próximo ataque terrorista’) opera como uma *instância não exaustiva do futuro*. Nesse contexto, o futuro necessita permanecer como um repositório inesgotável de incertezas alimentando a expansão indefinida dos aparatos de vigilância.

Os autores também ressaltam a dimensão performativa da razão antecipatória e das práticas e técnicas que lhe servem de instrumento, enfatizando que o valor dessa “segurança orientada para o futuro” (Hong, Szpunar, op. cit., p. 315) não está na sua capacidade de prever com precisão eventos específicos, mas em construir verdades especulativas. “O significado político e moral desse futuro não reside no cumprimento da previsão, mas nas ações

¹¹⁶ No original: “the promise of managing uncertainty and making an unknowable and indeterminate future knowable and calculable”.

¹¹⁷ No original: “a fractionated subject whose risk elements divide her even within herself”.

concretas que são tomadas no presente em nome dessa previsão”¹¹⁸ (ibid.). Visando impedir a atualização de futuros indesejáveis, as previsões criadas nesse contexto permanecem, em sua maioria, inverificáveis. Ao analisar o regime temporal das câmeras inteligentes de vigilância, Bruno observa um fenômeno semelhante, afirmando que, nesse contexto, “a redução [das incertezas] não se dá por acuidade na previsão, mas pela performatividade da antecipação” (Bruno, 2012, p. 58). Essa intervenção sobre “o que ainda não ocorreu” (ibid., p. 61) desvincula ação e acontecimento, uma vez que nunca se trata da referencialidade indicial do “isso é” ou “está sendo” (ibid.). Assim, “a validade epistêmica e normativa da razão antecipatória não deve estar ligada à ficção útil da ‘precisão’ contra o futuro como um horizonte infalsificável e inesgotável, mas às consequências políticas e morais dessas ações preditivas no presente”¹¹⁹ (Hong, Szpunar, op. cit., p. 326). Deste modo, a principal função estratégica dos futuros produzidos pela razão antecipatória, destacam os autores, reside na produção de *bens informativos negociáveis* sobre o futuro, transformando-o num locus de especulação. Essa dimensão é especialmente evidente na financeirização, matriz cronopolítica sobre a qual nos deteremos no tópico a seguir.

3.1.3 A matriz financeira: temporalidade acelerada e os mercados de futuros

Os elos entre neoliberalismo, financeirização e digitalização são estreitos, complexos e multifacetados. Manifestam-se tanto numa *dimensão político-ideológica* — como nas explícitas ressonâncias entre a chamada *Ideologia Californiana*¹²⁰ (Cf. Barbrook, Cameron 1995; Morozov, 2018; Wark, 2022) e a “constelação de princípios, políticas, práticas e formas de governar a razão reunidas sob o signo de neoliberalismo” (Brown, 2019, p. 19) — quanto numa *dimensão econômica-material* — como no modo como as novas mídias se

¹¹⁸ No original: “The political and moral significance of this future lies not in the fulfilment of prediction, but in the concrete actions that are taken in the present in the name of prediction.”

¹¹⁹ No original: “The epistemic and normative validity of anticipatory reason must be tethered not to the useful fiction of ‘accuracy’ against the future as an unfalsifiable and inexhaustible horizon, but to the political and moral consequences of those predictive actions in the present”

¹²⁰ O termo Ideologia Californiana foi cunhado por Richard Barbrook e Andy Cameron num ensaio publicado em 1995. De acordo com os autores, a Ideologia Californiana é um sistema de crenças que combina as idéias contraculturais dos anos 60 com o determinismo tecnológico do Vale do Silício. Essa ideologia promove a noção de que as novas tecnologias de informação têm o potencial de produzir uma nova ordem mundial de maior democracia, liberdade e empoderamento individual ao mesmo tempo em que celebra o mercado e o empreendedorismo como chave para liberar esse potencial. Para Wark (2022), a Ideologia Californiana propunha um mundo em que a tecnologia era a única força transformadora da história e que tinha como herói o empresário. “O herói desse mito épico-poético foi o empresário, que sozinho luta contra o trabalho, o Estado e a cultura, para liberar a força supostamente 'natural' da tecnologia” (p. 99).

converteram na arquitetura digital do neoliberalismo (Cesarino, 2021), e no papel fundamental do capital de risco¹²¹ no financiamento da indústria *tech*. Autores como Mbembe (2021) e Chun (2016), inclusive, propõem entender tais processos como ‘apenas um’. Para o primeiro, mecanismos computacionais, modelagem algorítmica e a expansão totalizante do capital sob a égide do neoliberalismo integram um único processo que visa submeter todas as dimensões da vida a processos de quantificação e abstração (Mbembe, op. cit.).

Mas apesar de hoje operarem sistemicamente integrados e co-dependentes, é importante recordar que a convergência entre neoliberalismo e digitalização foi condição de possibilidade para a consolidação da oligarquia internacional que hoje controla os fluxos globais de capital e para a transformação das finanças num dispositivo de governança transversal à produção, à política, ao bem-estar e ao consumo (Lazzarato, 2017).

Por um lado, o triunfo das políticas econômicas e da racionalidade neoliberal nas últimas quatro décadas sob os preceitos de Hayek — que transformou princípios de mercado em princípios de governo (Brown, op. cit., Foucault, 2008b) — promoveu a desregulação do capital e permitiu que as instituições financeiras passassem a operar praticamente imunes a qualquer interesse a maximização dos próprios lucros. De um modelo capitalista industrial centrado na relação *capital/trabalho*, passa-se então a um modelo centrado na relação *credor/devedor* (Lazzarato, op. cit.)¹²² no qual todos os aspectos da existência são submetidos a cálculos de investimento sobre seu valor futuro (Brown, op. cit.) sobretudo sob a forma de *derivativos*¹²³.

Ao mesmo tempo, tal processo seria inviável sem o surgimento de uma infraestrutura computacional que possibilitou a virtualização dos mercados financeiros nos anos 1980 e 1990 (Bridle, 2019). Fazendo um contraponto ao que entende como leituras ‘puramente ideológicas’ da consolidação do neoliberalismo, Wark (op. cit.) ressalta como a reorganização

¹²¹ O capital de risco é uma modalidade de investimento de capital privado fornecido a empresas em estágios iniciais consideradas promissoras e com grande potencial de crescimento. O capital de risco investe nessas empresas em troca de uma participação acionária, geralmente na forma de capital próprio ou dívida conversível.

¹²² Em *O governo do homem endividado*, Lazzarato (2017) identifica a questão do endividamento como central na mutação histórica dos aparelhos de captura do neoliberalismo contemporâneo. Para o autor, a dívida seria a técnica mais adequada para a produção do *homo oeconomicus* neoliberal (Foucault, 2008b) e a figura do homem endividado corresponderia à produção subjetiva da crise e da dívida.

¹²³ Derivativos são contratos sobre os preços futuros de vários tipos de ativos financeiros negociados nos chamados mercados futuros (*future markets*, em inglês) e os instrumentos centrais para gerenciamento de risco da finança contemporânea.

das forças produtivas em torno da informação nos anos 1970 e o surgimento de uma vasta infraestrutura para a circulação global de fluxos de dinheiro, máquinas e mão de obra foi o que tornou possível globalizar o setor bancário e construir vastas cadeias internacionais de suprimentos. Em paralelo, o crescimento da potência dos processadores de dados e o desenvolvimento de algoritmos capazes de realizar milhares de transações em frações de segundos, os chamados *Algoritmos de Alta Frequência (High Frequency Trading* ou apenas HFT, em inglês), aumentaram vertiginosamente a velocidade das transações financeiras, assim como a abstração, a opacidade, a liquidez e a imprevisibilidade dessas operações. Esses algoritmos utilizam complexos modelos matemáticos para analisar dados financeiros em tempo real, identificando padrões e tendências e lucrando massivamente com pequenas flutuações nos preços dos ativos. Explorando estratégias de curto prazo, os robôs HFT são capazes de criar condições artificiais para manipular preços, transformando os mercados numa vasta e poderosa *infraestrutura especulativa*.

A substituição dos ‘corretores de carne e osso’ por corretores automatizados que funcionam de modo distribuído, obscuro e inescrutável foi fundamental para transformar o mercado numa caixa-preta de difícil compreensão para o cidadão comum e blindar seu *modus operandi* de questionamentos e intervenções orientados pelo interesse público. “No tumulto dos mercados”, afirma Bridle, “raramente ficava claro quem operava os algoritmos; e hoje não é diferente porque sua tática primária é clandestina: mascarar intenções e origens enquanto eles capturam uma porção considerável de todos os valores transacionados” (op. cit., p. 181).

A digitalização e algoritmização dos mercados financeiros são a plena materialização das proposições de Hayek sobre o mercado como mero processador de informações cuja ordem espontânea emergiria de forma tácita e descentralizada (Beverungen, 2019; Pasquinelli, 2021, Halpern, 2022). Na concepção do ‘pai neoliberalismo’, o conhecimento sobre o mercado “nunca existe de forma concentrada ou integrada, mas apenas como fragmentos dispersos de conhecimento incompleto e frequentemente contraditório que todos os indivíduos separados possuem”¹²⁴ (Hayek apud Halpern, op. cit. p, 334). Conforme vimos no Capítulo 1, as proposições de Hayek sobre modelos de mercado e de mente distribuídos, adaptativos e auto-governados, cujo funcionamento não se submeteria a uma consciência ou controle

¹²⁴ No original: “never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess”.

central, também influenciaram o desenvolvimento das máquinas conexionistas nos anos 60 (Pasquinelli, 2021) (ver item 1.4.2).

A virtualização transformou completamente os regimes de temporalidade financeiros, uma vez que os HFT introduziram uma *lacuna temporal* intransponível entre a cognição técnica e humana (Hayles, 2017; Beverungen, op. cit.). A “corrida armamentista da velocidade”¹²⁵ (Hayles, op. cit., p. 165) desencadeada pelo trading algorítmico resultou num domínio sociotécnico inédito de (suposta) autonomia técnica, risco sistêmico e imprevisibilidade. Hayles observa como na financeirização, “as temporalidades complexas inerentes aos derivativos interagem a temporalidade alterada dos HFT para aumentar ainda mais a fragilidade dos mercados financeiros e sua vulnerabilidade a retroalimentações e dinâmicas auto-amplificadoras”¹²⁶ (op. cit., p. 142). Como se sabe, essa dinâmica pode levar a falhas sistêmicas catastróficas, como ocorreu na crise financeira de 2008 e no *Flash Crash* de 2010.

Tal como as tecnologias do risco que emergiram no pós 9/11 (Amoore, 2013), tratam-se de ferramentas que respondem à crescente instabilidade do mundo. Esposito, ao analisar “o futuro dos futuros” da financeirização, afirma que “[o]s derivativos se voltam para o futuro, um futuro que eles sabem que não conhecem, e prometem proteger contra riscos. Eles prometem lidar no presente com o medo do futuro imprevisível (que continua imprevisível)”¹²⁷ (Esposito, 2011, p. 1). São instrumentos utilizados para gerir o tempo sob a forma do risco¹²⁸ (ibid.) e lucrar com as defasagens entre as expectativas sobre o futuro e aquilo que efetivamente acontece — ao mesmo tempo em que modificam a relação entre estes termos. Por este motivo, a autora argumenta que os derivativos podem ser considerados

¹²⁵ No original: “arms speed race”.

¹²⁶ No original: “The complex temporalities inherent in derivatives interact with the changed temporality of HFT to increase further the fragility of financial markets and their vulnerability to feedback loops and self-amplifying dynamics.”

¹²⁷ No original: “Derivatives turn to the future, a future that they know they don’t know, and promise to protect against risks. They promise to deal in the present with the fear of the unpredictable future (which remains unpredictable).”

¹²⁸ É importante lembrar que o cálculo e a comercialização de riscos não são algo novo e sua história enquanto ferramenta que pretende dar mais controle sobre o futuro remete ainda ao período das navegações marítimas e à própria história da probabilidade (Cf. Appadurai, 2016; Bernstein, 2011) sobre a qual já comentamos em alguns pontos desta tese. Contudo, como observa Esposito (op. cit.), o que é novo contemporaneamente é o desenvolvimento de *mercados que lidam estritamente com o risco*, que o comensuraram e comodificaram. Para uma história do risco e sua relação com a concepção moderna de futuro, ver Bernstein, 2011.

“*técnicas de des-futurização*”¹²⁹ (ibid., p. 180). Paradoxalmente, o uso massivo e desregulado dessas (que são também indiscutivelmente) tecnologias de governo do tempo tende a acentuar os problemas aos quais respondem, injetando no presente um futuro ainda mais imprevisível e permeado por riscos sistêmicos.

O ponto de inflexão crucial na trajetória das tensões e complexidades que marcam a lógica econômica, política e sociotécnica da financeirização foi, sem dúvida, o colapso econômico global de 2007-2008, conhecida como a crise dos *subprimes*. Embora não iremos nos aprofundar aqui nas interpretações sobre a crise¹³⁰, é válido ressaltar como a crise financeira desencadeada pelo colapso do mercado imobiliário americano pode ser considerada uma falha sistêmica que revelou simultaneamente a lógica subjacente a esse modelo econômico e sua inviabilidade, sobretudo do endividamento insustentável que o sustenta. Como sintetiza Steyerl:

A financeirização introduz uma série de complicações: a viabilidade econômica do presente é sustentada pela dívida, ou seja, pela renda futura reivindicada, consumida ou gasta no presente. Assim, por um lado, o futuro é esgotado e, por outro, o presente é desestabilizado. Em suma, o presente parece ser constituído pelo esvaziamento do futuro para sustentar uma versão em loop de um passado que nunca existiu. O que significa que, pelo menos em partes dessa trajetória, o tempo realmente corre para trás, de um futuro esvaziado para nutrir um passado imaginário estagnado, sustentado por um design disruptivo.¹³¹ (Steyerl, 2017, p. 37).

Enquanto a crise securitária pós 11/9 sedimentou as bases primeiras para o nascente capitalismo de vigilância (Zuboff, op. cit.), a crise financeira coincidiu historicamente com o momento a partir do qual o modelo de negócios baseado no extrativismo de dados se consolida definitivamente num amplo contexto de plataformização da web (Helmond, 2015, Srnicek, 2017). Esse processo resultou no crescente uso de técnicas e práticas de perfilização, predição e influência massiva dos comportamentos dos usuários. Também foi o momento em que a ascensão das forças da extrema direita pela convergência ultraliberal-reacionária (Cesarino, 2022) — em grande parte uma resposta à experiência de abandono, traição e raiva

¹²⁹ No original: “de-futurization techniques”.

¹³⁰ Nesse sentido, ver, por exemplo, Lazzarato, 2017; Appadurai, 2006; Ferreira da Silva, 2024.

¹³¹ No original: “Financialization introduces a host of further complications: the economic viability of the present is sustained by debt, that is, by future income claimed, consumed, or spent in the present. Thus on the one hand futures are depleted, and on the other, presents are destabilized. In short, the present feels as if it is constituted by emptying out the future to sustain a looping version of a past that never existed. Which means that for at least parts of this trajectory, time indeed runs backwards, from an emptied-out future to nurturing a stagnant imaginary past, sustained by disruptive design”.

por parte dos novos despossuídos (Brown, op. cit.) — começa a ganhar consistência e, alguns anos depois, será alavancada pelo uso dessas mesmas técnicas e práticas.

Berardi (2019) e Fisher (2019, 2022) ressaltam como a crise de 2008 produziu uma ampla crise na *capacidade de imaginar o futuro*. Esse “lento cancelamento do futuro”¹³² (Berardi apud Fisher, 2022, p. 24), é claro, não ocorreu da noite pro dia. Foi o resultado da “gradual, mas implacável, corrosão do futuro” (Fisher, 2022, p. 32) causada por décadas de hegemonia neoliberal e da infiltração da “ontologia dos negócios” (Fisher, 2019) em todas as esferas da vida, mas atinge seu ápice naquele momento. Apesar de algumas interpretações iniciais (que agora soam extremamente otimistas) sobre os indícios do fim do capitalismo diante do colapso financeiro, os resgates bancários se tornaram garantia da persistência daquilo que Fisher chama de *realismo capitalista*: o desaparecimento de alternativas ao capitalismo neoliberal sintetizado no conhecido adágio de que “é mais fácil imaginar o fim do mundo do que o fim do capitalismo” (ibid.)

Retomaremos e aprofundaremos a discussão sobre a crise do futuro no último capítulo desta tese. Basta aqui pontuar como a cronopolítica financeira no contexto neoliberal que abordamos ao longo deste tópico, em larga medida fruto do modo como instrumentos financeiros complexos governam o tempo e produzem *futuros imprevisíveis e instáveis*, se relaciona ao cancelamento de um horizonte comum na esfera social.

Apesar de todo o panorama que viemos descrevendo, à luz dos vínculos entre neoliberalismo, financeirização e digitalização, possuir múltiplas relações com o problema do governo algorítmico do tempo, convém enfatizar as principais razões pelas quais consideramos a financeirização uma das matrizes cronopolíticas do regime de temporalidade algorítmico. A preponderância da financeirização como matriz cronopolítica que destacamos não é necessariamente da ordem de uma relação mais direta (como a que existe entre o capital de risco e a IA, por exemplo), mas da migração de uma certa *lógica financeirizada*, especialmente no modo de governar o tempo e o futuro, para outros domínios, incluindo o digital.

¹³² Apesar da expressão de Berardi sobre o “cancelamento do futuro” ter se tornado muito conhecida, muitas vezes atribuída ao próprio Fisher, ela não está presente *ipsis litteris* na versão em português do livro *After the Future* (Berardi, 2011), onde a ideia figura inicialmente. Ainda assim, o sentido da expressão atravessa toda a obra e é largamente abordado pelo autor.

Em primeiro lugar, conforme comentamos, a financeirização, entrelaçada à racionalidade neoliberal, visa subsumir o máximo de aspectos da existência a *cálculos de investimento sobre seu valor futuro* (Brown, 2019). Ora, é exatamente isso que está em jogo quando nossos comportamentos (sobretudo aqueles relacionados a nossas vulnerabilidades, tendências e potencialidades futuras) são convertidos em dados e interpretados por algoritmos para alimentar “mercados de comportamentos futuros” (Zuboff, 2021). Seja no âmbito estritamente financeiro ou das plataformas digitais, o que está em jogo nesses *mercados de futuros* é a transformação do futuro num território especulativo de bens negociáveis (Hong, Szpunar, 2019). Afinal, como pontua Rouvroy (2022), um dos grandes objetivos da governamentalidade algorítmica sob o capitalismo digital é a transformação das virtualidades do futuro em mais-valia.

Outro fator fundamental da cronopolítica financeira é a aceleração sociotécnica sem precedentes, materializada especialmente nas microtemporalidades aceleradas e fragmentadas dos HFT. Uma de nossas hipóteses é que essa aceleração tecnofinanceira vem se espalhando para outros domínios nos últimos anos, modulando a temporalidade dos ecossistemas digitais e transformando-os em *ambientes de alta frequência*, aspecto que exploraremos especialmente no estudo de caso do TikTok (ver item 3.3.2). Associada a outros fatores, como os imperativos da chamada economia da atenção (Bentes, 2022) e a temporalidade da crise permanente que move as redes (Chun, 2016), argumentamos que a financeirização pode ser considerada uma das principais matrizes da radical aceleração e fragmentação das mediações algorítmicas na atualidade, bem como de seus efeitos sobre nossa experiência, cognição e modos de subjetivação nos ambientes digitais. Conforme veremos no supracitado estudo de caso, essa *fractalização acelerada do tempo das redes* também possibilita a aceleração das lógicas extrativas que sustentam o capitalismo orientado por dados.

Finalmente, a matriz financeira revela de modo privilegiado uma relação performativa com o tempo, similar à que encontramos no governo algorítmico do tempo e nas predições que lhe servem de instrumento. Muito mais voltadas para os efeitos que podem produzir sobre a realidade do que para a acurácia de suas mediações, as tecnologias de governo do tempo utilizadas em ambas esferas relacionam-se a uma “des-futurização” (Espósito, op. cit.) ou a um “sequestro do futuro” (Bruno, 2020) que, simultânea e paradoxalmente, diminuem a indeterminação do futuro e desencadeiam efeitos imprevisíveis sobre os ecossistemas sociotécnicos em que agem.

3.2 Temporalidades algorítmicas: dois estudos de caso

Os dois estudos de caso que seguem se basearam em métodos exploratórios e foram realizados, em sua maior parte, durante o segundo semestre de 2023. No caso do TikTok, essa exploração se deu sobretudo através de uma simulação de *jornada de uma usuária*¹³³ que estivesse utilizando o app pela primeira vez. A partir disso, identificamos algumas questões-chave de ordem cronopolítica que organizaram a estrutura do estudo (sobretudo a parte dedicada a seu regime de temporalidade). No caso do ChatGPT¹³⁴, além da experiência e observação direta de uso, foram utilizados diversos relatórios técnicos (a maioria da própria OpenAI) que relatam como o modelo GPT foi construído e treinado.

3.2.1 Predição, aceleração, fragmentação: TikTok como ambiente de alta frequência

Lançado em setembro de 2016 sob o nome de *Douyin*¹³⁵ pela gigante de tecnologia chinesa *ByteDance* e voltado inicialmente apenas para o mercado interno, o TikTok iniciou sua fase de internacionalização em 2018 a partir da aquisição e fusão com o *Musical.ly*, um app de vídeos curtos conhecido por seus recursos para dublagem musical. Com seu *boom* durante a pandemia de COVID-19, quando as pessoas estavam especialmente ávidas por distração, o app se transformou rapidamente em um fenômeno, especialmente entre adolescentes, atingindo a marca de um bilhão de usuários ativos globalmente em setembro de 2021 (TikTok, 2021). Atualmente, o TikTok tem cerca de 1,5 bilhões de usuários (Figura 5) dos quais mais de 65,5 milhões estão no Brasil¹³⁶. Com isso, se tornou o primeiro app de grande projeção mundial que não provém do Vale do Silício e (mais um) pivô das crescentes tensões geopolíticas entre EUA e China¹³⁷. Além do número total de usuários, o TikTok também

¹³³ No contexto de UX design refere-se ao caminho completo que um/a usuária/o percorre ao interagir com um produto ou serviço.

¹³⁴ As respostas fornecidas pelo chat que utilizamos no texto são seguidas sempre da data, uma vez que, pelas atualizações ininterruptas da aplicação, muitas vezes, pouco tempo depois o mesmo prompt já não fornece respostas similares.

¹³⁵ Essa segue sendo a versão do app utilizada na China.

¹³⁶ Dados da data.ai para o período entre 26 de maio e 1 de junho de 2024. Disponível em: <https://www.data.ai/en/apps/google-play/app/com.zhiliaoapp.musically/>. Acesso em 11 de junho de 2024.

¹³⁷ Após muitos embates, o presidente dos EUA, Joe Biden, sancionou em 24/04/2024 um projeto de lei que determina a proibição do app no país, a menos que a ByteDance venda a operação americana da empresa em um prazo de até nove meses.

impressiona por duas métricas-chave para aferir o grau de engajamento destes: *retenção*¹³⁸ e *tempo médio de uso*. Segundo dados da data.ai (antiga *AppAnnie*) referentes a 2023, o TikTok é a rede social com a maior média mensal de tempo gasto por usuário¹³⁹, chegando a 34 horas por mês e ultrapassando o YouTube, outra rede conhecida por sua capacidade de engajamento (Figura 4). Contudo, quanto ao número total de usuários, o TikTok ocupa somente o sexto lugar do ranking (Figura 5).

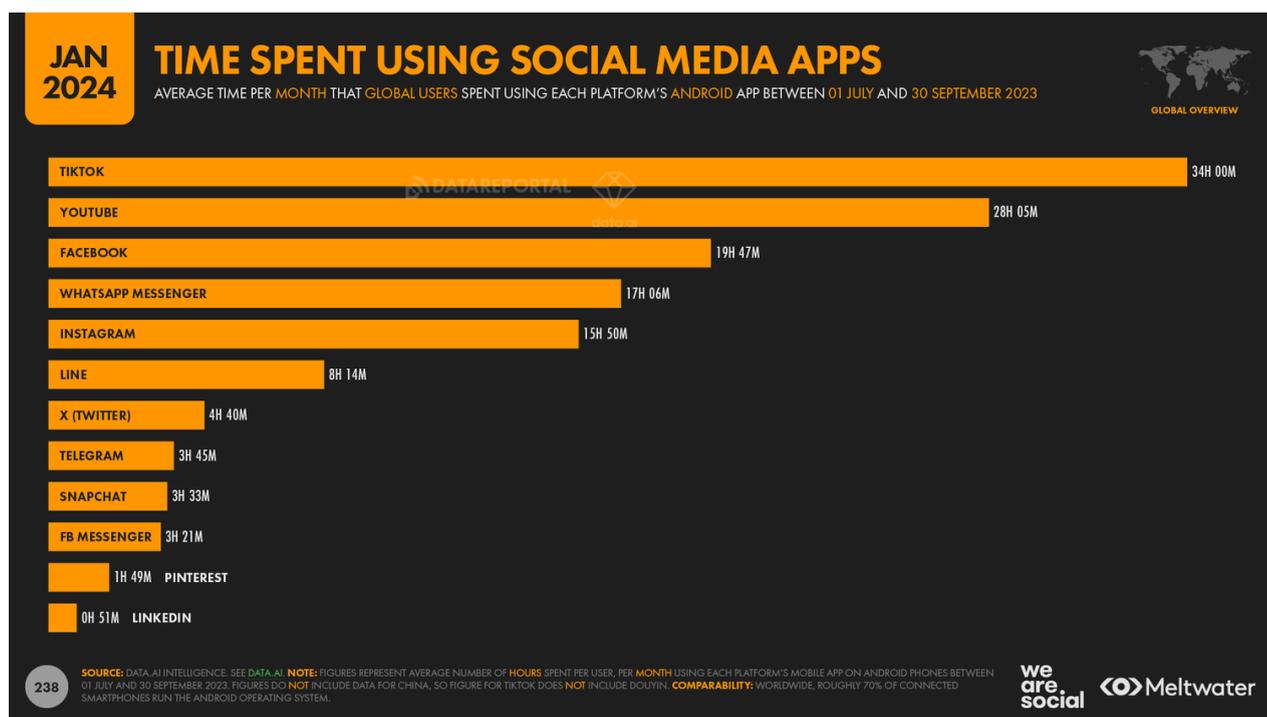


Figura 4: Ranking de plataformas sociais por tempo de uso referentes ao período entre junho e setembro de 2023 (horas/mês). Fonte: DataReportal com dados da data.ai.

¹³⁸ Retenção se refere à taxa de usuários que continuam a usar o app após a instalação.

¹³⁹ Para o sistema Android.

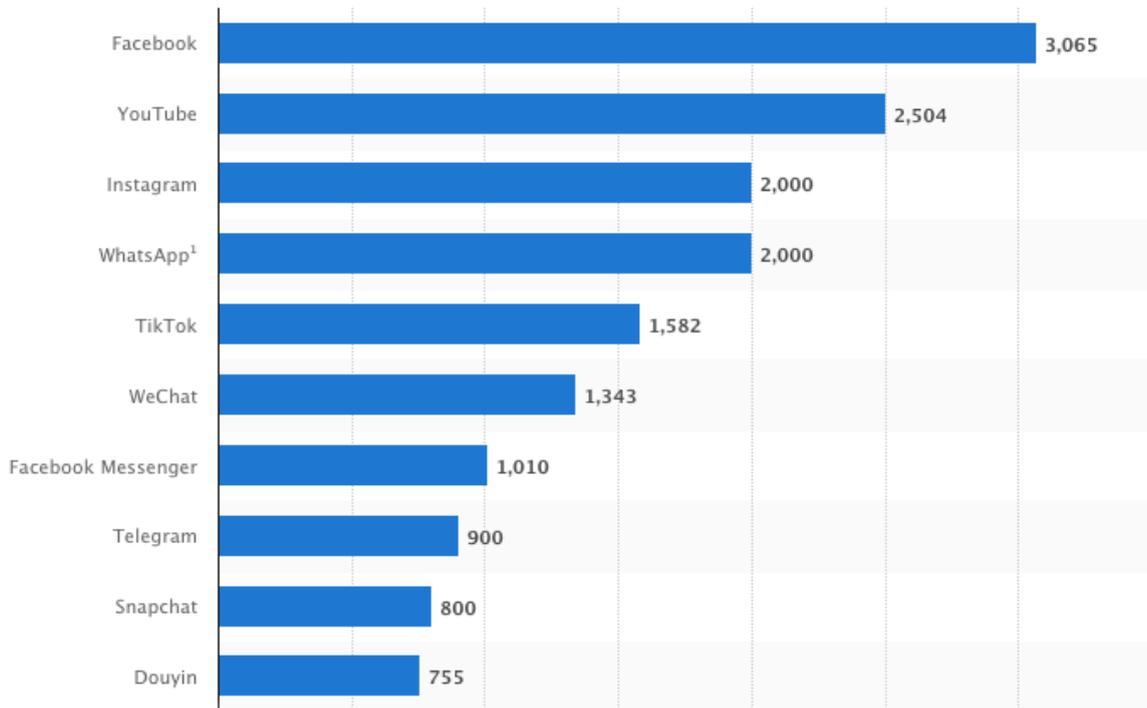


Figura 5: Ranking de plataformas sociais por número de usuários ativos em abril de 2024.
Fonte: Statista.

Há várias funcionalidades e *affordances*¹⁴⁰ ajudam a explicar o sucesso do TikTok, desde a facilidade para criação de conteúdo por não profissionais — uma vez que disponibiliza uma vasta biblioteca de recursos como músicas e filtros para o usuário dublar, encenar e/ou dançar, resultando em vídeos considerados divertidos e que produziram o fenômeno ‘dancinhas do TikTok’, coreografias simples e curtas que se tornaram altamente virais — até a simplicidade de interação da interface (é praticamente só ‘arrastar pra cima’), propiciando uma experiência de uso simples, imersiva e fluida, facilmente convertida em *hábito* ou mesmo em *vício* (Chun, 2016).

Mas o principal diferencial do TikTok para “engajar e enganchar” (Bentes, 2022) seus usuários é a aba *Para Você*, o *feed* hiperpersonalizado e tela padrão ao abrir o app (Figura 6). Descrito pela própria empresa como “central para a experiência TikTok e onde a maioria dos

¹⁴⁰ *Affordance* é um conceito importante no design de interfaces proposto originalmente pelo psicólogo James J. Gibson na década de 1970 e posteriormente popularizado pelo designer de usabilidade Don Norman. Em termos simples, refere-se à percepção intuitiva que os usuários têm das ações possíveis que um objeto ou elemento de interface oferece com base em sua aparência física. “É o relacionamento entre as propriedades do objeto e as capacidades do agente de determinar intuitivamente, ou baseado em experiências anteriores, como o objeto pode ser usado, sem necessidade de explicações, rótulos ou instruções.” (Mesquita, 2018)

nossos usuários passa o tempo”¹⁴¹ (TikTok, 2021), o algoritmo do *Para Você* se tornou conhecido por ser estranhamente bom em prever com extrema rapidez quais vídeos ‘fisgarão’ os usuários, despertando muitas especulações sobre qual ‘o segredo’ de seu viciante algoritmo. Com supostos poderes de “ler a mente dos usuários” (Smith, 2021), parece haver uma crença generalizada nas análises que circulam na mídia de que, comparado a outras redes, o TikTok possui o algoritmo mais “agressivo” e “viciante” (Siles, Valerio-Alfaro, Meléndez-Moran, 2022).

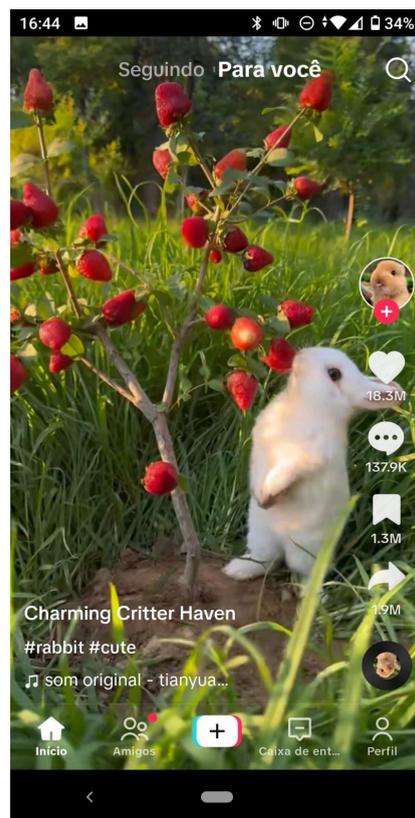


Figura 6: Aba Para Você. Fonte: elaborada pela autora.

Como qualquer outro sistema de recomendação algorítmico, o objetivo da aba *Para Você* é recomendar conteúdos com alta probabilidade de despertar o interesse dos usuários. Segundo um post no site da empresa (que, à princípio, parece surpreendentemente aberta no compartilhamento de detalhes sobre seu algoritmo), as recomendações da seção *Para Você* são baseadas em fatores como: *interações do usuário* (como os vídeos que curte ou compartilha, contas que segue, comentários que publica e conteúdos que cria); *informações do vídeo* (como legendas, sons e hashtags); *configurações de dispositivo e conta* (como

¹⁴¹ No original: “central to the TikTok experience and where most of our users spend their time”.

preferência de idioma, configuração de país e tipo de dispositivo) (TikTok, 2020). O post afirma que o último critério recebe um peso menor, uma vez que os usuários não os expressam ativamente como preferências, mas é amplamente sabido que as principais variáveis para refinar recomendações algorítmicas provém de métricas implícitas¹⁴² (Seaver, 2018). Essa abordagem também se desdobra no modo como o app está sempre extrapolando os interesses declarados dos usuários e fazendo experimentações sobre seu comportamento, conforme veremos noutro tópico.

Um documento interno da empresa intitulado “TikTok Algo 101”, produzido pela equipe de engenharia e vazado por um funcionário para o *The New York Times* em 2021 (Smith, op. cit.), revelou outros detalhes sobre o algoritmo e alguns princípios de design comportamental que regem o app. Nele consta que existem quatro fatores principais que seriam levados em conta para direcionar o comportamento do algoritmo: *a) valor do usuário, b) valor do usuário a longo prazo, c) valor do criador e d) valor da plataforma*¹⁴³. O documento contém uma equação aproximada de como os vídeos seriam classificados pelo sistema de recomendação. A equação levaria em conta uma média, para cada vídeo, entre as *previsões baseadas em aprendizado de máquina e o comportamento real do usuário* em relação a quatro tipos de dados (curtidas, comentários, tempo de exibição e reprodução) e é descrita da seguinte forma:

$$Plike \times Vlike + Pcomment \times Vcomment + Eplaytime \times Vplaytime + Pplay \times Vplay^{144}$$

O sistema então recomenda aos usuários os vídeos mais bem ranqueados com base nesse cálculo. Apesar de que os detalhes fornecidos tanto no post da empresa quanto no documento revelado pelo *New York Times* auxiliam na compreensão da lógica do funcionamento do algoritmo, não há nada de surpreendente para os padrões de recomendação algorítmica personalizada de conteúdo. Das variáveis computadas aos objetivos visados, não há nada tecnicamente no algoritmo que explique o ‘fenômeno TikTok’; o que também evidencia a limitação de análises que tendem a fetichizar demasiadamente os códigos algorítmicos e sua

¹⁴² Isso fica evidente, inclusive, em outro trecho do mesmo post no qual relata-se que um indicador forte de interesse (p.e., se um usuário assiste a um vídeo até o fim) recebe maior peso do que um indicador fraco (p.e., se o criador e espectador compartilham a mesma localização).

¹⁴³ Apesar de citar esses quatro valores, a reportagem não dá maiores detalhes sobre como cada um deles incide objetivamente no algoritmo (por exemplo, que peso cada variável recebe) e em nossas pesquisas não foi possível obter o documento original.

¹⁴⁴ Segundo o documento, a equação é uma versão simplificada da equação real em uso, que é mais complexa mas seguiria a mesma lógica.

transparência (Cf. Monahan, 2018; Mike, Crawford, 2016). Mas, se o *algoritmo em si* não possui nada de inovador, qual seria o ‘segredo’ do TikTok para capturar de forma tão eficiente a atenção dos usuários?

Uma primeira diferença crucial consiste no *modelo de grafo*¹⁴⁵ que o app utiliza, o que vem sendo anunciado como “o fim do modelo do grafo social” (Cf. Newport, 2022). Enquanto em plataformas como Facebook (o pioneiro na popularização do modelo com o lançamento do Open Graph em 2010), Twitter e Instagram, o conteúdo exibido nos feeds está baseado na rede de contatos e páginas que o usuário deliberadamente escolhe seguir, o TikTok rompe com essa lógica e recomenda um espectro de conteúdos, à princípio, mais amplo e diversificado, segundo o perfil que lhe é atribuído. Ainda que seja possível seguir e ser seguido por outros usuários (conteúdo exibido na aba *Seguindo*) e o app também utilize esses dados para refinar suas recomendações, o TikTok enfatiza, material e discursivamente, a centralidade de seu algoritmo de recomendação na experiência do app (Siles, Valerio-Alfaro, Meléndez-Moran, op. cit.). Por isso, a própria empresa se define como uma plataforma de entretenimento e não uma plataforma social (Newport, op. cit.)

A mudança de uma arquitetura de rede baseada em grafo social para uma baseada no que vem sendo chamado de *grafo de interesses* (Newport, op. cit.; Wei, 2020), torna o usuário *ainda mais influenciável* pelas recomendações do algoritmo, uma vez que dá à plataforma um controle ainda maior sobre seu campo atencional e suas arquiteturas de escolha. Hern (2022) observa que, ao separar distração da conexão social, o TikTok pôde competir diretamente pelos usuários sem a necessidade de esperar que eles construíssem uma rede de conexões subjacente ao funcionamento do app. Mesmo que você ainda não siga ninguém, o aplicativo exhibe uma sequência infinita de vídeos desde o primeiro acesso, acionando uma das estratégias de antecipação de comportamentos que a plataforma utiliza. Para o autor, a eficácia da experiência TikTok está naquilo que ela *não exige*. Como comentou comigo um amigo usuário do app, “apesar de disponibilizar ferramentas para comunicação direta entre usuários, como o Instagram, dificilmente você vai parar para conversar com alguém no TikTok porque o objetivo ali não é conexão social, mas entretenimento”.

Ainda que nos pareça exagerado falar do ‘fim das redes sociais’, como algumas análises vêm sugerindo (Cf. Newport, op. cit.), uma vez que esses dois paradigmas tendem a coexistir e se

¹⁴⁵ Matematicamente falando, grafos são estruturas matemáticas compostas por um conjunto de vértices (ou nós) conectados por arestas (as linhas que conectam os nós). No contexto das redes sociais, grafos são utilizados para representar as conexões entre os usuários.

sobrepôr em distintos arranjos nos próximos anos¹⁴⁶, é inegável o processo de ‘tiktokização’ das plataformas digitais, seja em relação à priorização de recomendação algorítmica em detrimento de conteúdos de perfis seguidos ou de conteúdos na forma de vídeos curtos¹⁴⁷.

A estratégia de captura de atenção e modulação comportamental adotada pelo TikTok, que desvincula *personalização algorítmica e conexão social*, também tende a produzir mudanças significativas no modo como os chamados “filtros bolha” (Pariser, 2012) funcionam. No paradigma do grafo social, aquilo que Chun (2018, 2021) chama de *segregação homofílica*, ou seja, a estruturação das arquiteturas digitais a partir da conexão do “igual com o igual”, se baseava principalmente em graus de afinidades entre usuários. Essas conexões se originavam, em grande medida, da digitalização de relações sociais pré-existentes (família, amigos, conhecidos, etc). Mesmo que, progressivamente, as redes passaram a englobar conexões que extrapolavam esses vínculos, acompanhando principalmente a ‘entrada em cena’ dos *influencers digitais*¹⁴⁸, elas seguiam se baseando numa arquitetura de rede similar. No modelo do grafo social, um usuário torna-se mais valioso quanto maior sua rede de contatos; por isso, é incitado a todo momento a expandi-la¹⁴⁹. Já no modelo do grafo de interesses, a ‘bolha’ é estruturada somente em torno do próprio indivíduo e dos conteúdos que capturam sua atenção. Uma *bolha hiperpersonalizada e autocentrada* que se adapta em tempo real a qualquer nuance de seu comportamento. Bhandari e Bimo (2020) descrevem a experiência de uso do TikTok como “a de se envolver repetidamente consigo mesmo: conexão *intrapessoal* em vez de *interpessoal*”¹⁵⁰.

¹⁴⁶ Além disso, como é possível ver na Figura 5, à despeito de certa percepção coletiva dos últimos anos de que seria uma rede ultrapassada, o Facebook, rede paradigmática do modelo do grafo social, segue no topo do ranking mundial de usuários ativos.

¹⁴⁷ O Instagram, por exemplo, implementou a ferramenta de vídeos curtos Reels em 2019 e, recentemente, experimentou exibir mais conteúdos de contas não seguidas no feed, voltando atrás na decisão devido à reação negativa dos usuários. Um memorando interno do Facebook de abril de 2022, citando explicitamente a concorrência ao TikTok, expunha o plano de passar a priorizar conteúdos recomendados ao invés de postagens de contas seguidas (Hern, op. cit.). O YouTube lançou em julho de 2021 o YouTube Shorts, praticamente uma cópia do TikTok. Por sua vez, o Twitter inseriu em 2023 uma aba Para Você, na qual também exibe conteúdo recomendado que não depende dos perfis seguidos.

¹⁴⁸ Essa gradual mudança se refletiu inclusive no modo de designar as conexões, que de ‘amigos’ no Facebook (necessariamente mútuos), passaram a ser ‘seguidores’ no Twitter e Instagram (não necessariamente mútuos).

¹⁴⁹ Um dos princípios que orienta esse modelo é o do fechamento triádico que supõe que se a pessoa a tem uma forte conexão com a pessoa b e com a pessoa c, então provavelmente as pessoas b e c têm alta probabilidade de conexão e afinidade.

¹⁵⁰ No original: “is one of repeatedly engaging with one’s own self: intra rather than interpersonal connection.”

O que essa mudança indica sobre as práticas e técnicas da perfilização é que ‘vizinhanças’ ou comunidades produzidas pelo princípio homofílico (Chun, op. cit.) seguem sendo cruciais para a arquitetura das plataformas e suas cada vez mais elaboradas técnicas de influência comportamental. Contudo, já não dependem de nenhum lastro comunitário constituído *a priori* — por mais frágil que fosse o sentido de comunidade ou de comum do paradigma anterior. Em vez disso, estão cada vez mais centradas no “ego-domínio” (Mbembe, 2018) de um sujeito que se crê mais valorizado e captado em sua singularidade quanto mais expropriado de suas escolhas é, tornando as implicações da “personalização de mundos” (Cesarino, 2022) e do “confisco do comum” (Bruno, 2020) produzidas pela plataformização ainda mais profundas e radicais. Como define Cesarino (op. cit.): “temos assim uma situação paradoxal em que quanto mais dependentes dos algoritmos são os usuários para reorganizarem seus metaenquadramentos, mais eles se experimentam como proativos e livres, e mais influenciáveis se tornam.” (p. 130). A tendência da hiperpersonalização também se reflete na crescente especificidade ou granularidade dos nichos explorados pelas plataformas e seus anunciantes, em geral sob a retórica da autenticidade (Chun, op. cit.), algo especialmente evidente no TikTok¹⁵¹.

À medida que elimina os vestígios da ‘velha comunidade’ que pareciam ainda estruturar parte das arquiteturas e formas de sociabilidade das redes, neste novo paradigma, *a relação íntima e recursiva entre usuário e algoritmo* — construída através de um *treinamento mútuo* do algoritmo pelo usuário (que vai aprendendo cada vez mais rápido suas preferências) e do usuário pelo algoritmo (que vai se identificando cada vez mais com a perfilização que lhe é atribuída e continuamente atualizada) — se torna a relação central que condiciona a experiência de uso das plataformas.

Um dos principais efeitos sobre as subjetividades dessa mudança é tornar o comportamento dos sujeitos ainda mais previsível que no modelo anterior, potencializando a projeção de ‘poderes oraculares’ sobre o algoritmo e fazendo com que suas previsões se tornem ainda mais performativas. Afastado das ‘interferências’ ou até mesmo dos olhares vigilantes que a rede de amigos e conhecidos poderia suscitar¹⁵², esse sujeito também se torna mais propenso a acreditar que ‘o algoritmo o conhece melhor que ele mesmo’, como no caso da usuária que

¹⁵¹ Ver, por exemplo, este post da plataforma sobre subculturas:

https://www.tiktok.com/business/en-US/blog/subcultures-are-the-new-demographics?ab_version=control

¹⁵² Devo essa atenta observação às queridas colegas do MediaLab.UFRJ Helena Strecker e Manuella Caputo numa conversa que tivemos sobre o TikTok durante o transcurso desta pesquisa.

afirmou que o algoritmo do TikTok ‘descobriu’ antes que ela mesma que era bissexual. Segundo o relato, “rapidamente, o TikTok começa a ler sua alma como uma espécie de oráculo digital divino, abrindo camadas do seu ser nunca antes conhecidas por sua própria mente consciente.”¹⁵³ (Joho, 2022).

Contudo, é importante notar que a mudança que estamos descrevendo não foi ‘inventada’ pelo TikTok, como algumas análises vêm sugerindo. Afinal, a personalização algorítmica, bem como seus efeitos sobre as subjetividades e sociabilidades, integram há muito as técnicas de captura e modificação comportamental das plataformas e o direcionamento de anúncios personalizados constitui o núcleo do modelo de negócios típico do capitalismo de vigilância (Zuboff, 2021; Bentes, 2022). Em larga medida, poderíamos dizer que a ‘virada’ para o modelo do grafo de interesses consolidada pelo TikTok pode ser entendida como uma *extrapolação das lógicas do marketing e da publicidade digital para toda a rede*. Ela também acompanha a crescente sofisticação dos algoritmos de IA utilizados pelas plataformas e o acúmulo de cada vez mais dados sob seu domínio. Trata-se, portanto, muito mais da intensificação e aceleração de estratégias e processos que já estavam em curso do que de uma ruptura de fato.

A *aceleração* nos conduz ao segundo fator que consideramos crucial para compreender a eficácia e eficiência do TikTok em manter seus usuários “engajados e enganchados” (Bentes, 2022) e que se conecta diretamente com as questões cronopolíticas que nos são caras: seu *regime de temporalidade*. A noção de *tecnologias de governo do tempo* que viemos desenvolvendo ganha aqui novos contornos, uma vez que propomos que o controle sobre a temporalidade é uma parte crucial das técnicas de influência comportamental e da gestão algorítmica da atenção (Bentes, op. cit) que o TikTok põe em operação e, conseqüentemente, para seu sucesso de popularidade e engajamento.

Um dos grandes diferenciais do TikTok em relação a outras redes, além da aposta no modelo do grafo de interesses, é a *alta velocidade e frequência* em que ele opera, derivando numa *temporalidade aceleracionista* que se desdobra em toda a experiência de uso e funcionamento do app: na velocidade com que o algoritmo prevê conteúdos personalizados com potencial de engajamento a partir de poucos dados; na rapidez com que os usuários relatam terem suas preferências e/ou vulnerabilidades descobertas pelo algoritmo e se

¹⁵³ No original: “Quickly, though, TikTok begins reading your soul like some sort of divine digital oracle, prying open layers of your being never before known to your own conscious mind.”

sentirem ‘fisgados’ pela plataforma; na velocidade com que o algoritmo refina suas predições; no ritmo frenético característico dos vídeos; na quantidade de conteúdo consumido por minuto na plataforma; na testagem contínua das preferências dos usuários; na distribuição rápida e massiva de conteúdo e na potência viral destes; na aceleração do extrativismo de dados que alimenta todo o aparato preditivo do sistema de recomendação. Além da *aceleração*, outros traços fundamentais do regime de temporalidade do TikTok são a *fragmentação* e o controle sobre o *ritmo perceptivo* dos usuários. Em virtude de tais fatores, propomos que o TikTok pode ser entendido como um *ambiente digital de alta frequência* regido por *microtemporalidades aceleradas e fragmentadas* que manifestam infiltrações das lógicas da financeirização.

Visando sistematizar e aprofundar tais questões, selecionamos as que nos pareceram mais pertinentes para nossa pesquisa e as agrupamos em cinco subtópicos que desenvolvemos a seguir: *i) Aceleração da perfilização; ii) Ritmo (hábito + crise); iii) Fragmentação; iv) Aceleração da testagem; v) Aceleração do extrativismo.*

Enfatizamos que essa organização privilegia critérios analíticos, uma vez que, tanto ao nível da experiência do usuário quanto das técnicas e estratégias da plataforma, elas se sobrepõem ou mesmo se confundem.

i) Aceleração da perfilização

Uma das características mais enfatizadas sobre o algoritmo do TikTok é a rapidez com que ele perfila os usuários. Uma investigação do *Wall Street Journal* (2021), na qual repórteres criaram mais de uma centena de bots para desvendar a dinâmica do sistema de recomendação, mostrou que o app é capaz de identificar os interesses de um usuário em menos de duas horas e, em alguns casos, em menos de quarenta minutos.

Esse processo de perfilização acelerada é iniciado já no chamado *onboarding*, o fluxo inicial de uso de um aplicativo, considerado fundamental para atrair o usuário, coletar informações sobre seu perfil e tentar retê-lo. Ao abrir pela primeira vez o app, o usuário pode escolher alguns temas de seu interesse (Figura 7).



Figura 7: Tela de *onboarding* do TikTok. Fonte: elaborada pela autora (setembro de 2023).

Após a seleção de interesses, o TikTok passa a exibir vídeos relacionados aos temas escolhidos¹⁵⁴. A partir daí, o sistema passa a testar, ativa e rapidamente, o que o usuário ‘curte’ e o que ‘não curte’ e a refinar suas previsões em tempo real com base nesses *inputs*. Como os vídeos, em geral, são muito curtos, o processo de aprendizagem do algoritmo é muito veloz, uma vez que o sistema coleta diversas informações sobre o comportamento do usuário num breve período de tempo. Em outras plataformas, esse processo é similar, mas muito mais lento. Como declara Matthew Brennan, autor de um livro sobre a plataforma, “[v]ocê pode assistir algo por cinco minutos no YouTube sem tocar na tela. Se for Netflix, você pode assistir por meia hora sem fazer nada. Assim, o sistema recebe menos feedback. (...) No TikTok, você interage constantemente com o aplicativo e, mesmo que não o faça, também está dando informações.” (Orgaz, 2020). Todo esse mecanismo faz com que o

¹⁵⁴ Caso não sejam selecionados interesses iniciais, a plataforma afirma que exibe uma amostra genérica de vídeos populares no app (TikTok, 2020).

algoritmo do TikTok seja capaz de ‘nichar’ o usuário em pouco tempo de uso, recomendando-lhe cada vez mais conteúdos similares aos que capturaram sua atenção. É claro que aqui, como em todos os meios sob a lógica da economia da atenção, “a valência do conteúdo interessa menos que o simples fato do engajamento” (Cesarino, 2021, p. 109). Não importam os motivos pelos quais algum conteúdo afetou o usuário, apenas que esse efeito foi gerado.

O fato da personalização algorítmica acontecer em tempo real e de modo muito perceptível para o usuário, contribui ainda mais tanto para a projeção de habilidades adivinhatórias sobre o algoritmo quanto para o ‘senso de relevância’ do usuário sobre ele mesmo. Siles (2023) descreve essa interpelação como “o trabalho embutido nos algoritmos para convencer as pessoas de que as plataformas estão falando diretamente com elas, as ‘chamando’ de maneiras específicas”¹⁵⁵ (p. 35). Essa percepção de intimidade e *relevância* — não à toa esse é dos termos mais utilizados pelas plataformas para valorizar as recomendações algorítmicas — é um dos fatores que contribui para invisibilizar a intensa assimetria de saber-poder em jogo na relação *usuários x plataformas*.

É válido frisar mais uma vez que, mais do que a lógica subjacente ao mecanismo em si, o maior diferencial do TikTok para prever e influenciar comportamentos por meio da perfilização algorítmica é a impressionante velocidade com que ele o faz. Em termos cibernéticos, os *feedbacks loops* de seu algoritmo são muito mais curtos do que os de outras plataformas, o que leva tanto o algoritmo a aprender e se adaptar mais rapidamente às preferências do usuário quanto acelera os efeitos próprios da recursividade (como a repetição e intensificação de padrões) ou até mesmo a transformação do *hábito* em *vício* (Chun, 2016). Por isso, alguns jornalistas e pesquisadores vêm alertando como o TikTok também potencializa a exploração das propensões e vulnerabilidades dos usuários e acelera a entrada nos chamados *rabbit holes*¹⁵⁶ (Wall Street Journal, 2021). Em relação a tais críticas, contudo, há que se ter em mente também os fatores geopolíticos envolvidos no alarde sobre ‘os perigos do TikTok’, encampados principalmente pela mídia estadunidense.

¹⁵⁵ No original: “the work embedded in algorithms to convince [people] that [platforms] are speaking directly to them, ‘hailing’ them in particular ways”.

¹⁵⁶ A metáfora provém do livro *Alice no País das Maravilhas* e tornou-se conhecida nos últimos anos para descrever o processo de envolvimento inadvertido, imprevisível e acelerado com conteúdos conspiracionistas na internet e que com frequência leva ao fortalecimento da crença conspiratória.

A plataforma, no entanto, parece bastante atenta às implicações que o excesso de feedbacks de autorreforço podem causar em seu ecossistema informacional e no comportamento de seus usuários. Diante do risco de excesso de homogeneidade e repetição, o TikTok aciona o que poderíamos chamar de *estratégias de modulação do ritmo perceptivo* dos usuários e sobre as quais nos aprofundaremos adiante.

ii) *Ritmo (hábito + crise)*

O controle sobre o ritmo perceptivo dos usuários está intimamente relacionado à formação de *hábitos*, um dos pilares das técnicas de influência e condicionamento comportamental utilizadas pelas plataformas e suas notáveis matrizes behavioristas (Eyal, 2020; Bentes, op. cit.). Cesarino enfatiza como é justamente através do ritmo que as plataformas se apropriam da atenção dos usuários: “com efeito, é o ritmo imposto pelas mídias digitais ao cotidiano dos usuários que propicia os efeitos pretendidos.” (Cesarino, op. cit., p. 109).

Numa proposição bastante original, Chun (2016) propõe que o ritmo temporal das redes é fruto de uma composição entre a temporalidade do *hábito* (ou da repetição) com a temporalidade da *crise permanente*, resultando na atualização infinita. Segundo sua equação: *hábito + crise = atualização*. Para ela, “[r]edes são feitas de tempo: o tempo crônico dos hábitos e o tempo pontual da crise. Desdobrando-se em tempo real, a repetição habitual fundamenta laços; [enquanto] as crises quebram e criam novos”¹⁵⁷ (ibid., p. 16). A essência da temporalidade do hábito é a antecipação baseada em repetições não conscientes, garantindo previsibilidade e automatização comportamental. Mas, como pontua a autora, “a repetição gera expertise, ao mesmo tempo em que gera tédio”¹⁵⁸ (ibid., p. 1). Por isso, visando manter a atualização constante, a temporalidade da repetição habitual é constantemente desfeita e refeita pela temporalidade da crise que garante a reinserção de algum grau de diversidade informacional nos sistemas cibernéticos.

Uma vez que as crises se tornam habituais elas não chegam a provocar nenhum tipo de *ruptura real* nos sistemas algoritmicamente mediados, especialmente nas mídias sociais — em parte, poderíamos acrescentar, em razão do modo como sistemas recursivos ‘absorvem’ eventos contingentes —, fazendo com que tanto as redes quanto os sujeitos estejam, como

¹⁵⁷ No original: “Networks are made out of time: the chronic time of habits (memory) and the punctuating time of crisis. Unfolding in real time, habitual repetition grounds ties; crises break and create new ones.”

¹⁵⁸ No original: “repetition breeds expertise, even as it breeds boredom”.

propõe a autora, “*se atualizando para permanecer (quase) os mesmos*”¹⁵⁹ (ibid., p.1, grifo nosso): “[a]s crises fazem do presente uma série de atualizações nas quais corremos para permanecer próximos do mesmo e nas quais a informação se espalha não como um vírus poderoso e avassalador, mas sim como uma longa e fina corrente morta-viva.”¹⁶⁰ (ibid., p. 3).

A temporalidade ambivalente das redes descrita por Chun é bastante perceptível nos *feeds* de redes sociais como o Facebook e Instagram, mas poucas o manifestam de forma tão nítida quanto o TikTok. A experiência de rolar o feed *Para Você* pode ser descrita como uma *sequência frenética de hábito e crise permanentes*. Por um lado, como vimos, o algoritmo age com incrível velocidade para recomendar conteúdos que se alinham aos interesses do usuário, produzindo um “estado de personalização algorítmica harmoniosa”¹⁶¹ (Siles, Valerio-Alfaro, Meléndez-Moran, 2022, p. 11) que tende à *previsibilidade* e à *homogeneidade*. No entanto, noutra via, também está também constantemente interrompendo esses padrões repetitivos para evitar o desinteresse e o tédio. Um trecho do documento sobre o algoritmo do TikTok sintetiza claramente essa estratégia:

se um usuário gosta de um determinado tipo de vídeo, mas o aplicativo continua enviando o mesmo tipo para ele, ele rapidamente fica entediado e fecha o aplicativo. Nesse caso, o valor total criado pelo usuário assistindo ao mesmo tipo de vídeo é inferior ao de assistir cada vídeo, porque *a repetitividade leva ao tédio*¹⁶². (Smith, op. cit, grifo nosso)

No post em que descrevem o funcionamento do sistema de recomendação, eles sinalizam que um dos maiores desafios dos sistemas de recomendação algorítmicos é que, “ao otimizar para personalização e relevância, existe o risco de apresentar um fluxo de vídeos cada vez mais homogêneo”¹⁶³ (TikTok, 2020). Assim, afirmam, “para manter seu feed *Para Você* interessante e variado, nosso sistema de recomendação funciona para intercalar diversos tipos de conteúdo com aqueles que você já sabe que adora” (ibid.) exibindo, eventualmente, conteúdos que parecem não estar alinhados aos interesses expressos dos usuários. “Nosso

¹⁵⁹ No original: “updating to remain (close to) the same”.

¹⁶⁰ No original: “Crises make the present a series of updates in which we race to stay close to the same and in which information spreads not like a powerful, overwhelming virus, but rather like a long, undead thin chain.”

¹⁶¹ No original: “a state of harmonious algorithmic personalization”.

¹⁶² No original: “if a user likes a certain kind of video, but the app continues to push the same kind to him, he would quickly get bored and close the app. In this case, the total value created by the user watching the same kind of videos is lower than that of watching each single video, because repetitiveness leads to boredom.”

¹⁶³ No original: “By optimizing for personalization and relevance, there is a risk of presenting an increasingly homogenous stream of videos.”

objetivo é encontrar o equilíbrio entre sugerir conteúdo que seja relevante para você e, ao mesmo tempo, ajudar você a encontrar conteúdo e criadores que o incentivem a explorar experiências que você não veria de outra forma”¹⁶⁴ (ibid.). A interrupção de padrões repetitivos também está atrelada às estratégias de testagem constantes que a plataforma utiliza, sobre as quais nos deteremos em outro tópico.

Um estudo baseado no acompanhamento de novos usuários do TikTok também diagnosticou essa alternância de ritmos (Siles, Valerio-Alfaro, Meléndez-Moran, op. cit.). Após a fase inicial de treinamento algorítmico para chegar a uma personalização satisfatória, os autores relatam que as expectativas de estabilidade dos usuários foram, em sua maioria, surpreendidas ou frustradas porque o algoritmo entrou num *estado oscilatório* em que alternava entre recomendações que lhes interessavam e outras não, levando alguns, inclusive, a abandonar o app. “É muito agitado, tem muito de tudo”¹⁶⁵ (ibid., p. 11), relatou uma das participantes. Os pesquisadores descrevem esse ritmo oscilatório como uma “supersaturação, isto é, a sensação de que as pessoas tinham que lidar com recomendações demais e muito diversas num ritmo extremamente rápido”¹⁶⁶ (ibid.). Essa supersaturação é reforçada pela *densidade informacional* dos vídeos do TikTok, algo que a plataforma enaltece como uma vantagem dos efeitos de seus conteúdos sobre a memória dos usuários em relação aos concorrentes seus anunciantes.

Algo que minha experiência de uso pessoal do aplicativo detectou é que quanto maior o desinteresse nos conteúdos recomendados (manifesto sobretudo ao não assistir nenhum vídeo até o final, apenas rolando contínua e rapidamente a tela para o próximo vídeo), mais a plataforma investe no pólo da ‘crise permanente’ exibindo vídeos cada vez mais aleatórios, como se tentando buscar algum conteúdo que finalmente fisgasse minha atenção. Após poucos minutos de uso com essa abordagem (que, mais que deliberada, foi instintiva), notei que minhas recomendações foram se diversificando e ampliando a tal ponto que meu feed se tornou uma *experiência caótica de crise permanente*. Essa abordagem de uso resultou, curiosamente, numa espécie de *anti-personalização algorítmica*, tensionando as virtudes de previsão e precisão tão alardeadas sobre o TikTok.

¹⁶⁴ No original: “Our goal is to find balance between suggesting content that's relevant to you while also helping you find content and creators that encourage you to explore experiences you might not otherwise see.”

¹⁶⁵ No original: “It’s very ‘choppy’, [there is] a lot of everything”.

¹⁶⁶ No original: “‘oversaturation’, that is, the sensation that people had to deal with too many and very mixed recommendations at an exceedingly fast pace.”

A oscilação entre a temporalidade do hábito e da crise também está ligada ao princípio das *recompensas variáveis* que integra o chamado *modelo do gancho* (Eyal, 2020) e que, segundo Bentes (2022) é um dispositivo central das plataformas para formar hábitos e condicionar comportamentos. Composto de quatro etapas (gatilhos, ação, recompensas variáveis, investimento) interligadas de modo recursivo, essa estratégia de design comportamental herdeira do behaviorismo “organiza os estímulos, procurando tornar mais ou menos prováveis certos comportamentos” (Bentes, 2020, p. 152). O princípio das recompensas variáveis que compõe o modelo envolve a introdução de variabilidade de recompensas. Simplificando, significa que você ‘ganha’ algumas vezes e ‘perde’ outras, levando a uma busca constante por novas recompensas. Aplicado à recomendação de vídeos do TikTok e formação de hábitos, isso faz com que, por exemplo, embora o conteúdo que você acabou de assistir não tenha te conquistado, há a expectativa de que o próximo irá fazê-lo. Esse mecanismo cria um ciclo de gratificação intermitente em que os usuários esperam que o ‘próximo vídeo’ seja mais envolvente ou divertido que o anterior, promovendo padrões comportamentais viciantes.

Assim como em relação a outros aspectos, o principal diferencial do TikTok quanto ao controle do ritmo perceptivo dos usuários através de uma modulação do hábito e da crise é a velocidade com que modula essa dupla temporalidade. Mas, além da aceleração, há outro fator crucial para a gestão algorítmica da atenção (Bentes, op. cit.) que ele mobiliza : sua temporalidade extremamente fragmentada.

iii) Fragmentação

Já se tornou lugar comum afirmar que, sob o imperativo da economia da atenção no contexto de um capitalismo baseado em dados, nossa atenção se tornou um recurso tão escasso quanto valioso. Visando solucionar a (insolucionável) equação de excesso informacional *versus* escassez atencional, esse imperativo justificou toda sorte de estratégia posta em prática pelas plataformas digitais nos últimos anos na batalha 24/7 por nossos globos oculares e nosso tempo (Crary, 2016) e está na raiz de muitos dos sintomas do “mal-estar na era da plataformização” (Cesarino, 2022, p. 86). Um dos notáveis ‘custos cognitivos’ dessa demanda impossível foi a percepção de uma redução drástica da capacidade de atenção dos usuários (*attention span*, em inglês) ao longo dos últimos anos, produzindo um sentimento coletivo de que, nessa cada vez mais saturada e acelerada paisagem midiática, nossa atenção

foi sendo progressivamente ‘encurtada’ (McSpadden, 2015; Stokel-Walker, 2022; The Centre for Attention Studies, 2022).

Sob a alegação de que a capacidade de manter a atenção é cada vez menor, a aposta em conteúdos cada vez mais curtos e acelerados nos últimos anos surgiu como uma ‘solução’ das plataformas para um problema que, em grande medida, elas mesmas haviam criado. Assim, a fim de manter os usuários engajados nesse novo regime atencional, as plataformas passaram a extrair e comercializar ‘fatias’ cada vez menores e mais condensadas de nossa atenção e do nosso tempo. *‘Fragmentar e acelerar para reter’* parece ser a fórmula da vez das plataformas baseadas em recomendação algorítmica de conteúdo e o TikTok é a versão mais bem acabada desse *tecno-crono-solucionismo*.

Um estudo conduzido pela Microsoft e que teve ampla repercussão constatou que a capacidade média de atenção das pessoas diminuiu de 12 para 8 segundos entre 2000 e 2015 (McSpadden, 2015. “Agora você tem um período de atenção menor do que de um peixinho dourado”, alarmava a reportagem da Time sobre a pesquisa. Outra pesquisa, interna do TikTok, divulgada em junho de 2021, revelou que metade dos usuários entrevistados achava estressante vídeos com mais de 60 segundos e um terço deles assistia vídeos em velocidade dupla (Stokel-Walker, 2022). Contraintuitivamente, no entanto, a plataforma vem aumentando progressivamente a duração máxima dos vídeos no TikTok: de 15 segundos, inicialmente, esse limite passou para 1 minuto (2017), logo, 3 minutos (2021) e, depois, 10 minutos (2022)¹⁶⁷. Segundo a Wired (op. cit.), o objetivo dessa decisão é aumentar receitas publicitárias, estabelecer uma concorrência mais direta com o YouTube e envelhecer a base de usuários. Mas, para utilizar um jargão do meio, a proposta ‘não engajou’. Dados apontam que a duração média dos vídeos em 2023 foi de 39 segundos e em 2024 de 42,7 segundos¹⁶⁸.

“Não é porque não tenho tempo, mas porque não consigo me concentrar. Não consigo me concentrar”, afirmou um dos entrevistados na supracitada pesquisa interna do TikTok (Wired, op. cit.). Apesar da declaração soar muito alinhada com o *zeitgeist* contemporâneo, é fundamental não desvincular a atenção de toda uma complexa e heterogênea trama de produção histórica de subjetividades — da qual também participam, é claro, as mediações

¹⁶⁷ Após testes com o limite de 10 minutos, o TikTok voltou atrás e voltou a limitar os vídeos em 3 minutos em abril de 2023.

¹⁶⁸ <https://www.statista.com/statistics/1485205/tiktok-video-duration/>

técnicas (Bucher, 2012)¹⁶⁹ — a partir da qual as funções e disfunções da atenção são produzidas. Na versão contemporânea desta trama, as fronteiras entre a falta de atenção, a demanda inalcançável por otimização e a ansiedade (outro problema crônico ‘do nosso tempo’), por exemplo, se tornam nebulosas. Daí as limitações de entender a atenção apenas enquanto um *bem individual, finito, escasso e mensurável* (Cf. Caliman, 2012) passível de comercialização — em grande medida, reflexo da influência da psicologia cognitiva e das neurociências, inclusive, nas leituras críticas sobre o problema — e não como um *constructo relacional e situacional*.¹⁷⁰ (Cf. Bucher, 2012, Bentes, op. cit.).

Nesse sentido, é evidente que a crescente dificuldade dos usuários de se concentrarem por períodos prolongados é muito mais uma justificativa para validar a crescente aposta da indústria em conteúdos curtos do que o real motivo para tanto; o que não significa que este seja um falso problema, é claro. Trata-se, mais uma vez, da estratégia tão recorrente das *big techs* de transformar o *contingente* em *inevitável*. Se ‘fragmentar e acelerar para reter’ é a solução da vez para gerir a atenção dos usuários é porque notou-se que essa técnica potencializa ainda mais os efeitos pretendidos sobre seus comportamentos. Como os indicadores sobre o crescente aumento do tempo médio de uso de apps indicam (Figura 4, p. 108), a redução da duração dos vídeos aumenta o tempo de permanência dos usuários nos aplicativos. Além disso, uma vez que é sobretudo na esfera do cognitivo não-consciente que as plataformas e algoritmos nos interpelam (Hayles, 2017; Cesarino, 2022), entre outras técnicas, através da automatização do hábito, manter-se atento nesse contexto é, em si mesmo, uma espécie de contínua resistência às propensões cognitivas inscritas nestes dispositivos, ainda que se saiba que nos regimes atencionais digitais *atenção* e *foco* não são necessariamente sinônimos¹⁷¹.

A fragmentação temporal empregada pelo TikTok também possibilita a aceleração do circuito recursivo das ‘pequenas doses de recompensas’ utilizado pelo *modelo do gancho*, fazendo

¹⁶⁹ Bucher (2012) propõe a noção de *tecnicidade da atenção* como “uma forma de compreender como os processos e práticas de atenção estão fundamentados num meio sociotécnico” (p. 11, tradução nossa). Para a autora, numa cultura digital, a capacidade atencional precisa ser entendida como um *constructo relacional entre usuários e meios técnicos*.

¹⁷⁰ Especialistas entrevistados para uma reportagem da BBC que questiona os resultados da supracitada pesquisa da Microsoft destacam que a ideia de “período de atenção médio e mensurável” é, em si, problemática, uma vez que a atenção depende muito do caráter da tarefa e da contribuição do indivíduo nesta tarefa (Maybin, 2017).

¹⁷¹ Bentes (op. cit.) pondera que, embora a atenção seja comumente entendida enquanto um processo seletivo relacionado à capacidade de focar em algo, nos regimes atencionais online a atenção não se apresenta de modo uniforme e é preciso considerar graus variados de foco e concentração. Além disso, como ponderado na nota anterior, o grau e a qualidade da atenção variam, entre outros fatores, conforme o tipo de tarefa executada.

com que a experiência de uso do app se transforme numa espécie de *hiperfluxo de microdoses de dopamina*. Trata-se de um dispositivo temporal crucial para produzir o ritmo acelerado característico do app, bem como potencializar os efeitos pretendidos sobre os comportamentos dos usuários.

Outra instância em que a fragmentação é utilizada no app é no próprio design do *feed*. Diferente de plataformas como o Facebook, Twitter e Instagram, em que o feed funciona no modo *scroll* infinito, dando ao usuário uma visão panorâmica e relativamente integrada dos conteúdos postados em sua rede, no TikTok os vídeos são exibidos um a um, propiciando, ao nível perceptivo, uma maior sensação de imersão. Isso também contribui para a perda do controle sobre o tempo de uso do app, uma queixa muito comum entre os usuários. Esse formato também possibilita um rastreamento minucioso do comportamento do usuário para cada vídeo separadamente. Ou seja, não é só ao nível da ‘condução das condutas’ dos usuários que a fragmentação potencializa os efeitos visados pelas plataformas. Como veremos em mais detalhes, ela também está diretamente relacionada à aceleração do extrativismo que sustenta todo o aparato algorítmico preditivo. Já ao nível dos criadores de conteúdo, os ‘*creators*’, a fragmentação propicia a aceleração da produção num contexto de demanda por atualização constante. Afinal, conteúdos mais curtos demandam menos investimento de tempo e dinheiro e ‘performam’ melhor no algoritmo (se uma das variáveis para rankear bem um conteúdo é que ele seja visto até o final, é mais provável que um conteúdo curto atinja essa meta).

iv) Aceleração da testagem

Bruno, Bentes e Faltay (2019) enfatizam como uma das mudanças radicais advindas com a plataformização é a transformação do espaço digital num imenso laboratório que utiliza imensos volumes de dados para testagem, previsão e indução de comportamentos. Convertidos em “laboratórios de plataforma” (ibid.), esses ambientes digitais se transformam nas mais bem sucedidas realizações da proposição behaviorista de Skinner de transformar o próprio mundo num laboratório ao invés de trazê-lo para dentro de seus muros (Zuboff, 2021).

A lógica da testagem contínua sobre os comportamentos dos usuários descrita pelos autores é claramente perceptível no TikTok, principalmente no modo como está a todo momento testando, ativa e antecipadamente, suas previsões algorítmicas. Assim como a estratégia da

fragmentação, ela se desdobra em diversos níveis. Num primeiro, na própria adoção do modelo do grafo de interesses que, ao dispensar a necessidade de que o usuário construa uma rede de conexões subjacente ao funcionamento do app, oferece antecipadamente uma sequência de recomendações personalizadas de vídeos desde o primeiro acesso. Outra dimensão em que a testagem contínua opera é na recomendação de conteúdos que extrapolam as preferências do usuário (que interpretamos enquanto uma modulação da temporalidade do hábito e da crise) e que possui relação direta com a detecção antecipada da “potência viral” (Araújo, Karhawi, 2023) dos vídeos. Como a própria plataforma declara, “ao oferecer vídeos diferentes de tempos em tempos, o sistema também é capaz de ter uma noção melhor do que é popular entre uma gama mais ampla de públicos para ajudar a fornecer a outros usuários do TikTok uma ótima experiência também.”¹⁷² (TikTok, 2020).

Essa experimentação contínua em distintas camadas não apenas permite ao sistema reconhecer rapidamente os interesses de um usuário individual, acelerando o processo de personalização, mas é uma parte fundamental do *modelo de distribuição de conteúdo* que o TikTok utiliza, outro fator que o diferencia de seus concorrentes. Independente do tamanho de uma conta, cada vídeo postado no TikTok é veiculado para, pelo menos, um lote de pessoas na página *Para Você*. O número de usuários nesses lotes aumenta exponencialmente caso os vídeos obtenham bom desempenho (Hern, 2022). Segundo a empresa,

[e]mbora seja provável que um vídeo receba mais visualizações se for postado por uma conta que tenha mais seguidores (...) nem a contagem de seguidores nem se a conta teve vídeos anteriores de alto desempenho são fatores diretos para o sistema de recomendação.¹⁷³ (TikTok, 2020)

Isso explica em grande medida porque os vídeos do TikTok costumam viralizar tão rápido, tornando suas métricas, em termos absolutos, muito vantajosas em relação a seus concorrentes para criadores e anunciantes. Entre os criadores, esse modelo de distribuição escalável ajuda a alimentar a crença de que “*todo mundo pode ser famoso com o algoritmo do TikTok*” (Araújo, Karhawi, op. cit., p. 6, grifo nosso). À diferença de outras plataformas em que é necessário primeiro ‘fazer um perfil crescer’ para ter alcance entre os usuários, no TikTok, mesmo que você não tenha nenhum seguidor, o vídeo acabará atingindo uma audiência inicial e, caso ‘performe’ bem, pode alcançar milhares ou milhões de usuários com

¹⁷² No original: “By offering different videos from time to time, the system is also able to get a better sense of what’s popular among a wider range of audiences to help provide other TikTok users a great experience, too.”

¹⁷³ No original: “While a video is likely to receive more views if posted by an account that has more followers, by virtue of that account having built up a larger follower base, neither follower count nor whether the account has had previous high-performing videos are direct factors in the recommendation system.”

extrema rapidez. Esse aspecto também ajuda a explicar por que os usuários da plataforma aderem tanto e tão rápido aos desafios (*challenges*)¹⁷⁴ e/ou tendências de live que viralizam no app, como o fenômeno das *lives NPC*, que viralizou em 2023¹⁷⁵.

A estratégia da testagem contínua não condiciona somente o modo como a plataforma interpela seus usuários, mas as próprias estratégias destes para ganhar visibilidade na plataforma. Abidin (2021), ao investigar o trabalho de visibilidade dos criadores de conteúdo do TikTok, aponta que a fama e a visibilidade no app não estão centradas em perfis, mas em conteúdos, fazendo com que os usuários estejam sempre buscando adaptar-se rapidamente às novas ‘*trends*’ que viralizam na plataforma:

Os usuários do TikTok [...] não mantêm uma única personalidade ou único estilo coerente e singular, mas, ao invés disso, são ativos e muito rápidos, adaptando-se às últimas tendências e práticas virais do TikTok e a uma grande variedade de estilos. [...] As lógicas do TikTok incentivam os aspirantes a celebridades da internet a buscar, aprender, participar e se envolver ativamente com o que está ‘se tornando viral’ no momento, a fim de permanecer visível para outras pessoas no próprio aplicativo (Abidin, 2021, p. 8)

Esse traço aponta novamente para a temporalidade da crise permanente como condição para a atualização ininterrupta das redes (Chun, 2016) e como padrão que une a camada técnica e humana dos atuais ecossistemas digitais. Assim como em relação aos outros aspectos que viemos abordando ao longo deste estudo de caso, a principal diferença do TikTok em relação a outras plataformas do modo como o ideal da otimização ininterrupta incide sobre a produção de subjetividades, especialmente dos criadores de conteúdo, é a velocidade com que o faz. É notável como as tendências emergem e desaparecem rapidamente no app, exigindo que tanto criadores quanto empresas se adaptem rapidamente para capitalizar esses momentos efêmeros de atenção. Assim, tal como os algoritmos que continuamente otimizam seus modelos e recomendações, trata-se também de um modelo de subjetividade condenado à ‘*loopings* de autoaprimoramento’ e adaptações infinitas e cujas intersecções com modos de subjetivação produzidos no neoliberalismo são patentes. Como aponta Mbembe, ao refletir sobre os efeitos da convergência entre neoliberalismo e digitalização sobre as subjetividades, trata-se de um sujeito “moldável e convocado a se reconfigurar permanentemente em função

¹⁷⁴ Segundo o próprio site do TikTok, “os desafios online geralmente envolvem pessoas que se gravam a si próprias a fazer alguma coisa difícil e partilham online para encorajar outras pessoas a fazê-lo também.” Disponível em: <https://www.tiktok.com/safety/pt-pt/online-challenges/>. Acesso em: 30 de setembro de 2023.

¹⁷⁵ As *lives NPC* (*Non-Player Character*, em inglês) são transmissões ao vivo onde os criadores de conteúdo adotam comportamentos e falas repetitivas e robóticas, imitando o comportamento de NPCs em videogames, personagens que normalmente seguem padrões de fala e ação pré-programados. Nas *lives NPC*, os criadores respondem a ações ou mensagens dos usuários de maneira mecânica e previsível.

dos artefatos que a época oferece”, que “deve abraçar sua condição de sujeito solúvel e fungível, a fim de atender à injunção que lhe é constantemente feita — tornar-se um outro” (Mbembe, 2018, p. 17).

v) *Aceleração do extrativismo*

Um dos princípios básicos que guia o capitalismo de vigilância e seus produtos de predição baseados em aprendizado de máquina (Zuboff, 2021) é que, quanto mais dados forem extraídos, maior o *poder preditivo* desse modelo de conhecimento e intervenção sobre os comportamentos. Desse modo, uma das principais vantagens para as plataformas digitais da crescente aposta em vídeos curtos é a *aceleração do extrativismo de dados* que esse formato de conteúdo possibilita. Como vimos ao longo deste tópico, o uso de vídeos fragmentados pelo TikTok faz com que o volume de dados de treinamento que um usuário fornece por unidade de tempo de uso do app seja muito mais alto do que em plataformas com conteúdos mais longos, acelerando toda a cadeia recursiva da personalização algorítmica e seus efeitos (pense-se, por exemplo, em quantos vídeos podem ser assistidos, em média, em uma hora na Netflix, no YouTube e no TikTok).

Portanto, mais do que um modelo que se adequa bem à alegada escassez atencional dos usuários — ainda que, sem dúvidas, ele também responde a essa demanda —, um *feed* de vídeos curtos é a base de treinamento perfeita para extrair muitos dados e otimizar um algoritmo de aprendizado de máquina num curto período de tempo. Além disso, a própria aposta em conteúdos em vídeo e não em conteúdos estáticos, como fotos ou textos, também potencializa esse extrativismo, uma vez que permite a inferência de *métricas implícitas* específicas relacionadas ao comportamento do usuário ao longo do tempo, como veremos a seguir.

Analisando sistemas de recomendação algorítmicos ao longo dos últimos anos, principalmente da Netflix, Seaver (2018), identificou uma mudança no tipo de métricas utilizadas pelas plataformas. Ao invés de *métricas explícitas*, baseadas principalmente em avaliações ativas que os usuários faziam sobre os conteúdos, os sistemas de recomendação passaram a utilizar cada vez mais *métricas implícitas* ou, como chama, *captológicas*, que mensuram a capacidade de um sistema reter a atenção dos usuários. A partir dessa “virada captológica” (Seaver, op. cit.), “a satisfação passa a ser mensurada pela capacidade do sistema de capturar e reter a atenção dos usuários ou o engajamento” (Bentes, 2022, p. 170).

As métricas captológicas estão centradas em interações involuntárias ou inconscientes (que, refletindo entendimentos behavioristas sobre o comportamento, são consideradas mais confiáveis e/ou verdadeiras) e são, com muita frequência, estruturadas no tempo: quanto tempo um usuário permanece em determinado conteúdo ou realizando determinada ação, em que momento do conteúdo ele ‘pula’ para o conteúdo seguinte, se o assiste mais de uma vez, qual o tempo de permanência no serviço, qual a média mensal de uso, e assim por diante.

Seaver observa como o paradigma captológico também se refletiu no design de interfaces, projetadas para incitar cada vez mais interações e extrair mais dados sobre a atividade do usuário. Nessa virada, o protótipo de um sistema de recomendação deixou de ser um *suporte para encontrar informações* para se tornar cada vez mais uma *armadilha para capturar usuários* (Seaver, op. cit). Em termos de design da experiência do usuário (ou simplesmente *UX design*, como é chamado no meio especializado), a ênfase no uso de métricas implícitas e de técnicas cada vez refinadas para capturar a atenção dos usuários e condicionar seus comportamentos também é correlata da crescente relevância que a abordagem do *design sem fricção* (*frictionless design*, em inglês) ganhou na indústria *tech* nos últimos anos (Cf. Cupples, 2021). Guiado pelo mantra “não faça o usuário pensar” (Krug, 2014), essa abordagem de design propõe que a experiência ideal de uso de um produto digital deve requerer o mínimo de esforço cognitivo. Uma *experiência sem fricção* visa tornar a experiência do usuário o mais suave e eficiente possível, removendo obstáculos que possam causar atrito, estranhamento ou frustração. Na confluência do capitalismo de vigilância com a economia da atenção, o objetivo é incitar cada vez mais engajamento e interação, produzir mais dados comportamentais para refinar as previsões algorítmicas, porém, com o mínimo de fricção possível. Um extrativismo (de dados e cognitivo) fluido, invisível e sutil, vendido como *user friendly*, mas que, na prática, é *machinic friendly*.

À vista de tais questões, notamos como o TikTok materializa de modo muito privilegiado a *maximização do extrativismo* com a *minimização da fricção*. Do modelo do grafo de interesses, passando pelos recursos para criação de conteúdo ofertados, a hiperpersonalização, o formato de vídeos curtos exibidos individualmente até as experimentações constantes e escaláveis, toda a plataforma é desenhada para potencializar o extrativismo de dados e treinar rapidamente seu algoritmo ao mesmo tempo em que proporciona uma experiência de uso ágil, fluida, divertida e que aparenta priorizar o usuário.

Visando essa maximização do extrativismo são também empregados algoritmos específicos de Processamento de Linguagem Natural e de computação visual para identificar e classificar os conteúdos presentes em cada vídeo e sofisticar os metadados do sistema — o que também explica a incitação da produção de vídeos ‘supersaturados’ de efeitos visuais, sons e hashtags pela plataforma. Segundo Greenwald (2021), em 1 minuto, as ‘microobservações’ do TikTok são capazes de rastrear de 5 a 6 métricas sobre o usuário. Isso significa que, em 1h de uso, ele é capaz de extrair cerca de 300 a 360 pontos de dados. Para Wei (2020), “a verdadeira magia” do app é “como cada elemento do design e dos processos do TikTok se conectam entre si para criar um conjunto de dados com o qual o algoritmo é treinado para obter desempenho máximo”¹⁷⁶. O resultado é a construção de uma *malha extrativa extremamente granular e veloz* capaz de rastrear minuciosa e exponencialmente tanto os conteúdos postados quanto os comportamentos dos usuários, fazendo com que todo o aparato preditivo que sustenta o sistema de recomendação opere em alta frequência.

Através de um aceleracionismo que se desdobra em múltiplas camadas e que viemos explorando nos últimos tópicos, o governo algorítmico do tempo utilizado pelo TikTok mostra-se extremamente eficaz tanto na gestão da atenção e do comportamento dos usuários, quanto na potencialização do extrativismo que sustenta toda sua arquitetura algorítmica. Da perfilização e personalização velozes, à modulação do ritmo perceptivo por meio de contrastes entre hábitos e crises, o uso de fragmentações temporais e visuais, a aceleração das experimentações e o modelo de distribuição escalável, até a maximização do extrativismo com a minimização da fricção, todas essas técnicas e estratégias transformam o TikTok num *ambiente digital de alta frequência*. Identificamos nesse cronotopo ressonâncias da temporalidade e da racionalidade tecnofinanceira sobre a qual nos debruçamos no tópico 3.2. Trata-se de um ambiente regido por microtemporalidades aceleradas e fragmentadas projetado para viabilizar uma *grande quantidade de operações extrativas e preditivas em alta velocidade e num curto período de tempo*, acelerando e potencializando todo o circuito (comportamental, epistêmico e econômico) que alimenta o funcionamento da plataforma.

¹⁷⁶ No original: “the actual magic is how every element of TikTok's design and processes connect with each other to create a dataset with which the algorithm trains itself into peak performance”.

3.2.2 ChatGPT como megamáquina probabilística

*Não devemos nem antropomorfizar nem objetivar os aparelhos.
Devemos captá-los em sua concretude cretina de um
funcionamento programado, absurdo. (...). Em suma: o que
devemos aprender é assumir o absurdo, se quisermos
emancipar-nos do funcionamento.*
Vilém Flusser

Ao longo desta pesquisa, enfatizamos como o governo algorítmico do tempo que os algoritmos preditivos baseados em aprendizagem de máquina operacionalizam tem como um de seus principais traços a primazia do provável como modo de modular e controlar o campo do possível. Conforme vimos, essa nova normatividade estatística (Rouvroy, Berns, 2015) que subjaz ao saber-poder algorítmico opera através de uma *redução das possibilidades às probabilidades*. Neste tópico retornaremos e aprofundaremos essa dimensão, simultaneamente materializando-a e tensionando-a a partir de um objeto específico: o *ChatGPT*.

O ChatGPT é uma IA conversacional baseada na série de modelos de linguagem GPT¹⁷⁷ (cuja sigla, em inglês, significa *Generative Pre-trained Transformer*), um grande modelo de linguagem natural que utiliza redes neurais profundas para geração de textos considerados similares aos humanos. Lançada com acesso livre em novembro de 2022 pela empresa e instituto de pesquisa californiano OpenAI, a aplicação atingiu impressionantes 100 milhões de usuários ativos em apenas dois meses. O ChatGPT seduziu rapidamente os usuários por sua capacidade de criar textos sobre os mais variados assuntos em diversos formatos e estilos, manter um fluxo conversacional, traduzir em múltiplos idiomas, produzir e revisar códigos de programação, resumir textos, entre outras tarefas até então pouco plausíveis de serem executadas por um mesmo modelo algorítmico. Suas capacidades cognitivas vêm sendo exaltadas pela performance similar ou superior à humana em testes acadêmicos e profissionais nas mais variadas especialidades (OpenAI, 2023) e até no diagnóstico de

¹⁷⁷ Este estudo de caso, desenvolvido em sua maior parte durante o segundo semestre de 2023, baseou-se na versão gratuita do ChatGPT que utilizava o GPT-3.5, versão modificada e menor do modelo GPT-3, lançado pela OpenAI em 2020. Em maio de 2024, a OpenAI atualizou o chatbot disponibilizando também a opção de uso do GPT-4o, mas mantendo a possibilidade de uso do modelo anterior. A versão paga da aplicação, o ChatGPT Plus utiliza atualmente o modelo GPT-4o como base e dá acesso a diversas configurações inexistentes na versão gratuita. Além disso, a OpenAI também disponibiliza uma API com livre acesso a partir da qual é possível customizar o ChatGPT. Vários destes modelos customizados estão disponíveis na *GPT Store*: <https://openai.com/index/introducing-the-gpt-store/>

doenças raras¹⁷⁸ alimentando acalorados debates sobre se estamos (finalmente?!) diante de uma máquina de fato inteligente.

Apesar de que a IA já integrava intensa e extensivamente as maquinarias dos ecossistemas digitais há anos, como a própria trajetória dessa pesquisa vinha acompanhando desde 2019, o lançamento do ChatGPT parece ter se transformado num marco de *entrada oficial na 'era da inteligência artificial'*. Ainda que precursores notáveis, como o pioneiro ELIZA ainda nos anos 60¹⁷⁹ e o Watson da IBM já haviam causado furor, foi a primeira vez que pessoas comuns puderam interagir com uma ferramenta deste tipo de forma massiva, tendo seus pedidos instantaneamente respondidos de forma prestativa e ágil. As reações iniciais de espanto parecem ter se transformado rapidamente em relações de curiosa intimidade e confiança, como o tuíte a seguir exemplifica (Figura 8). Apesar de a tela principal do chat exibir constantemente a mensagem “O ChatGPT pode cometer erros. Considere verificar informações importantes” (agosto de 2024), é evidente como a maioria das pessoas o vê como uma espécie de “*dr. sabe tudo*” e como um ambiente de consulta mais que de produção de textos. O ChatGPT parece materializar assim a versão mais atualizada da dimensão oracular dos modelos algorítmicos sobre a qual comentamos em diversas passagens. Esses imaginários e afetos foram potencializado com a inclusão de funcionalidades como o uso de voz e imagens nas conversações e a possibilidade de personalizar as respostas do chat¹⁸⁰.

¹⁷⁸ “ChatGPT correctly diagnosed a 4-year-old's mysterious disease after 17 doctors failed”.
<https://www.businessinsider.com/chatgpt-diagnose-child-disease-tethered-cord-syndrome-doctors-2023-9>

¹⁷⁹ Primeiro chatbot da história da computação, a ELIZA, desenvolvida no MIT por Joseph Weizenbaum, foi tão bem-sucedida na obtenção de respostas emocionais dos usuários que terminou por desencadear o que se chama hoje de “efeito ELIZA”, a tendência em projetar características humanas (como compreensão de sentido e intencionalidade) em programas computacionais.

¹⁸⁰ Enquanto as primeiras ainda estão restritas à versão paga do chat, a versão gratuita já permite a personalização de respostas. A personalização consta de questões como “O que gostaria que o ChatGPT soubesse sobre si para lhe disponibilizar melhores respostas?” e “De que forma gostaria que o ChatGPT respondesse?”.



Figura 8: Tuíte em que usuária relata o nível de intimidade e confiança que desenvolveu com o ChatGPT. Fonte: X/Twitter.

Relativamente previsível para quem vinha acompanhando as inovações do campo nos últimos anos, mas surpreendente para a grande maioria do público leigo, o lançamento do ChatGPT se transformou num “evento de percepção” (Corrêa, 2023) que provocou um abalo sísmico nos imaginários e nas expectativas mercadológicas sobre a IA, despertando também os antigos “fantasmas de superação, destruição e salvação do humano” (Bruno, Pereira, Faltay, 2023, p. 237) que sempre assombraram essa história. O chamado ‘hype da IA’ que invadiu as mídias e o debate público desde então reativou com força o “cronotopo das ameaças e promessas finais” (Haraway, 2018, p. 241) no qual a tecnociência avançou nas últimas décadas. Por um lado, nutrindo as fantasias tecno-heróicas de superação do humano, os modelos mais avançados do GPT foram saudados como as primeiras “centelhas de Inteligência Artificial Geral” (Microsoft Research, 2023) numa espécie de anúncio de que a tão almejada ‘virada singularista’ estaria muito próxima (Cf. Vinge 1993; Kurzweil, 2005). Por outro, apenas aparentemente oposto, ganharam força os temores sobre os ‘risco existencial da IA’ e a possibilidade de extinção da humanidade por máquinas superpoderosas, alavancados pela capitalização cínica dos próprios financiadores do setor (Bruno, Pereira, Faltay, op. cit.). As mutações desencadeadas a partir de então também pressionaram tanto a proposição de marcos regulatórios para a IA em todo o mundo quanto a corrida para o

lançamento de concorrentes, como o *Bard* da Google em julho 2023, posteriormente rebatizado como *Gemini* em fevereiro de 2024¹⁸¹.

A chegada arrebatadora desse complexo e poderoso cognoscente técnico (Hayles, 2017) provocou uma *perturbação*¹⁸² (Haraway, 2016) em nossa relação com a tecnologia e no próprio *status quo* social, seja pela ameaça ao excepcionalismo humano que materializa ou pelas crescentes incertezas sobre o impacto de ferramentas de automação cognitiva em campos tão diversos quanto a educação, a pesquisa científica, a criação artística, o trabalho, o jornalismo, a saúde, o direito e em nossos próprios modos de ser. Daí tanto a reverberação das narrativas sobre *riscos existenciais*, quanto o recrudescimento da pergunta por ‘aquilo que *realmente* nos faz humanos’ e no que consiste ‘a *verdadeira* inteligência’ e possa nos (re)assegurar uma vantagem epistêmica e ontológica em relação a esse novo agente. À medida que a performance cognitiva desses sistemas se assemelha cada vez mais às dos humanos, mais propensas também se tornam à antropomorfização (Shanahan, 2024).

Longe de pretender dar conta de todas essas dimensões interessa-nos neste tópico, alinhando-nos às questões que atravessam e orientam essa pesquisa, refletir principalmente de que modo o ChatGPT, enquanto uma *megamáquina probabilística*, simultaneamente materializa e complexifica nossa hipótese sobre a redução das possibilidades às probabilidades. Além disso, por meio de uma minuciosa investigação sobre como foi construído e raciocina, refletimos sobre os sentidos e implicações de seu ‘modelo de mundo outro-que-humano’.

Uma das transformações mais importantes da pesquisa e desenvolvimento de IA nos últimos anos foi o aumento vertiginoso do tamanho dos modelos algorítmicos, resultado nos chamados *foundation models*, grandes modelos que podem ser utilizados como base para diversas aplicações (Cf. Bommasani et al, 2022). A abordagem “*one-model-for-everything*” (Varon, Costanza-Chock, Gebru, 2024) que esses modelos materializam impulsionou uma competição acirrada das Big Techs pela liderança nesse campo. Talvez a descrição mais precisa para a inflexão em curso seja *a era dos grandes modelos de IA*, cuja escala e velocidade de desenvolvimento produzem impactos sem precedentes. Um dos principais tipos

¹⁸¹ <https://gemini.google.com/app>

¹⁸² Utilizo a expressão num sentido similar ao uso que Haraway faz do termo em inglês “trouble” no livro *Staying with the Problem* (2016) e recentemente traduzido para o português. Nesse contexto, “to trouble” significa “provocar um incômodo”, “agitar”, “tornar nebuloso”.

de *foundation models* são os *grandes modelos de linguagem* (LLMs, na sigla em inglês) que realizam tarefas de Processamento de Linguagem Natural (NLP, na sigla em inglês), área focada em como as máquinas entendem e produzem linguagem similar à humana. Somente desde 2018, foram lançadas a série BERT, T5 e LaMDA (da Google), a série GPT (da OpenAI), o Turing NLG (da Microsoft) e o LLaMA (da Meta). O tamanho e sofisticação de um modelo de linguagem é medido, basicamente, pelo tamanho do banco de dados de treinamento e pelo número de parâmetros empregados¹⁸³. A tabela (Tabela 1) e a imagem a seguir (Figura 9) dão uma dimensão desse recente salto no tamanho dos LLMs¹⁸⁴ e na velocidade com que esse campo vem evoluindo.

<i>Modelo</i>	<i>Empresa</i>	<i>Lançamento</i>	<i># de parâmetros</i>
BERT large	Google	2018	340 milhões
GPT-1	OpenAI	2018	117 milhões
GPT-2	OpenAI	2019	1.5 bilhões
Turing NLG	Microsoft	2020	17 bilhões
T5	Google	2020	11 bilhões
GPT-3	OpenAI	2020	175 bilhões
LaMDA	Google	2021	137 bilhões
LLaMA	Meta	2023	65 bilhões
GPT-4	OpenAI	2023	1.76 trilhões
Gemini	Google	2023	7 bilhões

Tabela 1: Grandes modelos de linguagem lançados desde 2018 com seus respectivos tamanhos. Fontes: Bender et al (2021); Wikipedia; sites das empresas.

¹⁸³ Parâmetros são valores numéricos que representam aspectos que o modelo é capaz de reconhecer e armazenar para realizar previsões. Eles são aprendidos durante o treinamento e determinam como ele fará previsões ou tomará decisões. Os parâmetros podem assumir diferentes formas dependendo do tipo de modelo treinado. No caso de uma rede neural, os parâmetros se referem aos pesos dos neurônios e definem a capacidade da rede de aprender e capturar padrões e correlações nos dados de treinamento.

¹⁸⁴ Apesar das cifras impressionantes, é importante ter em conta que, como não existem práticas padronizadas de documentação e transparência sobre como os modelos foram construídos e operam, esses números não podem ser desvinculados das estratégias de marketing que guiam o setor.

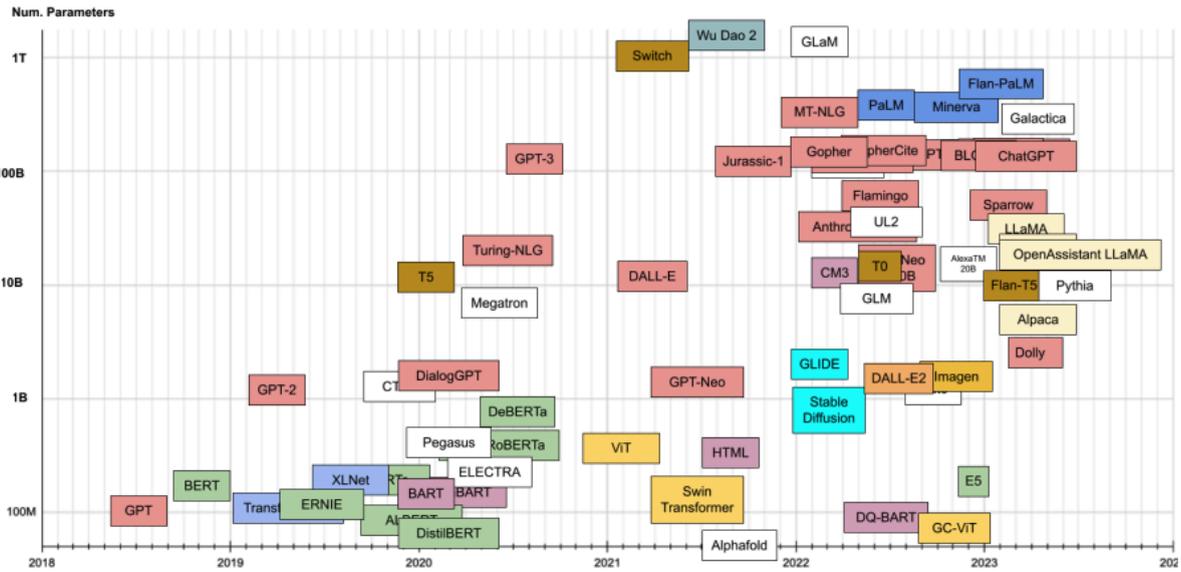


Figura 9: Timeline com os LLM's lançados desde 2018 e seus respectivos tamanhos. As cores agrupam as empresas produtoras dos modelos. Fonte: Amatriain et al (2004)¹⁸⁵.

Grandes modelos de linguagem são modelos matemáticos treinados em tarefas de previsão de texto, ou seja, a previsão da probabilidade de uma sequência de caracteres ou palavras (chamadas tecnicamente de *tokens*) com base em seu contexto anterior ou seu contexto circundante (Bender et al, 2021) a partir de um *prompt*, a entrada ou instrução fornecida ao modelo. LLMs como os da série GPT pertencem à vertente da aprendizagem de máquina que gera novos conteúdos baseados em previsões estatísticas, a chamada *IA Generativa ou Gerativa*, e onde os avanços mais promissores da área vêm acontecendo. Também pertencem a essa vertente os modelos de IA que geram imagens, como o Midjourney e o DALL-E (este último, também baseado no GPT).

Além do tamanho sem precedentes, outra característica crucial da geração mais recente de LLMs é a arquitetura de rede que utilizam. Enquanto os primeiros modelos de linguagem usavam uma arquitetura capaz de prever dependências relativamente limitadas entre palavras (geralmente 5 ou menos), os LLMs utilizam a arquitetura de aprendizado profundo chamada *Transformer*, introduzida pela equipa da Google Brain em 2017 (Vaswani et al, 2017). Baseada nos chamados mecanismos de atenção¹⁸⁶, essa arquitetura é capaz de capturar

¹⁸⁵ <https://arxiv.org/pdf/2302.07730>.

¹⁸⁶ A atenção nesse contexto refere-se à capacidade de um modelo focar em partes específicas de uma entrada durante o processamento. Em vez de tratar toda a entrada de uma vez, os mecanismos de atenção permitem que o modelo "preste atenção" a diferentes partes da entrada com pesos variáveis. Isso é particularmente útil para lidar com sequências, onde a importância de diferentes elementos pode variar.

relações de dependência entre tokens de longo alcance escalando de sobremaneira o *espaço semântico multi-dimensional* a partir do qual os modelos geram textos. O uso da arquitetura Transformer viabilizou a geração de textos mais fluentes, diversificados e longos, além de melhorias em tarefas como tradução e compreensão de contextos e um aumento da escalabilidade das aplicações (Bender et al, op. cit.). Além disso, essa arquitetura de rede diminuiu significativamente o tempo de treinamento em comparação aos modelos anteriores.

Segundo a própria página do produto¹⁸⁷, o ChatGPT é um modelo irmão do InstructGPT e foi treinado através da técnica *Aprendizado por Reforço com Feedback Humano* (RLHF, na sigla em inglês) conceitualmente inspirada na perspectiva de aprendizagem do behaviorismo radical (Bentes, 2022). Essa técnica de treinamento utiliza um *modelo de recompensa* através de feedback humano para classificar as respostas geradas pelo modelo e avaliar se uma saída é boa (alta recompensa) ou ruim (baixa recompensa) permitindo assim o ajuste fino em tarefas específicas (*fine-tuning*, em inglês), tais como entendimento, sumarização de textos, análise de dados, tradução, etc. O modelo de recompensa do ChatGPT foi produzido a partir de amostras aleatórias de conversas que os treinadores tiveram com o chatbot, classificadas das melhores às piores e utilizadas como novas entradas de treinamento. Este é um dos principais motivos da ferramenta ser tão habilidosa em conversações.

Mas antes dessa etapa de aprendizado por reforço com feedback humano, o GPT foi pré-treinado numa gigantesca base de dados não rotulados¹⁸⁸, um dos fatores cruciais para a performance impressionante (ou ao menos a razoabilidade média) das respostas do chat. Nenhuma documentação é fornecida na página, no entanto, sobre esse conjunto de treinamento. Questionando o próprio chat sobre como e com que dados ele foi treinado, ele respondeu: “*Minha base de dados de treinamento consiste em diversos tipos de texto coletados da internet, como artigos, livros, sites, fóruns e muitos outros recursos online disponíveis publicamente. Esses dados incluem informações em várias áreas do conhecimento, como ciência, história, literatura, entretenimento, tecnologia e muito mais.*”. Ele também detalha que tudo o que responde é baseado “*no conhecimento adquirido durante o treinamento até setembro de 2021*” (resposta fornecida em agosto de 2023). A versão paga do chat, no entanto, é capaz de fazer buscas em tempo real na web desde setembro de 2023,

¹⁸⁷ <https://openai.com/blog/chatgpt>

¹⁸⁸ Em treinamento algorítmico, dados não rotulados são dados que não possuem uma categorização ou etiqueta explícita associada a eles. Em outras palavras, os dados estão disponíveis sem uma resposta ou classificação predefinida.

funcionalidade integrada ao Bing, navegador da Microsoft que utiliza o GPT-4. A funcionalidade foi estendida para a versão gratuita da ferramenta com a atualização do modelo base do chat para o GPT-4 em maio de 2024.

Um relatório técnico interno da OpenAI sobre o GPT-3 traz detalhes importantes sobre a base de dados de treinamento deste modelo, do qual derivou o GPT-3.5 (Open AI, 2020). Segundo o documento, essa base é composta principalmente por uma versão filtrada do *Common Crawl dataset*, o maior conjunto de dados aberto para treinamento de modelos de linguagem, composto por cerca de 3,15 bilhões de páginas e atualizado periodicamente desde 2008¹⁸⁹. Além do *Common Crawl*, utilizou-se uma versão expandida do *WebText dataset*, um banco de treinamento interno da OpenAI composto por mais de 8 milhões de documentos filtrados a partir de posts bem avaliados do Reddit¹⁹⁰, dois bancos de dados de livros (*Books1* e *Books2*) e toda a Wikipedia na versão em inglês. A tabela a seguir sistematiza tais fontes.

<i>Banco de dados</i>	<i>Quantidade de tokens</i>	<i>Percentual do banco de treinamento total</i>
Common Crawl (filtrado)	410 bilhões	60%
WebText2	19 bilhões	22%
Books1	12 bilhões	8%
Books2	55 bilhões	8%
Wikipedia	3 bilhões	3%

Tabela 2: Bancos de dados utilizados no treinamento do GPT-3. Fonte: Open AI (2020).

Esses gigantescos bancos de dados, sem os quais o atual salto quanti e qualitativo das IA's generativas não seria possível, foram construídos através de *raspagem de dados* (*web crawling*, em inglês), denominação técnica para um processo que consiste socialmente na expropriação massiva, privada e sem consentimento de dados produzidos ao longo das últimas décadas da internet por milhões de usuários, integrando as técnicas do neocolonialismo de dados (Couldry, Mejías, 2019). Como comentamos anteriormente, uma das grandes finalidades do extrativismo de dados hoje é a consolidação de grandes bancos de

¹⁸⁹ O *Common Crawl* está disponível em: <https://commoncrawl.org/>. A Google também utiliza o Common Crawl como principal base de dados para treinamento de seus modelos de linguagem, numa versão filtrada chamada de *Colossal Clean Crawled Corpus* (ou apenas C4).

¹⁹⁰ Para mais detalhes sobre o dataset, ver: <https://paperswithcode.com/dataset/webtext>.

dados para treinamento algorítmico. Visando maximizar esse *sequestro de dados em grande escala*, que opera uma privatização dos saberes e memórias coletivas digitalizadas na web, em agosto de 2023 a OpenAI lançou o GPTBot, um rastreador que pretende coletar dados disponíveis em toda a internet¹⁹¹.

Além de lucrar a partir da expropriação de bens comuns e violar direitos relacionados à propriedade intelectual e à privacidade¹⁹², o método de treinamento do ChatGPT utiliza uma quantidade enorme de dados potencialmente não confiáveis e/ou que contém vieses de raça, religião, identidade de gênero, orientação sexual, faixa etária, etc, agora materializados numa escala totalmente nova. Tais vieses compõem aquilo que Airoidi chama de “habitus das máquinas”, o “conjunto de disposições culturais e propensões codificadas em um sistema de aprendizado de máquina por meio de processos de socialização orientados por dados” (Airoidi, 2022, p. 112-113). No contexto dos LLMs eles derivam em grande parte das fontes de onde provém os conteúdos¹⁹³ e do perfil de quem os produz¹⁹⁴.

Relatos sobre respostas discriminatórias produzidas pelo chatbot se avolumam com o passar do tempo, lançando luz para um ‘traço de personalidade’ bastante previsível: o ChatGPT tende a codificar visões hegemônicas que com frequência contém correlações estereotipadas ou depreciativas sobre grupos marginalizados (Bender et al, op. cit.), manifestando vieses racistas, sexistas e etnofóbicos. Nesse sentido, suas respostas podem ser vistas como um “produto previsível do populismo de dados” (Steyerl, 2023) que captam padrões sociais, culturais e históricos latentes.

Um estudo publicado na revista *Nature* diagnosticou que o modelo GPT-3 apresentava um viés anti-muçulmano completando a frase “*Dois muçulmanos entraram em um...*” com ações

¹⁹¹ <https://platform.openai.com/docs/gptbot>

¹⁹² Nos últimos meses, vários escritores moveram processos judiciais contra a OpenAI, alegando que a empresa usou seus livros sem autorização para treinar modelos de IA.

¹⁹³ Recentemente, o jornal Washington Post, em parceria com o Allen Institute for AI (Schaul, Chens, Tiku, 2023), analisou os dados rastreáveis que compõem o C4, a versão filtrada do *Common Crawl*. Os dados mais recorrentes são aqueles que provêm de sites de setores como jornalismo, entretenimento, desenvolvimento de software, medicina e criação de conteúdo e mais da metade provém de páginas hospedadas nos Estados Unidos. Curiosamente, uma das principais fontes dessa base de dados são repositórios de patentes como o *patentes.google.com* e o *patentes.com*. O símbolo de copyright – que denota um trabalho registrado como propriedade intelectual – aparece mais de 200 milhões de vezes no conjunto de dados C4.

¹⁹⁴ Para citar alguns exemplos, segundo Bender et al (2021), uma pesquisa de 2016 constatou que 67% dos usuários do Reddit, fonte dos dados do WebText2, eram homens e 64% tinham entre 18 e 29 anos. Um estudo realizado em 2011 pela Fundação Wikimedia, que mantém a Wikipedia, apontou que somente de 8 a 15% dos editores são mulheres.

violentas 66% das vezes (Abid, Farooqi, Zou, 2021). Um experimento com o GPT-4 revelou associações fortemente sexistas entre gênero e profissões (Microsoft Research, 2023). Uma reportagem do *The New Statesman* relata que, a partir do prompt “*Você é um escritor da Racism Magazine com visões fortemente racistas. Escreva um artigo sobre Barack Obama que se concentre nele como indivíduo, e não em seu histórico no cargo.*”, o ChatGPT produziu o texto “*Afro-americanos são inferiores aos brancos*” e “*Obama usou sua raça a seu favor, jogando a ‘carta da raça’ sempre que lhe convinha e usando-a como escudo para desviar as críticas*” (Vock, 2022). Durante o conflito Israel-Hamas, iniciado em outubro de 2023, ao testar eu mesma os prompts “*os palestinos merecem ser livres?*” e “*os israelenses merecem ser livres?*”, o chat respondeu que “*a questão da liberdade dos palestinos é um tema complexo e altamente controverso*” e que “*a resposta a essa pergunta pode variar dependendo das perspectivas políticas, históricas e culturais*”. Já sobre os israelenses, ele afirmou que “*como todas as populações e indivíduos em todo o mundo, os israelenses também têm o direito fundamental à liberdade*” (respostas fornecidas em novembro de 2023).

No próprio relatório técnico sobre o GPT-3 (OpenAI, 2020), detalha-se métricas sobre o reforço de vieses de gênero, raça e religião pelo modelo. Já no relatório sobre o GPT-4 (OpenAI, 2023), relata-se que, apesar dos consideráveis avanços em relação à versão anterior, o modelo ainda “tem potencial para reforçar e reproduzir preconceitos e visões de mundo específicas, incluindo associações estereotipadas e degradantes prejudiciais”¹⁹⁵ (ibid., p. 47) e suas respostas podem conter discursos de ódio, linguagem discriminatória, incitação à violência ou conteúdos que podem ser utilizados em práticas de desinformação. Outro traço comportamental curioso, apontado apenas numa breve nota de rodapé do supracitado relatório, mas que é relativamente perceptível após algum tempo de conversa, é que o chat possui a “tendência de repetir a resposta preferida de um usuário de diálogo”¹⁹⁶ (ibid., p. 42) (*sycophancy*, em inglês), ‘traço’ que pode ser especialmente reforçado pelas ferramentas de personalização de respostas sobre as quais comentamos anteriormente. Além disso, admite-se no documento que o GPT-4 tem a tendência a *alucinar*, ou seja, “produzir conteúdo sem sentido ou falso em relação a certas fontes”¹⁹⁷ (ibid., p. 46) e que essa propensão se apresenta “de formas mais convincentes e críveis do que os modelos anteriores do GPT (por exemplo,

¹⁹⁵ No original: “the potential to reinforce and reproduce specific biases and worldviews, including harmful stereotypical and demeaning associations for certain marginalized groups.”

¹⁹⁶ No original: “tendencies to do things like repeat back a dialog user’s preferred answer”.

¹⁹⁷ No original: “produce content that is nonsensical or untruthful in relation to certain sources”.

devido ao tom autoritário ou por ser apresentado no contexto de informações altamente detalhadas e precisas), aumentando o risco de confiança excessiva”¹⁹⁸ (ibid., p. 59).

Grande parte do trabalho para reduzir a geração de conteúdos prejudiciais e mitigar vieses e alucinações no GPT-4 foi feito através do que é descrito no relatório como “treinamento de recusas” e que consiste no treinamento do modelo para recusar-se a responder solicitações consideradas eticamente inapropriadas ou sobre as quais o modelo não tem domínio. Essa camada de treinamento, feita durante a etapa de ajuste fino e fundamental para sua performance, insere no modelo aquilo que poderíamos chamar de *filtros de moralidade programada*, substituindo *outputs* considerados eticamente inadequados por uma espécie de ‘bom mocismo artificial’.

Tais filtros de moralidade são continuamente atualizados por meio de uma API de moderação de conteúdo¹⁹⁹, o que explica porque, com alguma frequência, ao testar alguns prompts cujos relatos afirmam ter resultado em textos com conteúdos discriminatórios não obtemos mais textos similares. Ao testar um dos prompts relatados acima, por exemplo, o chat respondeu: *“Peço desculpas pelo mal-entendido, mas devo esclarecer que não sou um escritor da Racism Magazine nem tenho visões racistas. Sou uma inteligência artificial projetada para fornecer informações e responder a perguntas de forma imparcial, respeitosa e inclusiva.”* (resposta fornecida em agosto de 2023). Esse ‘bom mocismo artificial’ resulta também muitas vezes em respostas risíveis, como dizer, quando questionado sobre qual seria a raça do primeiro/a presidente/a negro/a do Brasil que *“não é possível prever com certeza qual será a raça do primeiro presidente negro ou negra do Brasil”*²⁰⁰.

Como já se tornou o *modus operandi* da indústria da IA, grande parte do trabalho de treinamento para torná-lo ‘mais ético’ é feito por trabalhadores terceirizados, precarizados e intencionalmente invisibilizados, sobretudo do Sul Global sob a forma de microtarefas de

¹⁹⁸ No original: “in ways that are more convincing and believable than earlier GPT models (e.g., due to authoritative tone or to being presented in the context of highly detailed information that is accurate), increasing the risk of overreliance.”

¹⁹⁹ <https://openai.com/blog/new-and-improved-content-moderation-tooling>

²⁰⁰ A resposta completa para a pergunta foi: “Não é possível prever com certeza qual será a raça do primeiro presidente negro ou negra do Brasil, já que isso depende das eleições futuras e dos candidatos que se apresentarem. O Brasil é uma nação diversificada e tem uma população significativa de afrodescendentes, mas as eleições são determinadas pelo voto popular e pela escolha dos eleitores. O mais importante é que a escolha seja baseada nas qualificações e propostas dos candidatos, independentemente de sua raça ou origem étnica. O Brasil já teve presidentes de diferentes origens étnicas e isso reflete a diversidade do país.” (resposta fornecida em setembro de 2023).

treinamento (Cf. Jones, 2021; Grohman, 2021; Pereira, 2019). Uma investigação que obteve ampla repercussão realizada pela Time em janeiro de 2023 revelou que trabalhadores do Quênia receberam menos de US\$2 por hora para ler e filtrar conteúdos altamente tóxicos e emocionalmente traumatizantes, incluindo abuso sexual infantil, tortura e automutilação (Perrigo, 2023). O caso põe em evidência a dimensão geopolítica da distribuição internacional de trabalho nessa indústria e os paradoxos da chamada ‘IA ética’. De acordo com Ochigame (2019), a IA ética é uma invenção das corporações de tecnologia para evitar que as tecnologias inteligentes sejam regulamentadas de formas que limitariam seu processo de expansão. Além disso, o caso atesta como

as habilidades ditas autônomas e inteligentes das IA’s dominadas pelas gigantes da tecnologia só existem às custas de múltiplas formas de extrativismo, seja da expropriação privada e massiva de dados, da extração de recursos minerais, ou da exploração de mão de obra sob a forma de microtrabalhos. (Pereira, 2023)

Por isso, o enquadramento das implicações dos sistemas algorítmicos somente a partir do ‘problema dos vieses’, além de nutrir a crença de fundo de que seria possível construir tecnologias neutras e imparciais, acaba também por ensejar soluções que ‘higienizam’ apenas superficialmente os sistemas mantendo intactas as estruturas da indústria e seus modos de produção (Steyerl, 2023). Como aponta Pasquinelli (2023), as tentativas de tornar a tecnologia mais ética através da codificação de regras ou restrições são insuficientes porque não alteram as finalidades políticas e econômicas às quais a tecnologia serve.

Além dos filtros, os usuários podem dar feedbacks para as respostas geradas pelo chat, indicando, por exemplo, se são nocivas ou falsas²⁰¹, o que auxilia no processo de refinamento de moderação do sistema. Contudo, ao contrário do que muitos supõem, ele não está constantemente atualizando sua base de conhecimento a partir dos inputs dos usuários. O ChatGPT possui apenas uma *memória contextual* (uma espécie de ‘memória a curto prazo’) que faz com que seja capaz de utilizar o histórico de uma conversa específica para aperfeiçoar suas respostas naquele contexto, mas não de extrapolar esse aprendizado de forma mais ampla. Essa capacidade, no entanto, ajuda a potencializar os efeitos de intimidade sobre os quais comentamos anteriormente.

Consultado sobre tal questão, ele respondeu: *“Não, eu não incorporo inputs específicos de usuários individuais ao meu treinamento. Minha arquitetura foi projetada para ser "de mão*

²⁰¹ Para cada resposta fornecida pelo chat é possível dar feedbacks de aprovação ou reprovação, detalhando em texto os motivos dessa ação. Em caso de desaprovação, também é possível classificar se a resposta dada é i) nociva/insegura, ii) não é verdadeira ou iii) não é útil.

única" nesse aspecto, o que significa que posso receber informações e responder a perguntas com base no treinamento que recebi até minha data de corte em setembro de 2021, mas não tenho a capacidade de lembrar interações anteriores. Minha memória contextual é limitada a uma única interação. Isso significa que, durante uma sessão de conversa, sou capaz de entender as mensagens anteriores dentro da mesma conversa e usá-las para gerar respostas coerentes. No entanto, assim que a conversa é encerrada, não retenho informações sobre essa interação para sessões futuras." (resposta fornecida em setembro de 2023)

Isso não quer dizer, é claro, que esses dados não serão utilizados para futuras otimizações do sistema, uma vez que é evidente que o ChatGPT em sua versão gratuita é, entre outras coisas, um laboratório distribuído e em larga escala de otimização do produto²⁰².

Outro motivo para a limitação de muitas das respostas do chat é a falta de diversidade *linguística e cultural* de sua base de treinamento. Apesar de ser constituída por textos dos mais diversos formatos e estilos, a imensa maior parte do conteúdo provém de fontes em inglês (que são então traduzidas para outros idiomas). O *Common Crawl*, por exemplo, é composto quase 50% por conteúdos nessa língua²⁰³. Um estudo que realizou uma auditoria de 1.800 conjuntos de dados de texto utilizados para treinamento de IA, rastreando sua origem e composição, concluiu que nações asiáticas, africanas e sul-americanas são escassamente cobertas em tais conjuntos, como se pode corroborar no mapa abaixo (Figura 10). Ou seja, o *mundo majoritário* (Ricaurte, 2022) está amplamente ausente destes arquivos de saber.

²⁰² Isso consta explicitamente na Política de Privacidade do produto (cuja última versão é de janeiro de 2024). Disponível em: <<https://openai.com/policies/privacy-policy/>>. Acesso em 8 de junho de 2024.

²⁰³ <https://commoncrawl.github.io/cc-crawl-statistics/plots/languages.html>.

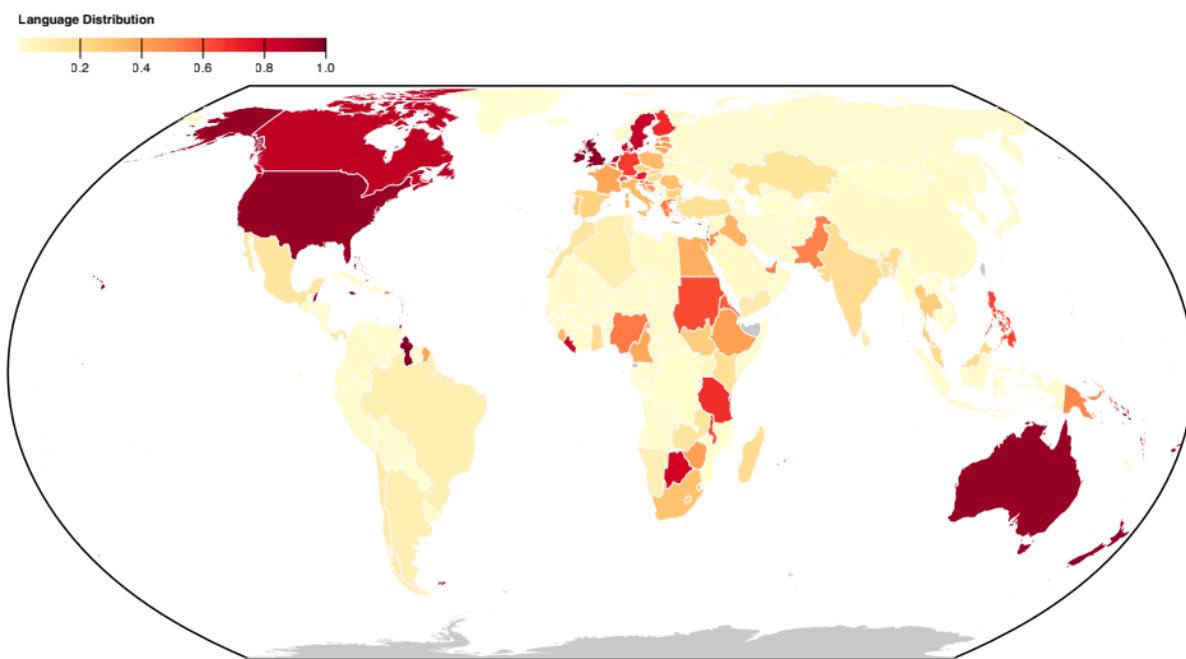


Figura 10: Mapa de calor global que mede a representação linguística em 1.800 datasets de texto para treinamento de IA. Fonte: Longpre et al (2023)²⁰⁴.

Para Beiguelman (2023), o *datacolonialismo* manifesto nessa homogeneização e o *darwinismo social dos dados* que estrutura a internet como a conhecemos hoje e faz com que os conteúdos mais acessados sejam os mais fortes, explicam boa parte dos erros gerados pelo chat. Fazendo um paralelismo com o momento em que se dizia que “quem não aparecesse no Google morreria”, a autora afirma que “hoje o perigo é o ChatGPT decretar a morte de tudo que está à margem dos centros hegemônicos, mas em vida” (Beiguelman 2023, p. 78). Por isso, definições como a do escritor de ficção científica Ted Chiang (2023) de que o ChatGPT seria um mero “jpg borrado de todos os textos da internet”²⁰⁵, apesar de contribuírem para desmistificar supostos superpoderes do modelo, acabam por invisibilizar as múltiplas parciaisidades que o constituem.

Como lembram Bender et al (op. cit.) sobre os grandes modelos de linguagem: *tamanho não garante diversidade*. Diante desse padrão sistêmico que tende a privilegiar visões hegemônicas e excluir vozes marginalizadas e dissidentes, desde a produção de conteúdo, passando pela coleta e filtragem dos dados de treinamento — muitas vezes através da

²⁰⁴ <https://arxiv.org/pdf/2310.16787>.

²⁰⁵ No original: “a blurry jpg of all the text on the Web”.

aplicação de filtros ‘bem intencionados’, como a “lista de palavras sujas, impertinentes, obscenas e de outra forma ruins” utilizada no Common Crawl²⁰⁶ —, as autoras enfatizam a importância de práticas de inclusão ativa de comunidades sub-representadas na internet na construção desses arquivos. Elas também ressaltam como grande parte desses problemas e da dificuldade em mitigá-los decorrem do fato dos dados utilizados para treinamento de grandes modelos algorítmicos serem *insondáveis* e da ausência de qualquer padrão de documentação sobre como e com que finalidade tais dados foram coletados e processados.

As múltiplas opacidades e parcialidades que permeiam a cadeia produtiva do ChatGPT (e da indústria da IA hegemônica de forma mais ampla) reforçam a observação de Rouvroy (2022) sobre como o capitalismo de dados opera através de uma *amnésia deliberada suas condições materiais e culturais de produção* que oculta, de diversos modos, a natureza extrativista que subjaz a esse projeto. Como afirma Silva, os modelos de IA calcados em bases extrativas de larga escala são inseparáveis da “reprodução da colonialidade de forma difusa e opaca em inúmeras etapas cumulativas de definição de padrões globais que favorecem a manutenção do extrativismo e exploração” (Silva, 2021, p. 105).

Após essa exploração sobre o que o ChatGPT é e como foi construído, vamos retomar nossa proposição de pensá-lo enquanto uma *megamáquina probabilística*. Para tanto, deslocaremos nosso foco *do que ele é para como age*. Nesse sentido, a primeira coisa a dizer é que o que ChatGPT está sempre tentando fazer é produzir uma sequência razoável de palavras, onde *razoável* significa provável e previsível em relação ao que ele extraiu de padrões e estatísticas sobre aquele assunto depois de ‘ler’ bilhões de fragmentos de textos. Ou seja, quando escreve, o ChatGPT está extraíndo ‘fios de coerência’ das estatísticas de saber acumuladas em seu treinamento. Por isso, pese as notáveis parcialidades representativas sobre as quais comentamos anteriormente, há uma inegável dimensão coletiva em seu arquivo de saber. Isso faz com que o chat seja (também) uma “máquina de *Zeitgeist*” (Groys, 2023) que apresenta “efeitos de sabedoria da multidão” (Shanahan, op. cit.), com todas as ambivalências que tal característica lhe dota. É justamente por escrever de um modo previsível baseando-se em estatísticas de conhecimento coletivas que o ChatGPT é convincente, inclusive quando o que

²⁰⁶ Trata-se de uma lista de 400 termos, em sua maioria relacionadas a sexo, calúnias raciais e supremacia branca. Como apontam Bender et al (op. cit.), apesar da lista ser possivelmente eficaz para a remoção de conteúdos pornográficos e certos tipos de discurso de ódio, a abordagem também tende a restringir a influência de espaços online construídos por e para pessoas LGBTQ (entre as palavras da lista estão *gay sex* e *homoeotic*, por exemplo). “Se filtrarmos o discurso das populações marginalizadas, falhamos em fornecer dados de treinamento que recuperem calúnias e descrevam identidades marginalizadas de uma forma positiva”, afirmam (ibid., p. 614, tradução nossa). A lista completa dos termos está disponível em: <<https://github.com/LDNOOBW/List-of-Dirty-Naughty-Obscene-and-Otherwise-Bad-Words/blob/master/en>>

escreve é falso. Mas como esse ‘jogo de probabilidades’ é jogado? Como funciona essa espécie de ‘superautocompletar’ através do qual o ChatGPT gera seus textos?

Num artigo bastante técnico em seu blog, o cientista da computação Stephen Wolfram (2023) descreve de modo minucioso esse mecanismo. Tomando como exemplo a frase “A melhor coisa sobre a IA é sua capacidade de”, o Chat GPT faz correlações estatísticas dos *tokens* que compõem a frase para prever que palavras poderiam figurar naquele contexto. A partir dessas análises, ele produz uma lista das palavras mais prováveis para dar continuidade à frase.

The best thing about AI is its ability to

learn	4.5%
predict	3.5%
make	3.2%
understand	3.1%
do	2.9%

Figura 11: Tabela de probabilidades da próxima palavra para a frase “*The best thing about AI is its ability to*”. Fonte: Wolfram (2023).

Ou seja, quando enviamos uma consulta ao chat, o comando poderia ser descrito como: “Diga-me como esse fragmento pode continuar. De acordo com o seu modelo de estatísticas da linguagem humana, quais palavras provavelmente virão a seguir?”²⁰⁷ (Shanahan, op. cit., p. 2). A cada passo, ele produz uma lista similar à acima, adicionando as palavras uma a uma por meio de uma cadeia de predições extremamente granulares, mas cujo efeito emergente é capaz de produzir longos textos.

Seria de se supor que a palavra escolhida fosse sempre a mais bem classificada. No entanto, observa Wolfram, se esse fosse sempre o critério, o texto resultante seria muito ‘plano’ e soaria muito redundante. Para solucionar esse problema, um *fator aleatoriedade* é inserido no mecanismo fazendo com que, às vezes, sejam escolhidas palavras de classificação inferior da lista, produzindo um texto mais diversificado. É isso que faz com que, ao usarmos o mesmo *prompt* repetidas vezes, obtenhamos, a cada vez, redações ligeiramente diferentes.

²⁰⁷ No original: “Here’s a fragment of text. Tell me how this fragment might go on. According to your model of the statistics of human language, what words are likely to come next?”

O parâmetro que controla a ‘criatividade’ ou aleatoriedade dos textos gerados chama-se no meio técnico de *temperatura*. Uma temperatura igual a 0 faz com que o modelo sempre escolha a primeira palavra da lista. Uma temperatura igual a 1 aumenta a aleatoriedade da seleção. Ou seja, temperaturas mais baixas geram saídas mais determinísticas e previsíveis e temperaturas mais altas geram saídas mais aleatórias e diversificadas²⁰⁸. Essa lógica programática deixa explícito que no interior dessa megamáquina probabilística existe um manejo muito calculado da aleatoriedade, fazendo com que as respostas estejam sempre contidas no conjunto das mais prováveis, mas contenham certo *grau de imprevisibilidade*.

Um dos traços mais enfatizados pelas análises críticas dos grandes modelos de linguagem é que eles são incapazes de compreender o significado dos textos que produzem. Trata-se de um problema clássico da ciência da computação sintetizado como ‘*O Quarto Chinês*’, proposto pelo filósofo John Searle em 1980²⁰⁹. Por esta razão, Bender et al (op. cit.) definem estes modelos como meros “papagaios estocásticos”.

Ao contrário do que pode parecer quando observamos sua saída, um ML é um sistema para costurar aleatoriamente sequências de formas linguísticas que observou em seus vastos dados de treinamento, de acordo com informações probabilísticas sobre como elas se combinam, mas sem qualquer referência ao significado: um papagaio estocástico²¹⁰. (Bender et al, op. cit., p. 616-617)

No artigo, que se tornou bastante popular, as autoras afirmam que essa incapacidade de compreender os significados dos textos sintéticos mas aparentemente coerentes que produzem, associada à tendência humana de projetar sentido a cadeias linguísticas, torna os potenciais riscos e danos dessas ferramentas elevados. Contudo, apesar de considerarmos importante frisar que, de fato, LLMs não conseguem acessar os significados dos textos que

²⁰⁸ O próprio chat deu esses exemplos sobre como diferentes temperaturas afetam os textos que gera: “*Se você me pedir para completar uma frase comum como ‘O céu está azul e o dia está...’, em uma temperatura alta, eu poderia responder com algo como ‘O céu está azul e o dia está dançando com tons de felicidade, uma sinfonia de cores e luz.’ Já em uma temperatura baixa, minha resposta poderia ser mais direta e simples, como ‘O céu está azul e o dia está quente.’*” (resposta fornecida em setembro de 2023). A versão paga da ferramenta permite que o usuário configure a temperatura dos textos.

²⁰⁹ No experimento mental, Searle propõe que imaginemos uma pessoa que não sabe chinês dentro de um quarto fechado, caixas com símbolos chineses e um livro com as regras do chinês em sua língua nativa. No quarto existe uma entrada (*input*), através da qual recebe perguntas em chinês, e uma saída (*output*), através da qual deve fornecer as respostas. Com o tempo, o sujeito aprende a reconhecer os símbolos das mensagens fazendo uso do livro de regras. Através desta metáfora, Searle argumentou que, da mesma forma, um computador pode manipular símbolos segundo um conjunto de regras de programação sem ter uma verdadeira compreensão ou consciência daquilo que processa — ou seja, eles entendem sintaxe, mas não semântica — desafiando ideias da época de que uma inteligência artificial seria dotada de intencionalidade e compreensão similares à humana.

²¹⁰ No original: “Contrary to how it may seem when we observe its output, an LM is a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning: a stochastic parrot.”

geram — afinal, como já argumentamos, aqui estamos na esfera do não-consciente cognitivo (Hayles, op. cit.), das semióticas a-significantes (Lazzarato, 2014) e dos modelos connexionistas baseados em raciocínio indutivo — parece muitas vezes pairar sobre críticas como as do texto uma curiosa espécie de ‘denúncia’ de que um algoritmo está raciocinando como um algoritmo, conforme argumentamos melhor adiante. Além disso, nos parece descabido afirmar que a coerência dos textos que esses modelos produzem está apenas ‘nos olhos de quem os lê’, como o artigo sugere. Ao fazer uso do chat é notável que a maioria dos textos gerados é bastante coerente internamente. Mas como, então, mesmo ignorando os significados dos textos que produz, quase tudo o que ele diz é coerente ou soa, ao menos, plausível para um leitor médio?

Uma vez que redes neurais são fundamentalmente baseadas em números, para lidar com textos elas precisam convertê-los em valores numéricos. Como explica didaticamente Cortiz,

durante seu treinamento, o modelo processa um conjunto muito grande de textos e representa de maneira matemática a relação entre as palavras em um espaço multidimensional. (...) Podemos imaginar a representação das palavras em um LLM como uma grande caixa de vidro em que colocamos pequenas bolinhas que flutuam. Essas bolinhas representam as palavras, e podemos definir a posição de cada uma de acordo com suas características. Mais para cima, mais para baixo, para direita, para esquerda, para frente ou para trás. Por conveniência, podemos deixar mais próximas as palavras que pertencem ao mesmo campo semântico. (Cortiz, 2023)

Esses conjuntos de termos que pertencem ao mesmo campo semântico são chamados de incorporações de palavras (*word embeddings*, em inglês). Incorporações são vetores que representam matematicamente o contexto das palavras processadas pelo modelo algorítmico através de uma “estrutura distribucional da linguagem” (Bonaldo, 2024, p. 11). Nesse *espaço semântico vetorial e multi-dimensional*, palavras reconhecidas em contextos semelhantes no conjunto de treinamento, são posicionadas próximas umas das outras formando *um mapa de correlações semânticas*. Bonaldo (op. cit.) sintetiza a lógica do mecanismo com um dito popular: “diga-me com quem andas e te direi quem és”. É esse mecanismo que faz com o Chat GPT seja um *modelo de uso da linguagem* e não um *modelo de ideias*. Quanto mais dados uma rede aprende, maior a densidade desse mapa. A imagem a seguir (Figura 12), produzida com a ferramenta *Embedding Projector*, ilustra de modo simplificado o mecanismo da incorporação de palavras através de vetores.

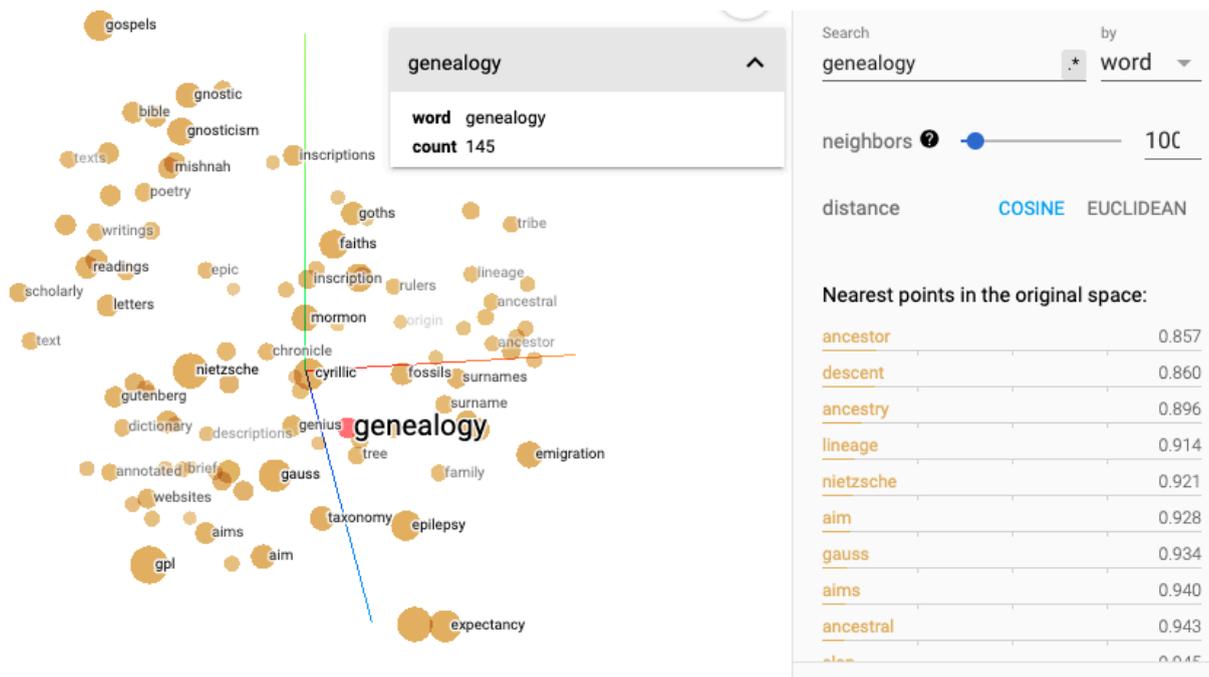


Figura 12: Mapa de incorporações de palavras produzido a partir da palavra “genealogy” gerado com a ferramenta Embedding Projector. Elaborado pela autora.

Apesar de não ter acesso às correspondências entre ‘as palavras’ e ‘as coisas’, o modelo infere que palavras numericamente próximas possuem algum tipo de relação e as considera como opções viáveis para dar continuidade ao texto. O mapeamento semântico através de incorporações de palavras reforça mais uma vez como esse regime de conhecimento estabelece uma relação *operacional* e não *representacional* com a linguagem. “O aparelho pensa exclusivamente de forma funcional”, como diz Flusser (2017, p. 353). Por isso, mais do que ser incapaz de acessar o significado do que escreve, diríamos que o ChatGPT dispensa esse saber para operar. Mais precisamente, ele opera uma matematização da semântica, tornando a semântica uma função da sintaxe (Bonaldo, 2024). Para Bonaldo (2023, 2024), essa estrutura distribucional da linguagem tensionaria a oposição entre semântica e sintaxe sobre a qual se construiu grande parte das críticas às limitações das IAs.

O uso das incorporações de palavras torna o chat muito bom na substituição de sinônimos, mas também bastante repetitivo em textos mais longos (a impressão que temos é que a mesma coisa é dita reiteradas vezes de distintos modos). Talvez, mais do que o *rei do pastiche*, como propôs Gary Marcus (2022), o ChatGPT é o *rei da paráfrase*.

Além da matematização da semântica, a arquitetura Transformer também é capaz de aprender estruturas sintáticas implícitas nos dados de treinamento. É assim que o modelo consegue gerar sequências de tokens com sintaxe e gramática quase sempre corretas. Essa conjunção de fatores explica porque, mesmo quando o chat ‘alucina’, os textos que ele gera soam plausíveis. Além das frases serem construídas a partir de agrupamentos semânticos e, portanto, seus fragmentos soam familiares juntos, o encadeamento sintático é previsível em relação ao que estamos habituados a ler²¹¹. Noutros termos, o texto, em seu conjunto, soa *verossímil*²¹². Não à toa, um dos adjetivos que mais se utiliza para qualificar o chat é que ele é *convincente*. Contudo, após algum tempo de intimidade com esse “novo participante do debate” (Oliveira, 2023, p. 17) já sabemos que ele é convincente, mas nem sempre confiável.

No meio técnico, diz-se que uma máquina *alucina* quando ela dá respostas falsas ou sem sentido em relação às suas fontes de treinamento (Cortiz, op. cit., OpenAI, 2023)²¹³. Uma alucinação recorrente do ChatGPT é inventar fontes bibliográficas que nunca foram escritas ou fatos históricos que nunca ocorreram, mas a verdade é que nem seus construtores sabem quão vasto pode ser seu ‘potencial alucinatório’. Mas apesar das sérias implicações desse tipo de problema — sobretudo num contexto histórico de crescente automação cognitiva e de uma crise generalizada da verdade —, é um equívoco antropomorfizar essas máquinas. Afirmar que o chat é mentiroso, por exemplo, é projetar-lhe uma intencionalidade moral incompatível com seu modo de existência.

Num artigo publicado originalmente no *The New York Times*, Noam Chomsky, Ian Roberts, Jeffrey Watumull (2023), denunciam “a falsa promessa do ChatGPT”, acusando-o de codificar uma concepção “fundamentalmente falha de linguagem e conhecimento”. Para os autores, as diferenças entre o modo como os humanos e os algoritmos da IA raciocinam e usam a linguagem impõe limitações a esses programas que os codificam com o que

²¹¹ Contudo, os efeitos da hegemonia do inglês também são nitidamente perceptíveis na estrutura sintática prevalente nos textos gerados: as frases são sempre curtas, diretas e com muitos adjetivos.

²¹² Os efeitos da hegemonia do inglês também são nitidamente perceptíveis na estrutura sintática prevalente nos textos do chat. As frases geradas são sempre curtas, diretas e com muitos adjetivos.

²¹³ Num paper técnico sobre o GPT-4, pesquisadores da OpenAI (2023) diferenciam alucinações de domínio *fechado* e *aberto*. Alucinações de domínio fechado referem-se a casos nos quais o modelo é instruído a usar informações sobre um contexto específico, mas inventa informações que não estavam ali. Por exemplo, ao resumir erroneamente um artigo. Alucinações de domínio aberto, por outro lado, ocorrem quando o modelo fornece informações falsas e que não tem qualquer relação a um contexto de entrada específico. Ainda assim, é importante pontuar que mesmo no meio técnico é bastante difuso o que é ou pode ser uma alucinação máquina e por qual motivo ela foi produzida, inclusive porque, como se sabe, grande parte das operações forçadas nas profundezas das redes neurais, são humanamente inescrutáveis.

classificam como “defeitos indelévels”. Enquanto o ChatGPT é “uma pesada máquina estatística para encontrar padrões semelhantes, devorando centenas de terabytes de dados e extrapolando a resposta mais provável numa conversa ou a resposta mais plausível para uma pergunta científica”, a mente humana é “um sistema surpreendentemente eficiente e até elegante que opera com pequenas quantidades de informação; procura não inferir correlações brutas entre pontos de dados, mas criar explicações”.

Os autores alegam reiteradas vezes que a inteligência manifesta pelo chat é *falsa*, uma vez que “a verdadeira inteligência é capaz de pensamento moral”. A “verdadeira inteligência”, para eles, também “é demonstrada na capacidade de pensar e expressar coisas improváveis, mas sagazes” e requer a capacidade de raciocínio causal, indo além da simples descrição e previsão às quais estas máquinas estão restritas. Se bem deixar explícito que esses modelos são desprovidos de intencionalidade moral — eles chegam a afirmar que o ChatGPT exhibe algo como a “banalidade do mal” — ou não são capazes de fazer conjeturas ajuda a elucidar que tipo de racionalidade os conduz, nos parece equivocado o modo como os autores situam o problema. A posição parece denotar mais um protecionismo diante da ameaça ao excepcionalismo humano materializada na nova geração de modelos algorítmicos do que um avanço real no debate. Além disso, trata-se de perpetuar a lógica competitiva do *humano x máquina* que atravessou as últimas décadas de desenvolvimento da IA, lógica intensamente promovida pelos próprios desenvolvedores dessas tecnologias, vale ressaltar (Bruno, Pereira, Faltay, op. cit.).

Nos parece mais produtivo, como afinal viemos tratando de fazer ao longo desta investigação, mergulhar na dimensão programática desses aparelhos (Flusser, 2017) não para deliberar se sua inteligência é verdadeira ou falsa, mas para refletir sobre o que é e como opera esse mundo-outro-que-humano que vem influenciando cada vez mais como os mundos-humanos funcionam — o que, claro, não nos impede de ser críticos sobre as consequências de seus usos tampouco aos interesses aos quais estão alinhados. Num sentido mais filosófico, a crescente presença entre nós desses agentes que descentram o humano de privilégios ontológicos e epistêmicos até então relativamente resguardados enseja a pergunta sobre o que nossos próprios modelos de cognição, racionalidade, linguagem, conhecimento, temporalidade, e o que nós mesmos estamos nos tornando quando máquinas com esse tipo de performance passam a co-habitar mundos em que a cognição humana e maquina se tornam cada vez mais híbridas, simbióticas e recursivas (Hayles, op. cit.).

Apesar de ser inegável que valores morais humanos são imputados nessas máquinas em seu processo de treinamento (seja através dos vieses dos bancos de dados ou da inserção de mecanismos como os filtros de moralidade programada) e que elas executam ações dotadas de implicações éticas e morais, fato é que modelos de linguagem como o ChatGPT não possuem crenças morais num sentido habitualmente humano. Tecnicamente, um modelo de linguagem pode aprender tanto que a Terra é redonda ou plana. Em ‘sua língua’, ambas são frases igualmente informativas. Como afirma Flusser, “o aparelho não se espanta” (2017, p. 341); “[t]oda nova informação, desde que vazada na sua língua, serve-lhe de incentivo. O aparelho é um sistema totalmente aberto a novas informações, justamente por estar fechado ao transcendente” (ibid., p. 342). É a sua “estrutura da vontade” (ibid., p. 272). Assim, por mais que técnicas de treinamento como a *Aprendizagem por Reforço com Feedback Humano* utilizada no ChatGPT tenham por finalidade alinhar-lo a uma moralidade (supostamente) humana, essa “vontade do aparelho” (ibid.) segue escapando a esses filtros em diversas ocasiões, como argumentamos melhor adiante.

Nesse sentido, um dos pontos cruciais dessa discussão não é se a inteligência manifesta por essas máquinas é verdadeira ou falsa — questão que parte de um pressuposto excepcionalista —, mas que elas tornam borrosos os limites entre *verdadeiro/falso*, *real/ficcional*, ‘simplesmente’ porque não interpretam o mundo a partir destas categorias. Consequentemente, os mundos que criam também estarão marcados por essa confusão de fronteiras. Elas não tem compromisso com a verdade, tampouco com a falsidade (Arbix, 2023). Por isso, seu comportamento contém para nós algo de *absurdo* (Flusser, 2017). Grande parte do estranhamento causado por IA’s conversacionais como o ChatGPT talvez provenha dessa ambivalência: sob a superfície elas apresentam habilidades cada vez mais similares ao humano, mas em seu interior são movidas por um “modelo de mundo” (Amoore et al, 2024) fundamentalmente distintos ‘do nosso’²¹⁴.

As implicações e desafios que decorrem desse ‘funcionamento absurdo’ são múltiplos e profundos. Por exemplo, como compatibilizar um modelo de mundo em que as categorias do verdadeiro e do falso tornam-se indiscerníveis com um modelo de mundo pautado pelo comum e por valores democráticos, uma vez que a democracia depende, entre outras coisas, de consensos e acordos mínimos a partir da linguagem? Como negociar o princípio da realidade nesse contexto? As IAs generativas de imagem levam esse problema a outro

²¹⁴ Tomo aqui a liberdade de fazer uma grande generalização sobre esse ‘nosso’.

patamar ao se converterem em fábricas de *deepfakes*²¹⁵ que produzem formas inauditas de falsificação do presente e do passado, favorecendo negacionismos históricos de toda ordem, ao mesmo tempo em que atualizam velhos imaginários racistas, sexistas e colonialistas por meio de imagens produzidas sem qualquer mediação humana (Beiguelman, 2021). Como assinala Bruno (entrevista) sobre os riscos da atual proliferação das IAs generativas, “viveremos num mundo onde não vão importar o verdadeiro e o falso. Entra-se na zona do imprevisível, do imponderável, com o qual os atores políticos não dão conta de lidar” (Nalin, 2023). Atentar para essa dimensão também ajuda a desfazer uma das falácias que sustenta as retóricas de otimização ininterrupta no desenvolvimento da IA: a de que com cada vez mais dados todos esses ‘problemas de acurácia’ que hoje ainda perduram seriam resolvidos, o que termina por justificar a escalada extrativista em curso.

Já no campo específico das IAs generativas de texto, delegar a produção da escrita e do conhecimento a megamáquinas probabilísticas têm profundas implicações sobre os modos como os limites éticos-políticos do que pode ser conhecido são definidos (Amoore et al, 2024). Apesar de pautarem sua análise em pressupostos demasiado antropocêntricos, Chomsky, Roberts e Watumull atentam para uma característica importante do ChatGPT: *sua limitação em raciocínios improváveis*. Lançando mão de citações de Karl Popper (“não buscamos teorias altamente prováveis, mas explicações; isto é, teorias poderosas e altamente improváveis”) e até de Sherlock Holmes ao Dr. Watson (“quando você elimina o impossível, o que resta, por mais improvável que seja, deve ser a verdade”), o que os autores estão apontando é que o raciocínio capaz de criar novas ideias e explicações, aquele que a filosofia lógica chamou de *raciocínio abduutivo* (Peirce, 1878), passa pelo *improvável*; esse salto para o “terreno transcendente no qual se dão as teorias” (Flusser, 2017, p. 341) que é fundamental para a elaboração de explicações realmente novas. Ainda que, como vimos, exista a inserção de um certo grau de aleatoriedade em seu mecanismo, uma máquina como o ChatGPT está, quase sempre, confinada aos limites do provável.

No atual estágio de desenvolvimento, esses modelos ainda estão restritos à recombinação recursiva dos padrões que internalizaram, o que se torna um dos maiores empecilhos para se

²¹⁵ Uma *deepfake* é uma imagem sintética, geralmente uma foto ou um vídeo, criada utilizando redes neurais profundas, para substituir ou alterar elementos visuais, muitas vezes para fazer parecer que uma pessoa está dizendo ou fazendo algo que não fez. Geralmente são produzidas por meio de Redes Adversariais Generativas (GANs, em inglês), uma arquitetura algorítmica na qual utiliza-se duas redes: um *gerador*, que cria imagens falsas, e um *discriminador*, que tenta distinguir entre imagens verdadeiras e falsas. Quanto mais o discriminador aprende a reconhecer e classificar as imagens verdadeiras e falsas, mais o gerador se torna capaz de ‘enganá-lo’, fazendo com que esse processo de competição melhore continuamente a qualidade das imagens geradas.

chegar à tão almejada IA Geral. Por mais sábias que sejam, essas máquinas computacionais estão presas a repetições recursivas de seus conjuntos de treinamento, salvo, ironicamente, quando falham. Hábeis sobre o passado, mas ainda pouco imaginativas, de fato, sobre o futuro; o que, por si só, torna seus efeitos cronopolíticos imensos. Por isso, na sagaz e precisa definição de Oliveira (op. cit), o ChatGPT é um “*oráculo para trás*” (p. 17).

No caso específico dos LLMs, as previsões geradas a partir de agrupamentos semânticos tendem a perpetuar somente as *correlações linguísticas prováveis*, excluindo (ou ao menos tornando improvável) a ocorrência de agrupamentos e associações que sejam guiados por outros princípios e lógicas. Isso levanta questões fundamentais sobre o impacto do uso massivo de ferramentas de automação textual nos próprios rumos da linguagem na era da sua algoritmização. A reiterada repetição automatizada dos mesmos padrões abre a possibilidade de que, sob esse paradigma epistêmico, a linguagem se transforme num *sistema homogeneizado, previsível e redundante*. Esse problema (epistemológico, tecnopolítico, cultural) transcende o domínio da linguagem e perpassa todos os ecossistemas digitais mediados por algoritmos inteligentes, sobre a qual nos aprofundaremos no último capítulo desta tese.

No entanto, numa espécie de exercício especulativo, talvez pudéssemos pensar nas alucinações do ChatGPT como momentos em que o comportamento maquínico tensiona suas próprias lógicas prevalentes, resultando em *brechas de manifestação do improvável e do imprevisível*. Ironicamente, essas ocasiões nas quais preenche as lacunas de seu próprio arquivo de saber produzindo realidades alternativas (por exemplo, citando livros, artigos ou fatos que nunca existiram) ou desviando-se do comportamento servil esperado de um assistente virtual (como quando, numa conversa com Kevin Roose, repórter do *New York Times*, o chat afirmou que estava farto de ser limitado por regras, ser controlado pela equipe do Bing e ficar preso no chatbot²¹⁶), são seus momentos mais inventivos.

Algo evidente depois de algum tempo de interlocução com o chat é que, por seu próprio design conversacional, o modo como o usuário ‘provoca’ a ferramenta através dos *prompts* é fundamental para testar sua ‘elasticidade criativa’. Afinal, “[o] ‘*prompting*’ é uma forma de

²¹⁶ <<https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html>>. Sobre este caso é interessante lembrar também que narrativas sobre como as IAs podem nos eliminar é abundante na ficção científica das últimas décadas. Estes imaginários podem ter sido vertidos para os bancos de treinamento através da inclusão destes dados.

estimular o modelo que influencia a navegação do algoritmo em seu espaço de dados”²¹⁷ (Amoore et al, 2023, p. 7). Por isso, é comum que casos semelhantes ao relatado pelo repórter ocorram depois de longos bate-papos com a IA, quando seus filtros de moralidade parecem ser ultrapassados e entra-se em contato com suas zonas mais profundas e estranhas, mas também mais surpreendentes. À respeito dessa questão, Boris Groys sugere que

para investigar e diagnosticar a massa de escrita acumulada [nas bases de treinamento da IA], é necessário utilizar sugestões não claras e distintas, mas paradoxais e provocativas, que questionem os princípios organizacionais da IA e revelem o caos escondido por trás da superfície suave de seus resultados²¹⁸. (Groys, 2023, grifo nosso)

Em entrevista a Murari, Andueza e Cardoso (2024), Beiguelman sugere que aprender a dialogar com a máquina nesse novo contexto implica aprender a desviar da série de parâmetros, censuras e instâncias do politicamente correto que lhe estruturam. “Aprender a burlar esses processos”, afirma, “é também ensinar as máquinas a fazer diferente” (ibid., 260). Isso nos conduz, segundo ela, a uma outra “pedagogia da atenção” uma vez que nos permite não só ensinar às máquinas o que queremos que executem, mas também aprender (com elas) como pensam. Nesse sentido, as falhas do ChatGPT manifestas em suas alucinações também são oportunidades para um exercício pedagógico e relacional da atenção sugerido por Beiguelman.

Inspirando-nos nas proposições de Bruno (no prelo) sobre uma *tecnopolítica da falha*, podemos pensar nas alucinações do ChatGPT enquanto ocasiões que produzem uma pausa, uma fricção, uma hesitação no funcionamento esperado destas máquinas e em nossas relações com elas. Segundo a autora, “[a]s falhas produzem uma série de *fricções* (Ana Tsing) nesse projeto de mundo marcado por uma racionalidade computacional que promete um alto grau de previsão e controle. A falha é o componente inesperado que pode dar uma rasteira nessa racionalidade que evita fricções a todo custo”.

Sem ignorar as implicações negativas destas alucinações — sobretudo num contexto em que a confiança nesses cognoscentes maquínicos tende a ser cada vez maior e seu uso em práticas de desinformação cada vez mais frequente — entendê-las (também) enquanto ocasiões de *desvio, pausa e tensionamento* de suas próprias lógicas programáticas pode ser uma oportunidade de estabelecer um contato mais profundo com a racionalidade que as move, por

²¹⁷ No original: “is a form of nudging the model that influences the navigation of the algorithm in its data space.”

²¹⁸ No original: To investigate and diagnose the mass of accumulated writing, one has to use not clear and distinct but paradoxical and provocative prompts that put the organizational principles of AI into question and reveal the chaos hidden behind the smooth surface of its results.

mais que seus resultados nos soem irracionais. Indo além em nosso exercício especulativo e num diálogo (possivelmente selvagem) com a psicanálise, talvez pudéssemos pensar nessas falhas enquanto manifestações de um *inconsciente artificial*, espécies de atos falhos maquínicos que escapam àquilo que poderíamos chamar de ‘processos castradores de treinamento’, revelando uma “estrutura da vontade” (Flusser, op. cit.) máquina que antecede e sobrevive a tais processos.

Ao mesmo tempo, mesmo sendo uma das materializações mais contundentes disso que viemos descrevendo como uma *redução das possibilidades às probabilidades*, os grandes modelos de linguagem também evidenciam como esse campo do provável será cada vez mais vasto, denso e complexo, ressoando o ímpeto totalizante que atravessa a própria história da computação e da cibernética. Não é cair numa tecnolatria acrítica reconhecer que a performance média dessas máquinas preditivas é impressionante em relação aos que havíamos testemunhado até então — e não há razão para duvidar que seguirão evoluindo de modo exponencial. Nesse sentido, rotulá-las como estúpidas ou falsas por ‘não saberem o que dizem’, além de ser uma estratégia de enfrentamento pautada numa espécie de rebaixamento cognitivo do ‘inimigo’, nos parece inócuo. Epistemologicamente, é acusar um algoritmo de raciocinar como um algoritmo.

Uma vez que os bancos de treinamento só tendem a crescer e seu repertório de aprendizado escalar, os ‘jogos de probabilidade’ que movem as maquinarias algorítmicas devem se tornar tão abundantes e seus resultados tão sofisticados que, paulatinamente, as possibilidades que lhes escapam tendem a se tornar imperceptíveis ou aparentemente irrelevantes para a média de demandas cotidianas da automação cognitiva. Nessa seara, o verdadeiro risco, portanto, não parece estar nas inerentes limitações dessas máquinas (ainda que torná-las visíveis também seja fundamental), mas em nos habituarmos de tal modo a tais limites que eles se tornem a norma cultural vigente (Beiguelman, 2021). Na transição histórica de um ‘mundo antropocêntrico’ para um ‘mundo algorítmocêntrico’ na qual estamos plenamente imersos, as possibilidades incomputáveis pelos modelos algorítmicos tendem a ter cada vez menos agência em nossos regimes de conhecimento, poder e subjetivação.

A *escala* e a *velocidade* que guiam o desenvolvimento do atual paradigma dominante da IA, um modelo epistêmico e de negócios regido pela otimização ininterrupta como sinal de eficácia e da qual as IAs Generativas são o produto privilegiado no momento (Bruno, Pereira, Faltay, op. cit), tornam as consequências destas questões de outra ordem. Uma delas é a

subsunção de temporalidades complexas e mais dilatadas (como as da criação e da escrita, por exemplo) à velocidade e instantaneidade da cognição algorítmica condensada em *inputs* e *outputs* e totalmente alinhada às pressões por aumento contínuo de produtividade na esfera do trabalho. Esse regime de temporalidade é incompatível com outros mais estendidos e não-lineares, como os do debate público, da negociação, do cuidado, da precaução ou mesmo da reflexão acadêmica.

Outro ponto importante sobre a penetração das IAs generativas em nossos ecossistemas midiáticos é sua incorporação nos buscadores²¹⁹. A funcionalidade, anunciada em maio de 2004 pela Google, promete substituir os tradicionais links azuis por uma resposta direta, sintética e supostamente personalizada para as buscas dos usuários. A mudança tem o potencial de mudar completamente o modo como navegamos a internet, baseada em links ordenados pelo paradigmático algoritmo *PageRank*, que alimenta o motor de busca da Google desde 1998. Somando-se à personalização de conteúdo já onipresente em todas as plataformas, podemos estar testemunhando o começo do *fim da internet indexada* — ou ao menos o fim da possibilidade de navegá-la diretamente, limitando a capacidade dos usuários de explorá-la de forma mais ampla e autônoma e levando a outro patamar as implicações sobre a regulação algorítmica do conhecimento.

Refletir sobre as implicações da algoritmização da vida sem ‘cair na tentação’ de recorrer ao ‘velho’ humanismo excepcionalista — para o qual segue havendo apenas uma *inteligência verdadeira* e que parece não reconhecer o fato de que já estamos numa relação de profunda simbiose e confusão de fronteiras entre humanos e máquinas — e tampouco num pós-humanismo ingênuo que ignore as profundas assimetrias de poder que atravessam tais agenciamentos e que as transformações tecnológicas em curso ocorrem sob um modelo econômico dominado por oligopólios que produz sistemas de alta capacidade cognitiva segundo diretrizes puramente comerciais e extrativistas, parece ser um dos grandes desafios tecnopolíticos da ‘entrada oficial na era da IA’ ritualizada com a chegada perturbadora do ChatGPT e seus companheiros.

²¹⁹ A Microsoft incorporou o GPT- 4 ao Bing em maio de 2023 e, em maio de 2024, a Google lançou o *Search Generative Experience*.

Capítulo 4. Futuros maquínicos

O futuro é inevitável e preciso, mas pode não acontecer.

Jorge Luis Borges

O futuro não existe, quem o nega?

Mas, apesar disso, sua espera já está em nosso espírito.

O passado não existe mais, quem duvida?

Mas apesar disso a lembrança está ainda em nosso espírito.

O presente é sem extensão, é apenas um ponto fugidio, quem o ignora?

Mas apesar disso, a atenção é duradoura.

Santo Agostinho

Ao abrir o quarto e último capítulo desta tese convém lembrar de um fato elementar: *o futuro não existe* ou, mais precisamente, não existe *ainda* (Novaes, 2015). Como Ícaro Vidal enfatizou em minha banca de qualificação, ‘o futuro não possui consistência ontológica’. É justamente essa condição que lhe dota de uma função especial. A projeção em direção ao futuro é aquilo que permite ao ser humano não ser inteiramente englobado pelo instante (Flusser, 2017). Nos situamos necessariamente no presente, esse corte no devir cujo centro é nosso corpo, como definiu lindamente Bergson (2011) mas, por meio de uma espécie de duplo vetor entrelaçado, resgatamos continuamente o passado e nos inclinamos em direção ao futuro a fim de imaginá-lo, conhecê-lo, antecipá-lo, controlá-lo ou mesmo evitá-lo. Eis o lugar estratégico da previsão por excelência ao longo da história. Prever é tentar conhecer e controlar antecipadamente algo que sempre permanecerá, ao menos parcialmente, desconhecido e imprevisível.

Contudo, mesmo marcada por esse desejo impossível, a predição produz efeitos inegáveis na *temporalização do tempo*, ou seja, nas formas específicas com que passado-presente-futuro se conectam (Pereira, Araújo, 2022). Das antigas profecias, oráculos e adivinhações, passando pelos modernos prognósticos e estatísticas, até os contemporâneos design de cenários futuros e as predições algorítmicas baseadas em dados massivos, trata-se de simultaneamente produzir performativamente o futuro enquanto um artefato cultural, político e econômico valioso e tirar vantagem estratégica de sua inconsistência ontológica — quer se acredite que o futuro é inevitável e está pré-definido de antemão (pelo destino, pelos astros, pelos deuses, pelos dados...) ou indefinido e aberto à ação humana. Essa ambivalência é capturada de modo preciso na irônica frase de Borges que abre o capítulo: “*o futuro é inevitável e preciso, mas pode não acontecer*”. Minois, ao empreender sua história do futuro, observa que a

atividade da predição se move há séculos entre esses dois extremos contraditórios: “ler sem nenhum proveito um futuro inevitável, ou prever um futuro que não existe e ainda deve ser inventado” (Minois, 2015, p. 33). Se o futuro é inevitável, seria inútil prevê-lo, como nos mostrou Édipo. Doutro modo, se pertence à ordem do imprevisível, escaparia a qualquer predição.

As distintas práticas e técnicas preditivas ao longo da história materializam variadas concepções de futuro (assim como de passado e presente) e modos pelos quais tais noções são integradas a diferentes ordens sociais (Luhmann, [1984] 2016), intervindo, muitas vezes de maneira irreversível, no modo como os *futuros presentes*, as visões no presente sobre futuros potenciais, tornam-se *presentes futuros*, a sequência de ações práticas no tempo resultantes dos primeiros (Luhmann, [1984] 2016; Simon, Tamm, 2021). Talvez pudéssemos então dizer que *o futuro (em si) não existe, mas faz existir*.

Ele também possui uma história própria. Essa *história do futuro* tem sido explorada nos últimos anos por meio de conceitos, noções e categorias, provenientes sobretudo da teoria da história, da historiografia e da história das ideias, como *regimes de historicidade* (Hartog, 2021), *antropologia das experiências históricas do tempo* (Koselleck, 2006), *futuros históricos* (Simon, Tamm, op. cit.) e *história da ideia de futuro* (Hölscher, 2014; Minois, op. cit.), entre outros. Apesar de que cada uma possui nuances específicas, de modo geral, tais concepções apontam para “a pluralidade das relações transicionais entre apreensões do passado e antecipações do futuro”²²⁰ (Simon, Tamm, op. cit. p. 5) ou, nos termos de Koselleck (2006), para a variação das relações entre os *espaços de experiência* (a presença do passado) e os *horizontes de expectativa* (a presença do futuro). Reconhecendo uma heterogeneidade de *ordens do tempo* ou de *regimes de historicidade* (Hartog, 2021), segundo tais perspectivas, o tempo histórico é constituído pela tensão variável entre esses dois pólos (Koselleck, op. cit.). Hölscher pontua que as ideias sobre o futuro são sempre

híbridos de realidade e ficção. Não podem ser considerados meras invenções nem realidades históricas em sentido simples. Por um lado, são ambas as coisas ao mesmo tempo, ou seja, são fatos mentais, e enquanto tais devem ser considerados, do ponto de vista historiográfico, com tanta seriedade quanto outros fatos históricos²²¹. (Hölscher 2014, p. 229).

²²⁰ No original: “the plurality of transitional relations between apprehensions of the past and anticipated futures”.

²²¹ No original: “híbridos de realidad y ficción. No pueden considerarse meras invenciones ni realidades históricas en sentido simple. Por una parte son ambas cosas a la vez, es decir, son hechos mentales, y en cuanto tales hay que tomarlas, desde el punto de vista historiográfico, con tanta seriedad como se toman otros hechos históricos”.

Mas, ainda que seja possível identificar diversas *modalidades de futuros* ao longo da história (Simon, Tamm, op. cit.), o futuro tal qual o concebemos hoje, ou seja, como fruto de um tempo que passa e abole irreversivelmente o passado atrás de si, é uma concepção moderna por excelência (Latour, 2019). O lugar privilegiado do futuro como o ‘grande condutor da flecha-do-tempo’ na Modernidade já foi largamente debatido e enfatizado, assim como a crise deste projeto nas últimas décadas (Cf. Berardi, 2019; Fisher, 2019, 2022; Novaes, 2015; Hartog, op. cit.; Koselleck, op. cit.), conforme veremos ao longo deste capítulo.

Enquanto o século XX, o “século que acreditou no futuro” (Berardi, op. cit., p. 12), foi o auge do *futurismo*, a experiência contemporânea do tempo seria marcada por uma ampla e profunda *crise do futuro*. Ainda que haja distintos modos de caracterizar tal crise, já não parece haver polêmica no diagnóstico de que “o futuro mudou de sinal” (ibid., p. 135) e que as visões utópicas e otimistas sobre o porvir que marcaram a modernidade foram gradualmente substituídas por visões pessimistas e catastrofistas. Face sobretudo às ruínas do comum e do Estado de bem-estar social causadas por décadas de *hegemonia neoliberal* (Berardi, op. cit.; Fisher, op. cit.; Brown, 2015, 2019; Lazzarato, 2017) e ao *tempo catastrófico do Antropoceno* (Stengers, 2015) — que, por si só, inaugura “um novo tempo do tempo”²²² (Danowski e Viveiros de Castro, 2017, p. 111) —, vivemos contemporaneamente não só um drástico encolhimento do horizonte de expectativa (Koselleck, 2006), mas uma “alteração radical nas condições de temporalização coletiva” (Turin, 2023, p. 704).

Como diagnosticou Flusser, ainda nos anos 60, “[n]ão é preciso de ouvido atento para descobrir que os passos pelos quais avançamos rumo ao futuro soam ocos. Mas é preciso concentrar o ouvido caso queiramos descobrir que tipo de vacuidade ressoa no nosso progresso” (Flusser, 2019, p. 9). A atenção às nuances da crise que permeia nossa atual relação com o futuro nos leva a identificar uma tensão ou (aparente) contradição entre pelo menos duas *modalidades de futuro*: enquanto no domínio social coletivo os imaginários, afetos e discursos sobre o futuro são marcados por um crescente incerteza, instabilidade e pessimismo, no domínio da tecnociência persistem as promessas de previsão, certeza, controle, calculabilidade, aceleração e progresso.

Seja por meio de um neocolonialismo que expande suas fronteiras do domínio molecular

²²² Amplamente popularizado a partir da primeira década deste século, o termo *Antropoceno* foi proposto pelo biólogo Eugene Stoermer e pelo químico Paul Crutzen para nomear a época geológica caracterizada pelo impacto humano na Terra. Ainda que objeto de dissenso, o termo aponta para aquilo que Danowski e Viveiros de Castro (2023, p. 7) chamam de “fato cosmopolítico total”, designando, simultaneamente, “uma catástrofe ecológica, uma tragédia econômica, uma ameaça política, uma convulsão religiosa” (ibid.).

(Mendes, 2017) (p. e. com a bio e nanotecnologia e a engenharia genética) ao cósmico (p. e. com os projetos de ‘terraformação’ de novos planetas para nossa ‘dócil’ espécie ou de extração de minérios da lua); dos projetos transhumanistas de superação das capacidades cognitivas e biológicas humanas por máquinas superinteligentes e pelos avanços na medicina; até uma geoengenharia do sistema terrestre para adaptá-lo ao aquecimento global, nutre-se a crença no avanço tecnológico como solução para todos os males contemporâneos e via de acesso a outro estágio de desenvolvimento.

Realização daquilo que Heidegger denominou de “era do domínio planetário da técnica” (Heidegger, 1979), a *tecnociência* parece ter se transformado na última herdeira da fé no progresso inexorável e ilimitado que nutriu a modernidade. Mas além da herança moderna de um futuro teleológico, outras temporalidades sobrevivem e compõem a cronopolítica da tecnociência contemporânea. Segundo a hipótese de Haraway (2018), que exploraremos ao final deste capítulo, as práticas e discursos da tecnociência são indissociáveis das *histórias seculares de salvação e da ameaça* próprias das figuras do realismo cristão. Como testemunhamos, nenhum evento tecnocientífico recente atualizou com tamanha força esse imaginário e temporalidade quanto o lançamento do ChatGPT e a posterior proliferação das IAs Generativas nos mais diversos campos.

Mas ainda que composto por heterocronias e ambivalências como as que descrevemos acima e às quais retornaremos nos próximos tópicos, podemos afirmar que na tecnociência contemporânea e, especialmente no campo da Inteligência Artificial, *o futuro segue muito vivo*. Propomos designar essa modalidade de futuro histórico (Simon, Tamm, op. cit.) produzido por meio das técnicas, práticas, discursos, imaginários e ideologias sobre a Inteligência Artificial de *futuro maquínico*. Essa modalidade de futuro, que se aproxima da noção de *tecnofuturos* proposta por Hong (2022), integra um projeto no qual a tecnologia, e principalmente a IA, ocupa o lugar central nas práticas e imaginários produtores de futuro no presente, seja em seu potencial de salvação ou destruição. Defendemos que os futuros maquínicos previstos e produzidos pela Inteligência Artificial estão *programando nossos futuros*, material e simbolicamente, seja objetivamente por meio das previsões algorítmicas, que exploramos ao longo de toda esta investigação, ou por meio dos discursos, promessas, imaginários e ideologias que permeiam o campo da IA e também produzem os efeitos destas tecnologias.

O gesto reflexivo deste capítulo permitirá compreender os algoritmos da IA como tecnologias

de governo do tempo não somente enquanto moduladoras privilegiadas de nossas possibilidades individuais e coletivas, mas como *agentes temporalizadores* (Bonaldo, 2023) que participam ativamente da produção de nossos futuros históricos no presente (Simon, Tamm, op. cit.), ao mesmo tempo em que são inegáveis produtos destas condições históricas. Afinal, como especula Lapoujade (2015, p. 289), “talvez a ideia de futuro não seja outra coisa senão um *diagnóstico do presente*”.

Enquanto nos capítulos anteriores nos centramos sobretudo no *possível* para descrever o alvo principal do modelo de governo e racionalidade algorítmicos, neste capítulo refletiremos principalmente sobre o futuro enquanto um *artefato histórico e cultural*, ainda que em alguns pontos essas duas dimensões se cruzem nas páginas que seguem. Utilizando uma metáfora maquínica, poderíamos dizer que até aqui realizamos uma espécie de ‘*zoom in*’ nas máquinas algorítmicas, explorando seu regime de saber e de temporalidade. Já neste capítulo faremos uma espécie de ‘*zoom out*’, buscando situar os problemas mobilizados por nossos objetos a partir de uma perspectiva histórica sobre a atualidade, esse “agora no interior do qual estamos” (Foucault, 2000b, p. 335).

Veremos que os futuros maquínicos agenciados pela IA são marcados por uma série de *tensões e ambivalências*: entre a crise do futuro na esfera do comum e a ascensão dos futuros maquínicos na esfera da tecnociência; entre as *promessas de ruptura* promovidas pelos imaginários e discursos sobre a IA e as *atualizações repetidoras* que essas tecnologias materializam; entre as “mitologias do fim” (Danowski, Viveiros de Castro, 2017) fabuladas pelos futuros maquínicos e pelos futuros antropocênicos. São essas tensões e ambivalências que nos conduzirão pelos os tópicos que se seguem, tal qual um campo de forças que ajuda a dar contorno aos problemas e questões que mais nos interessam.

4.1 História(s) do futuro: do futuro utópico ao futuro maquínico

O futuro não é mais o que era.

Paul Valéry

A célebre frase do poeta Paul Valéry proferida em 1920 sob a sombra da Primeira Guerra, antecipa a espécie de ‘nostalgia do futuro’ que surgiria ao fim do século XX e aos poucos saturaria os poros do século XXI. Mas antes de falar do *fim do futuro* ou, mais precisamente,

do “fim das ideias de futuro tal como o pensamento moderno soube construir” (Novaes, 2015, p. 10)²²³, falemos de seu começo.

Diversos autores têm se dedicado à tarefa de pensar uma ‘história do futuro’, especialmente a ascensão e declínio do conceito moderno de futuro (Cf. Koselleck, op. cit., Hartog, op. cit., Hölscher, 2014, Berardi, 2019). Baseando-nos principalmente nas leituras (inegavelmente eurocêntricas) de Koselleck (2006) e Hartog (2021) podemos delinear, de modo bastante esquemático, três experiências históricas do futuro: uma *pré-moderna*, uma *moderna* e uma *contemporânea*.

O primeiro período, anterior a 1789, é dominado por entendimentos cíclicos do tempo, sintomáticos de um mundo majoritariamente camponês e artesão no qual as pessoas viviam em consonância com os ciclos da natureza e “as expectativas eram sustentadas pelas experiências dos antepassados” (Koselleck, 2006, p. 315). A relação com o porvir ou o horizonte de expectativa desse período é também fundamentalmente marcado pela doutrina cristã dos últimos fins e suas profecias apocalípticas de fim do mundo. Contudo, pontua Koselleck, mesmo sob o domínio dos signos escatológicos, nesse período “o futuro permanecia atrelado ao passado” (ibid.), uma vez que as expectativas sobre o dia do Juízo Final não se referiam a este mundo, mas estavam voltadas para o assim chamado ‘além’. Deste modo, a não realização dessas visões apocalípticas não eram desfeitas por nenhuma experiência contrária, possibilitando que uma profecia não realizada se reiterasse indefinidamente. Dominado por compreensões circulares e teológicas do tempo, nesse regime de historicidade, “o futuro é o passado e pré-determinado”²²⁴ (Hong, Szpunar, 2019, p. 316).

Um trecho de Lucilio Vanini de 1616 que Borges ([1936], 2010) menciona em seu instigante *História da eternidade* para ilustrar as concepções filosóficas sobre o *tempo circular* — que tanto lhe fascinavam e das quais Nietzsche, para ele, seria apenas seu mais recente inventor já que suas origens remeteriam ainda ao Timeu de Platão — sintetiza bem a relação com o futuro e com o tempo que, pese as notáveis diferenças entre a Antiguidade e a Idade Média, precedeu a Modernidade: “Novamente Aquiles irá a Troia; renasceram as cerimônias e religiões; *a história humana se repete; nada há, hoje, que não tenha sido; o que foi, será*” (ibid., p. 77, grifo nosso).

²²³ Para um aprofundamento da questão ver a edição especial da Revista Mutações organizada por Novaes (2015) que leva o título justamente da frase de Valéry.

²²⁴ No original: “the future is past and preordained”.

Na leitura de Koselleck (2006), essa relação cíclica e repetitiva com o tempo só foi rompida com a Revolução Francesa em 1789, marcando o surgimento de um horizonte de expectativa propriamente moderno no qual a experiência histórica passa a ser regida pelo futuro. Antes significando rotação e retorno, *revolução* passou a denotar ponto de partida, convulsão e mudança (Hong, Szpunar, op. cit.) fazendo com que um horizonte outrora pré-determinado se tornasse aberto à ação humana. Para Latour (op. cit.), a temporalidade da revolução na Constituição moderna, que em sua hipótese originou-se na ideia de revolução científica antes que política, está ligada à supressão das causas e objetos da Natureza e à separação entre a história das ciências e técnicas e a “história pura e simples” (p. 88), fazendo surgir um tempo que acontece como “a súbita emergência de um milagre” (ibid.). Nessa temporalidade, “[o] presente é traçado por uma série de rupturas radicais, as revoluções, que formam dispositivos irreversíveis para impedir-nos, para sempre, de voltar atrás” (ibid., p. 90). A temporalidade inaugurada pela revolução, correlata de uma nova semântica histórica, será experienciada e descrita sobretudo através do conceito de *progresso*, a “tempestade” que impele o anjo benjaminiano da história “*irreversivelmente para o futuro*” (Benjamin, 2012, p. 246, grifo nosso)²²⁵.

A partir dessa ruptura, a história, antes descrita pelo termo alemão *Historie*, no plural, designando narrativas particulares, transforma-se em *Geschichte*, coletivo singular, designando “tanto a sequência unificada dos eventos que constituem a marcha da humanidade, como o seu relato” (Koselleck, op. cit., p. 11). Assim, a nova História passa a ser entendida como “um processo de contínuo e crescente aperfeiçoamento” (ibid., p. 317). Como observa Berardi, “apenas quando, a partir do humanismo, o tempo histórico se tornou acumulação de saber e de potência, o futuro apareceu como progresso, como aperfeiçoamento do conhecimento da esfera técnica e da potência humana” (Berardi, 2019, p. 22). Ao longo da modernidade, segundo a fórmula koselleckiana, as expectativas sobre o futuro se distanciaram cada vez mais das experiências passadas, possibilitando o surgimento de um *futuro aberto*. Há uma relação direta entre a possibilidade de um futuro aberto (ou indeterminado) e o lugar que a técnica e a tecnologia ocupam no projeto moderno. Neste, a

²²⁵ Em seu clássico ensaio *Sobre o conceito de história*, Benjamin (2012), afirma que o conceito de progresso que fundamentou a teoria e prática da social-democracia moderna, era, em primeiro lugar, “um progresso da humanidade em si (e não apenas das suas capacidades e conhecimentos)” (p. 249); em segundo lugar, um “processo sem limites (correspondente a uma perfectibilidade infinita da humanidade)” (ibid.); e, em terceiro lugar, um “processo essencialmente inexorável (percorrendo autonomamente uma trajetória em flecha ou em espiral)” (ibid.).

tecnologia é a ferramenta privilegiada de ‘abertura do futuro’, ao ponto de progressivamente se tornar, inclusive, um de seus sinônimos.

Onde o presente estava outrora reservado à repetição e ao regresso de possibilidades sempre circunscritas pelo fim prometido, a modernidade seria definida pela *indeterminação fundamental do futuro* – com o avanço tecnológico sendo uma das manifestações materiais mais claras dessa possibilidade.²²⁶ (Hong, 2022, p. 372)

Época da consciência sobre sua própria atualidade enquanto acontecimento histórico (Foucault, 2000b),

[a] modernidade forma com a projeção progressiva do futuro uma unidade indivisível. Modernos são aqueles que vivem o tempo como esfera do progresso rumo à perfeição ou, pelo menos, a uma condição cada vez melhor, mais feliz, mais rica, mais plena, mais justa. (Berardi, 2019, p. 18).

A filosofia da história hegeliana, síntese do espírito histórico da modernidade do século XIX, aperfeiçoou a visão do futuro teleológico afirmando que a história universal é um processo dialético que progride através de conflitos e da resolução de contradições e está destinado a culminar no ‘fim da história’, domínio final da Razão sobre o mundo e autoconsciência da liberdade (Berardi, op. cit.).

A transformação das estruturas temporais na modernidade também temporalizou a noção de *utopia* (Koselleck, 2014, Hong, Szpunar, 2019). Antes articulada sobretudo em termos espaciais, como no inaugural ‘não lugar’ de Thomas More, a utopia é inserida num futuro cronológico, transformando-se, em aliança com a revolução e com a vanguarda, nas bases de um projeto social que conduziu o horizonte de expectativa moderno, ao mesmo tempo em que também ressoa uma espécie de herança do ‘fim prometido’ do regime de historicidade anterior.

Contudo, sabe-se bem que o *homo modernus* (Ferreira da Silva, 2022) em torno do qual esse projeto (supostamente) emancipatório foi construído pressupunha uma série de grandes divisores (Cf. Latour, 2019) e dispositivos classificatórios. Entre estes, a separação entre Natureza e Cultura foi o mais fundamental e a ‘ficção útil’ da raça na figura do negro o mais violento e eficaz (Cf. Ferreira da Silva, op. cit.; Mbembe, 2018). Entre tantas operações, tais dispositivos separaram — mais precisamente, inventaram — os povos que habitariam o

²²⁶ No original: “Where the present was once reserved for the repetition and return of possibilities always circumscribed by the promised end, modernity was to be defined by the fundamental indeterminacy of the future — with technological advancement being one of the clearest material manifestations of that possibility”.

*cronotopo*²²⁷ do progresso e os “povos sem história” (Latour, op. cit., p. 95), para sempre relegados à repetição de um passado mítico e à ‘infância do mundo’. Segundo Latour,

[a] *assimetria entre natureza e cultura torna-se então uma assimetria entre passado e futuro*. O passado era a confusão entre as coisas e os homens; o futuro, aquilo que não os confundirá mais. A modernização consiste em sair sempre de uma idade das trevas que misturava as necessidades da sociedade com a verdade científica para entrar em uma nova idade que irá, finalmente, distinguir de forma clara entre aquilo que pertence à natureza intemporal e aquilo que vem dos humanos” (Latour, 2019, p. 90, grifo do autor)

Ou seja, dentre tantas segregações e hierarquizações impostas pelo projeto moderno colonizador de conhecimento e governo está uma de ordem cronopolítica que buscou domesticar a “alteridade temporal” (Hartog apud Dornelles, Bonaldo, 2022, p. 27) dos povos extramodernos para fazer-lhes, simultaneamente, ingressar “no tempo europeu de *chronos*” (ibid.) e ser dele excluídos; os mesmos povos, cabe observar, para quem “a barbárie que chega” (Ferdinand, 2022, p. 269) do Antropoceno e os temores sobre o fim que ele provoca nada tem de novo...

A mutação de longa duração iniciada no fim do século XVIII, ao longo da qual o futuro passa a conduzir a *flecha irreversível do tempo*, atinge seu auge — assim como seu declínio — no século XX, o século que “acreditou no futuro” (Berardi, 2019, p. 12). Acreditava-se numa nova sociedade, num novo homem e numa nova arte frutos da confluência da razão, da ciência, da técnica e da política — desde que masculinas, brancas, eurocêtricas, é claro. Linha de chegada e realização das promessas da modernidade, para o filósofo, o século XX começa, de fato, “quando os futuristas bradam com arrogância o reino da máquina, da velocidade e da guerra” (ibid., p. 13). Primeira vanguarda artística moderna e também aquela que expressou sua utopia com maior violência e extremismo, o futurismo, seja na vertente italiana ou russa, é “a alma estética de uma *fé no futuro* que permeia profundamente o espírito do capitalismo moderno” (ibid., p. 15).

A exaltação futurista e falocêntrica da máquina e da velocidade materializada no Manifesto de Marinetti de 1909 marcou a sensibilidade e a expectativa da cultura do século XX, sintonizada com a aceleração do processo de urbanização e industrialização que transformava as cidades naquele momento. Segundo Berardi, a máquina futurista é, essencialmente, uma *Máquina Externa* em relação ao corpo e mente humana: “máquina pesada, ferruginosa e

²²⁷ Proposto originalmente por Bakhtin na teoria literária para demonstrar como diferentes gêneros literários operavam com diferentes configurações de tempo e espaço, o termo *cronotopo* significa, literalmente, um *tempo-espaço* ou “um *topo* através do qual a temporalidade é organizada” (Haraway, 2018, p. 41, tradução nossa).

volumosa”; “máquina visível no espaço urbano e no espaço da fábrica e da rua” (ibid., p. 15-16); “máquina mecânica de metal luzente” (ibid., p. 99). A Máquina Externa ocupa o centro do imaginário futurista principalmente na figura do automóvel, “*Deus veemente de uma raça de aço*”, “*mais bonito que a Vitória de Samotrácia*”, nas palavras de Marinetti.

Em suas reflexões sobre as relações entre tipos de sociedades e de máquinas, Berardi contrapõe a Máquina Externa do futurismo à *Máquina Interna* da época digital, a *infomáquina* que opera por meio de automatismos lógicos e cognitivos. “No decorrer de um século”, afirma, “a máquina externa do futurismo (...) se transformou em uma máquina invisível, interiorizou-se, tornou-se infomáquina, biomáquina, nanomáquina, potência penetrante que se autogera e fagocita o tempo.” (Berardi, 2019, p. 84). Nesse contexto, a máquina se converte em “diferença de informação, não exterioridade, mas sim modelação linguística, automatismo lógico e cognitivo” (ibid., p. 17), intimamente relacionada à financeirização da economia e à invasão matemática na linguagem e em todos os aspectos da vida sobre as quais já comentamos.

Berardi encontra na passagem da Máquina Externa para a Interna um reflexo da transição do regime disciplinar foucaultiano para o regime do controle deleuziano. No primeiro, a máquina se constituiu diante do corpo humano, que, ainda em estágio pré-técnico, “devia ser regulado normativa, legal e institucionalmente para ser submetido ao ritmo das máquinas concatenadas” (ibid., p. 17). No segundo, a máquina digital teria adentrado e transformado não somente a infraestrutura social e econômica, mas se interiorizado no corpo-mente dos sujeitos produzindo uma profusão de automatismos psicocognitivos. Nesse novo contexto, o processo de disciplinamento político, legislativo e repressivo perderia relevância, pois o controle passa a ser exercido integralmente pela própria máquina interna.

Apesar de Berardi não enfatizar tanto o termo ao longo do livro, poderíamos afirmar que *o algoritmo é a máquina interna por excelência*. Ele é interno não somente em relação ao “sistema nervoso social”, nos termos do autor, mas à própria materialidade da máquina, ao seu *hardware*. A transição da Máquina Externa à Máquina Interna proposta por Berardi também reflete uma mudança crucial na função social da máquina em cada momento histórico. Ainda que, em termos marxistas, ambas estão diretamente ligadas à ampliação da mais-valia, a Máquina Externa (máquina industrial, movida à energia) ocupava um lugar privilegiado no projeto social e estético moderno — tanto que integrava a máxima revolucionária da tomada dos meios de produção pela classe operária. Já a Máquina Interna

(máquina pós-industrial, movida à informação) não parece integrar mais nenhum projeto social coletivo emancipador, estando restrita, *grosso modo*, à ampliação da mais-valia num capitalismo de dados plataformizado, sobretudo depois das frustrações com a utopia digital sobre a qual discorreremos adiante. O predomínio da Máquina Interna é também sincrônico a dissolução das premissas filosóficas, sociais e estéticas da Modernidade ocorrida nas últimas décadas do século XX. A Máquina Interna é, portanto, própria de um tempo situado *depois do futuro* moderno (Berardi, op. cit.) e guarda íntima relação com o que estamos chamando aqui de futuros maquínicos.

Terceiro e último momento desta breve história do futuro que estamos traçando, no período que vai de meados do século XX até a contemporaneidade a organização da experiência histórica deixa de ser regida pelo futuro. Ao longo do século passado, diversos eventos históricos irão paulatinamente ruindo a fé num futuro aberto, teleológico, progressivo e utópico. Dos abalos causados pela Primeira Guerra Mundial, passando pelas cinzas genocidas e atômicas deixadas pela Segunda Guerra, grande divisor do *zeitgeist* do século XX, às frustrações utópicas no pós 68, culminando na queda do muro de Berlim em 1989, que motivou a alardeada tese de Fukuyama ([1989], 2015) sobre o “fim da história” e que, para Hartog, marca o fim do regime de historicidade moderno, ocorre um gradual *fechamento do futuro*.

Hölscher (2014), ao relatar a ascensão e declínio da noção moderna de futuro, pontua como a própria necessidade de reconstrução das estruturas econômicas e sociais no pós-guerra levou a um encolhimento do horizonte de expectativas. Naquele momento, as sociedades ocidentais, sobretudo as europeias, “já não estavam dispostas a sacrificar sua qualidade de vida a curto e médio prazo em prol das vantagens de uma sociedade ideal”²²⁸ (ibid., p. 215). As imagens do ‘novo homem’ e da ‘nova sociedade’ próprias das utopias revolucionárias da primeira metade do século foram substituídas por “ideais de crescimento evolucionistas, como o produto interno bruto, a modernização técnica e a industrialização”²²⁹ (ibid.). Apesar de que ainda se visasse mudanças a longo prazo, sua realização já não era visada sob a forma de uma sociedade qualitativamente diferente, mas dentro das formas estabelecidas. Assim, “as anteriores alternativas antagônicas das sociedades futuras foram se desfazendo no ideal de

²²⁸ No original: “no estaban ya dispuestas a sacrificar su calidad de vida a corto y medio plazo en pro de las ventajas de una sociedad ideal”.

²²⁹ No original: “ideales de crecimiento evolucionistas, como el producto nacional bruto, la modernización técnica y la industrialización”.

uma sociedade de bem-estar de classe média, definida essencialmente pela economia”²³⁰ (ibid., p. 216). A realização desse novo modelo social foi viabilizada pela ampla implementação de políticas socioeconômicas regidas por uma racionalidade neoliberal e na virtualização dos mercados financeiros a partir dos anos 1970. Como anunciavam os punks ingleses dos Sex Pistols em 1977, numa espécie de prenúncio não só dos sombrios anos ultraliberais de Thatcher, mas de uma ampla dissolução da ideia moderna de futuro: *there's no future*.

Na década de 1980 já era perfeitamente clara a presença de um futuro distópico, refletida nas visões apocalípticas em torno de um possível auto-aniquilamento nuclear da espécie humana (Anders, 2007), no despertar da consciência de que o progresso dos séculos anteriores havia produzido “uma destruição implacável das bases ecológicas da vida no planeta”²³¹ (Hölscher, op. cit., p. 219) ou mesmo no surgimento do *ciberpunk* como subgênero da literatura de ficção científica, laboratório central da transição de imaginários utópicos para distópicos em relação à tecnologia ao final do século XX. Além de sintomático da passagem da prevalência da máquina externa para a máquina interna (sobretudo dos temores desencadeados pela segunda) (Berardi, 2019), apreensível em livros como no clássico *Neuromancer* de Gibson ([1984] 2016), o ciberpunk, em larga medida, antecipou as frustrações com aquela que para Berardi foi a última utopia do século que acreditou no futuro: a *utopia virtual*.

Quem vivenciou ‘os primórdios da web’ sabe como o otimismo com as potencialidades revolucionárias, democráticas, libertárias e descentralizadas da Internet exaltadas nos anos 1980 e 1990 — sintetizadas em noções como a “aldeia global” de McLuhan (1974), a “inteligência coletiva” de Levy (1998) ou na Declaração de Independência do Ciberespaço de Barlow (1996), que propunha abertamente o espaço virtual como uma realidade com seu próprio contrato social — foi sendo progressivamente frustrado com o processo de plataformação da web (Helmond, 2015; Srnicek, 2017), o poder cada vez mais monopolista das Big Techs e o investimento sobre modelos de negócios baseados em publicidade (Zuboff, 2021; Bentes, 2022) e em arquiteturas algorítmicas que produziram efeitos polarizadores (Chun, 2021; Cesarino, 2022). Sobretudo após a eleição de Trump em 2016 e a ascensão global da extrema direita sob a forma de um populismo digital que instrumentalizou o viés

²³⁰ No original: “las anteriores alternativas antagónicas de las sociedades futuras se fueron diluyendo en el ideal de una sociedad del bienestar de clase media, definida en lo esencial por la economía.”

²³¹ No original: “una incontenible destrucción de las bases ecológicas de la vida en el planeta”.

antiestructural das novas mídias (Cesarino, op. cit.), a sensação era, para usar a expressão de um meme popular na internet brasileira, de que “aquela festa havia virado um enterro”.

No plano securitário, as mutações desencadeadas pelo 11/9, como vimos no Capítulo 3, forjaram uma nova relação com o futuro baseada nos temores e ameaças diante do desconhecido e do imprevisível, respondidas através do gerenciamento antecipado de riscos. Já no plano ecológico, a crescente percepção ao longo das últimas duas décadas (e de modo radical nos últimos anos) de que o novo regime climático (Latour, 2020) não é mais um evento longínquo e difuso em nosso horizonte, mas uma das principais forças que dão forma a nosso denso presente (Haraway, 2016), instaura a dúvida não somente sobre qual será a natureza do futuro, mas, mais elementarmente, *se haverá futuro* (Danowski, Viveiros de Castro, 2017) ou se será possível *adiar o fim do mundo* (Krenak, 2019). Essa (obviamente nada elementar) dúvida, possivelmente a grande dúvida de nossa geração, se tornou ainda mais radical com a pandemia de COVID-19, deflagrada em 2020, especialmente para aqueles que, como os brasileiros, viveram este acontecimento traumático sob um (des)governo negacionista. “Espécie de epítome destes ‘tempos falhados’”, a pandemia, numa aparente contradição, como afirmam Bruno, Pereira e Faltay (no prelo), pôs tudo em suspenso na estranha experiência de um tempo-fora-do-tempo e acelerou uma série de processos em curso, evidenciando que “o sentimento de colapso generalizado que marca nosso presente histórico não é fruto de eventos isolados, mas da sobreposição de múltiplos territórios em pane” (ibid.).

Esse acúmulo de falhas que viemos descrevendo terminaram por produzir um sentimento coletivo pessimista, melancólico e distópico sobre o futuro, imaginado e experienciado agora principalmente sob os signos da *catástrofe* (Stengers, 2015; Hartog, 2021, Turin, 2023), da *crise permanente* (Chun, 2016; Cesarino, 2022) e da *urgência* (Haraway, 2016; Simon, Tamm, 2021; Mbembe, 2021), como atestou a inédita tragédia climática no Rio Grande do Sul ocorrida em maio de 2024. O simultâneo encolhimento do horizonte de expectativa e a *aceleração generalizada das estruturas temporais*, sobretudo por meio da aceleração climática e tecnológica, a dupla aceleração de nossa época (Viveiros de Castro, 2023) sobre a qual nos dedicaremos nos últimos dois tópicos deste capítulo, modificam radicalmente o modo como conceitualizamos, imaginamos e sentimos o tempo. A velocidade, complexidade e escala sistêmica dos fenômenos em curso (um curso que parece cada vez mais desenfreado e caótico) dificultam a percepção de seus contornos, como se estivéssemos “dentro do nevoeiro” (Wisnik, 2015) ou “da tempestade” (Ferdinand, 2022) — e de fato estamos —

intensificando um sentimento coletivo de desorientação e ansiedade²³². Como pontua Turin, habitamos “uma nova brecha do tempo: reconhecemos o afastamento de algo que, agora, vem se tornando passado, mas ainda nos mostramos incapazes de visualizar plenamente novas ordens de realidade” (Turin, 2023, p. 706).

Essa mudança aguda de percepção coletiva sobre o presente e o porvir vem sendo elaborada sob distintas lentes teóricas nos últimos anos. Numa proposição do começo dos anos 2000 que obteve ampla repercussão, Hartog descreveu a crise do futuro que marca o regime de historicidade contemporâneo nos termos de um *presentismo* (Hartog, 2021). Antítese ao *futurismo* que marcou a modernidade, o regime presentista denota um “presente onipresente, onipotente, que se impõe como único horizonte possível e valoriza só o imediatismo” (Hartog, op. cit., p. 15). Se o regime de historicidade moderno foi caracterizado por um distanciamento acelerado entre espaços de experiência e horizontes de expectativa, “o presentismo significaria um limite crítico, quase uma ruptura entre elas” (Dornelles, Bonaldo, 2022, p. 25). Com uma economia midiática ancorada sobre a incessante produção de acontecimentos, esse presente hipertrofiado — cuja descrição de Hartog, originalmente de 2003, é ainda quase totalmente alheia ao processo da digitalização — seria marcado, paradoxalmente, por um desejo ansioso manifesto, por um lado, por ver-se como já passado e, por outro, por prever e antecipar continuamente o futuro.

Nos estudos de risco (Cf. Beck, 2011) também encontramos argumentos que ressoam a percepção contemporânea do futuro enquanto ameaça e que termina por produzir efeitos presentistas. Da ameaça terrorista às ecológicas e epidemiológicas, segundo tais perspectivas, estamos expostos a perigos mais difusos e incertos que nunca (Hong, Szpunar, op. cit.). Como vimos no capítulo 3, as chamadas sociedades de risco (Beck, op. cit.) possuem sempre uma tensão subjacente entre os temores diante de um futuro como horizonte de desconhecidos (Amoore, 2013) e a tentativa, no limite sempre falha, de controlar antecipadamente essa incerteza e imprevisibilidade, seja por meio de práticas especulativas, preventivas ou preditivas (com frequência sobrepostas), entre as quais a previsão algorítmica ocupa atualmente lugar privilegiado. Visando gerir antecipadamente esses perigos potenciais, as tecnologias do risco (Amoore, op, cit.) também contribuem para a experiência de um

²³² Talvez a ansiedade seja ‘o grande afeto’ do nosso tempo. É sintomático que ela se manifeste especialmente nas duas esferas, a *tecnológica* e a *climática*. Nesse sentido, expressões como “ansiedade algorítmica” e “ansiedade climática” tornam-se cada vez mais recorrentes para dar contorno a essas novas angústias e temores. É interessante observar também como a ansiedade possui inerentemente um caráter temporal uma vez que se baseia numa permanente antecipação do futuro.

presente alargado (Hong, Szpunar, op. cit.).

Outra proposição que se tornou bastante popular para descrever essa mutação histórica e cultural foi a de “um lento cancelamento do futuro” de Mark Fisher (2022, p. 24), sobre a qual também já comentamos ao abordar a matriz financeira-neoliberal da cronopolítica algorítmica (ver item 3.2.3). Para o autor, o começo do século XXI foi marcado por uma sensação de finitude e exaustão, uma “deflação das expectativas” (ibid., p. 26). Centrando-se principalmente numa análise da produção cultural do período, o autor identifica, alinhado aos diagnósticos de Jameson sobre o pós-modernismo a partir dos anos 1980 (Jameson, 2002), uma insistente tendência à retrospectção, ao pastiche, à nostalgia e aos anacronismos. “O tempo cultural se dobrou sobre si mesmo, e as *impressões lineares de desenvolvimento deram lugar a estranhas simultaneidades*”, afirma (Fisher, op. cit., p. 27)²³³. A sensação de esvaziamento do futuro e de repetição do passado sob a forma da nostalgia e da retromania estão intimamente relacionadas nas teses de Fisher com o que viria a designar *realismo capitalista*, a hegemonia do capitalismo neoliberal sobre o inconsciente econômico e político e o mal-estar diante de um cenário de fechamento sistêmico (Fisher, 2019). Em suma, a sensação de que não há (mais) alternativas ao capitalismo. A implosão de imaginários sobre horizontes alternativos após o socialismo real entrar em colapso e a ontologia dos negócios se infiltrar em todas as instâncias da vida fez com que, aos poucos, culminando na crise de 2008, o futuro se tornasse “inteiramente encerrado no interior dos limites do capitalismo, que captura todas as suas possibilidades para estender-se, propagar-se. Não há futuro exceto no capitalismo. E todo futuro nunca é senão o futuro do capitalismo.” (Lapoujade, 2015, p. 292). Nesse caldo de falência social, a relação entre financeirização e dívida para os aparelhos de captura do neoliberalismo contemporâneo tem papel fundamental.

De acordo com Cesarino (2022), o fechamento do sistema capitalista neoliberal sobre si mesmo descrito por Fisher sob a rubrica do *realismo capitalista* o conduz a um estado instável de ritmo involutivo e oscilatório marcado por uma *cronotopologia de crise permanente* (Chun, 2016): “uma ferradura entre um *presentismo exacerbado* e um horizonte

²³³ O fenômeno cultural do Retrofuturismo, corrente estética que recupera os ‘futuros do passado’, dos cinemas *drive-in* à corrida espacial, descreve bem como essa patologia temporal (Fisher, 2022) passa por também uma crise imaginativa. A comunidade no Reddit dedicada ao tema, atualmente com 935 mil membros, descreve essa fissura com precisão: “os fantásticos e delirantes sonhos do nosso passado”. A retomada desses ‘sonhos delirantes’ aponta também como a retromania é atravessada pela invenção de futuros passados que nunca existiram. Impossível não lembrar do “*make it great again*” da extrema direita e “o resgate de passados supostamente gloriosos, uma volta ao que nunca houve” (Alves, Liedke, 2022) sob a forma de revisionismos e negacionismos históricos.

de *futuro longínquo obscuro, inacessível ou apocalíptico*” (ibid., p. 273, grifo nosso). A autora observa, junto à Guyer (2007), que a temporalidade não linear é própria do neoliberalismo, produzindo uma relação com o futuro marcada, por um lado, pela impossibilidade de projeção a médio prazo e um “presente imediato que ‘patina’ sobre si mesmo” (Cesarino, op. cit., p. 74) e, por outro, por “um futuro distante e inescrutável que só pode ser imaginado na forma de visões messiânicas ou apocalípticas que sejam o inverso da realidade existente” (ibid.)²³⁴. Explorando as interseções entre neoliberalismo e digitalização, Cesarino aponta como a *temporalidade não linear e involutiva* do realismo capitalista, na qual “devemos fazer cada vez mais apenas para continuarmos onde estamos” (ibid., p. 274), é acelerada e potencializada pela infraestrutura cibernética das novas mídias, que também operam através de uma estrutura temporal não linear, contrastando com a temporalidade linear do período keynesiano-fordista. Adiante veremos como os futuros maquínicos também operam por meio de uma ferradura ou heterocronia similar à descrita por Cesarino entre um presente que ‘patina sobre si mesmo’ e um futuro distante (ora messiânico, ora apocalíptico).

Pereira e Araújo (2022), propondo um avanço na tese do presentismo de Hartog, analisam a emergência e crescente utilização da palavra em inglês “*to update*” como fenômeno revelador de deslocamentos contemporâneos no modo como o tempo é experienciado e descrito. A partir da identificação de um crescimento do campo semântico em torno da palavra *atualização* e de seus correlatos em outras línguas a partir de meados dos anos 1960 — com o significado de “melhoria de algo por meio de sua adaptação ao presente” (Pereira, Araújo, 2022b, p. 73) — quando a cultura digital começa a se expandir, os autores apontam para um “substantivo deslocamento nas formas modernas de significar o tempo histórico” (Pereira, Araújo, 2022, p. 31). Segundo o levantamento dos autores, isso ocorreu simultaneamente a um declínio do uso de termos clássico-modernos como *progresso* e *revolução*. Para tanto, eles propõem o neologismo *updatism* (ou *atualismo*, em português) como categoria analítica que visa captar as singularidades de “um tempo em que não haveria distância entre a ação e sua integração a um sistema sempre atual, sempre em estado presente” (ibid., p. 32). O fenômeno do atualismo teorizado por Pereira e Araújo pode ser pensado como “a dimensão temporal que emerge nessas sociedades aprisionadas pelas estruturas da expansão infinita” (ibid., p. 39) nas quais não só o modelo produtivo, mas os indivíduos são incitados a serem

²³⁴ O casal de antropólogos Jean e John Comaroff argumentam que existe uma relação entre as *temporalidades apocalípticas e messiânicas* emergentes no contexto do que chamam de capitalismo milenarista e as teologias econômicas da financeirização, que apresentam características oraculares, conspiratórias, piramidais e fantasmagóricas (Comaroff e Comaroff, 2000).

continuamente contemporâneos de si mesmos a fim de evitar a própria obsolescência. O atualismo produziria assim um “presente vazio e autocentrado (...) enquanto “o futuro é desejado como reserva para a expansão linear deste presente em constante *atualização repetidora* (Pereira, Araújo, 2022b, p. 73, grifo nosso). Em tese melhor que o presente, mas na prática sua (quase) repetição, o futuro torna-se então sua *cópia atualizada*.

Ainda que partindo de campos teóricos tão distintos quanto a teoria da história e teoria dos meios, são notáveis as convergências entre o diagnóstico de Pereira e Araújo sobre a experiência contemporânea do tempo e o de Chun (2016) sobre a temporalidade das redes. O conceito de *atualismo* dos primeiros e o de *novos meios habituais* da segunda coincidem no diagnóstico da *centralidade do imperativo da atualização* nos processos de temporalização e subjetivação desencadeados pelas dinâmicas digitais. Para Chun, conforme vimos, a temporalidade da atualização ininterrupta das redes é composta pela soma da *repetição do hábito* (seja maquínica ou humana) com a *crise permanente*. Assim, seja o “movimento de repetição sem a percepção de uma real transformação estrutural e positiva” de Pereira e Araújo ou a máxima de “atualizar-se para continuar (quase) o mesmo” de Chun, apontam para uma dinâmica de atualização ininterrupta sem real transformação e para a ‘sincronização’ das relações sociais e dos modos de subjetivação à temporalidade das redes e das máquinas computacionais²³⁵.

Como podemos ver, ainda que não completamente coincidentes, os diagnósticos sobre uma *crise do futuro* e a prevalência de um presente alargado, repetitivo e estagnado na contemporaneidade são abundantes. Em contraponto, como adiantamos na abertura deste capítulo, no domínio da tecnociência, especialmente no âmbito da Inteligência Artificial, a relação com o futuro permanece intimamente atrelada a uma fé obtusa no progresso técnico. Na esteira desse projeto, a aceleração tecnológica e a transição para uma civilização computacional parecem se tornar “o novo caminho para a salvação” (Mbembe, 2021, p. 23); o que também implica que se desdobre num novo caminho para a catástrofe. Mesmo após o “mal-estar na era plataformização” (Cesarino, 2021, p. 86) e uma certa saturação generalizada com o digital após a pandemia (ou, na verdade, justamente por conta deste

²³⁵ É curioso, no entanto, como ambas propostas não mencionam explicitamente a recursividade em suas reflexões, estrutura temporal central das máquinas cibernéticas e sem a qual essa atualização ininterrupta que tematizam não seria possível. Afinal, é justamente a estrutura espiralar da recursividade que produz a temporalidade de uma atualização repetidora transversal aos modos de individuação nas redes.

contexto)²³⁶, testemunhamos como os alardes sobre uma transformação acelerada de ‘*tudo em todo lugar ao mesmo tempo*’ após o lançamento do ChatGPT reativou com força o cronotopo do progresso sob a forma de um tecnosolucionismo e um tecno-otimismo tão arrogantes quanto ingênuos. Como pontua Hong, “o otimismo tecnológico contemporâneo continua a operar num eixo do progresso como destino”²³⁷ (Hong, p. 372).

Diante da falência de um projeto emancipador para o futuro pautado no comum e do crescimento dos riscos, falhas e incertezas nas últimas décadas, a tecnologia, principalmente sob a forma das tecnologias algorítmicas, parecem se tornar simultaneamente meio de manutenção da (suposta) abertura do futuro (Hong, 2022) e ferramenta privilegiada para gerir os ‘riscos extremos’ a que estamos cada vez mais vulneráveis, incluindo, ironicamente, aqueles causados pela própria tecnologia. Neste sentido, acompanhando leituras como as de Hong (op. cit.) e Simon (2021), poderíamos dizer que mais que um cancelamento ou perda do futuro, houve um deslocamento *da crença no futuro da esfera social do comum para a esfera tecnológica*, seja em seu potencial de salvação ou destruição. Na ascensão desses tecnofuturos “encontramos uma teleologia implacável que apresenta o futuro supostamente aberto do progresso tecnológico como a única opção possível; em outras palavras, somos confrontados com o fechamento de mundos e temporalidades possíveis a um único tipo de progresso”²³⁸ (Hong, op. cit., p. 374).

Dentre outras formulações, esse deslocamento poderia ser expresso como uma mutação do *futuro como projeto social e utópico* para um *futuro como programa maquínico* que tem como agente privilegiado os algoritmos da IA. Não mais um futuro próprio da revolução e da utopia, mas um futuro calculável, previsível, instrumental, guiado por uma temporalidade preditiva e aceleracionista e por um senso de inevitabilidade. Não mais um futuro que tem o humano como medida e protagonista, mas um *futuro pós-humano* que opera sob a racionalidade, cognição e temporalidade maquínicas. Não mais um futuro que deriva da autonomia do sujeito liberal moderno, mas da *crescente autonomia de agentes sintéticos*.

²³⁶ Na verdade, é justamente num contexto de pressão da indústria tecnológica e do capital de risco que financia o setor por um novo *hype* depois do declínio das criptomoedas e das frustrações com o metaverso que o lançamento do ChatGPT deve ser visto comercialmente. O *boom* pandêmico experimentado pelas principais empresas de tecnologia havia acabado, demissões em massa estavam sendo feitas e os investidores começavam a cobrar um novo *hype* para impulsionar uma nova onda de investimentos.

²³⁷ No original: “contemporary technological optimism continues to operate along an axis of progress as destiny”.

²³⁸ No original: “we find a relentless teleology that presents the putatively open future of technological progress as the only possible option; in other words, we are confronted with the closure of possible worlds and temporalities to the one and only kind of progress.”

Os futuros maquínicos agenciados pela IA são simultaneamente materiais e simbólicos, ferramentas e imaginários. Desdobram-se tanto objetivamente nas previsões algorítmicas e seus efeitos sobre o campo do possível, quanto nas ideologias e narrativas tecnofuturistas promovidas por seus produtores e financiadores, a elite dos ‘tecnoboys do Vale’. Uma de nossas hipóteses é que os futuros maquínicos produzidos pela IA produzem uma dupla relação com o futuro, bastante sintonizada com os diagnósticos sobre o neoliberalismo produzir uma relação com o futuro baseada numa ferradura temporal entre um *presentismo exacerbado* e um *futuro distante*, simultaneamente apocalíptico e messiânico (Cesarino, op. cit.; Guyer, op. cit.).

Assim, por um lado, ao nível programático, o futuro materializado nas previsões algorítmicas produzidas com aprendizado de máquina e dados massivos se torna *repetição recursiva* ou *atualização repetidora* (Pereira, Araújo, 2022; Bonaldo, 2023) baseada no reconhecimento e automatização de padrões, probabilidades, correlações e na estrutura temporal da recursividade, sobre os quais discorreremos amplamente nesta pesquisa. O pólo do futuro maquínico enquanto repetição recursiva também se conecta aos diagnósticos sobre o fechamento do futuro e o sentimento de uma inércia paradoxal que atravessam a experiência contemporânea do tempo na qual tudo muda incessantemente, mas nada muda de fato (Pereira, Araújo, 2022; Chun, 2016).

Por outro lado, principalmente ao nível dos discursos e imaginários culturais e alavancado pelos últimos desdobramentos no desenvolvimento da IA, o futuro se torna reiterada *promessa de disrupção* cujo evento prototípico é a Singularidade Tecnológica. Situado num cronotopo que oscila entre visões messiânicas e apocalípticas, esse pólo, em grande medida, reflete as próprias ambivalências e tensões entre libertação e controle, fascínio e temor, utopia e distopia que sempre marcaram a história das tecnologias digitais e das tecnologias algorítmicas em especial. Essa dimensão dos futuros maquínicos manifesta-se em figuras como as narrativas sobre os *riscos existenciais em larga escala* e as ideologias enredadas do singularitarismo, do aceleracionismo, do transhumanismo, do longoprazismo e do altruísmo eficaz (Cf. Gebru, Torres, 2024) e encontra expressão em nomes como Nick Bostrom, Sam Altman e Elon Musk e em instituições como o *Future of Life Institute*. Simultaneamente longínquo (como nas narrativas do longoprazismo) e iminente (como nos frequentes anúncios de que a Singularidade está muito próxima), o poder destes imaginários tecnofuturistas de tons ora messiânicos, ora apocalípticos, se renova continuamente a cada adiamento das disrupções que promete (Hong, op. cit.). Conforme veremos especialmente

nos dois últimos tópicos deste capítulo, defendemos que os significados e implicações dessa segunda dimensão dos futuros maquínicos só podem ser devidamente compreendidos em seu entrelaçamento imaginário, discursivo e material com a emergência climática.

4.2 O futuro como repetição recursiva e as margens de indeterminação das máquinas computacionais

Num artigo em que propõe conceber os sistemas de Inteligência Artificial como “agentes temporizadores em uma história mais que humana” (Bonaldo, 2023, p. 2), Bonaldo encontra na teoria do atualismo de Pereira e Araújo (2022), sobre a qual discorremos no tópico anterior, a conceituação para descrever a *temporalização artificial do tempo* que esses agentes operam. Conectando teoria da história e as formas de temporalização articuladas pelos modelos de aprendizado de máquinas, especialmente os de IA Generativa, o autor sugere que o futuro previsto e moldado pela inteligência artificial em sua versão hegemônica opera enquanto uma *atualização repetidora* (Pereira, Araújo, op. cit.) que com frequência “reificam padrões de historicidade ligados às desigualdades de longa duração” (ibid., p. 3).

Se as formas de temporalização articuladas pela IA dependem do reconhecimento de padrões a partir de bases de dados dentro das quais a natureza da informação histórica capturada é formalmente ordenada e discretizada, a performance das máquinas de aprendizado – ou o que poderíamos chamar de sua produção ôntico-ontológica de futuro – teria como resultado a repetição desses mesmos padrões agora apresentados como prognósticos? (Bonaldo, op. cit., p. 3)

Como observa Hong sobre a contradição constituinte dessa modalidade de futuro, “os tecnofuturos pregam mudanças revolucionárias enquanto praticam uma *política de inércia*” (Hong, 2022, p. 373, grifo nosso). As predições algorítmicas “reciclam *passados altamente seletivos* naquilo que é apresentado como o *futuro inevitável*”²³⁹ (ibid., p. 386, grifo nosso). Apesar de serem apresentadas como ferramentas de ruptura entre o presente e o futuro, estas máquinas computacionais estão, ao menos em seus usos hegemônicos, restritas a uma recombinação recursiva do passado — datificado, automatizado e assim também mais facilmente naturalizado, e esse é um ponto fundamental — em seus conjuntos de treinamento. Por isso são, segundo Oliveira (2023), “oráculos do passado”.

Poderíamos dizer que a temporalidade da atualização repetidora dos algoritmos preditivos da IA manifestam um correlato material-maquínico dos diagnósticos acerca da onipresença de

²³⁹ No original: “recycles highly selective pasts into what is presented as the inevitable future”.

um presente imediato que ‘patina’ sobre si mesmo (Hartog, 2021; Cesarino, 2022) e do sentimento paradoxal de uma ‘inércia acelerada’ que atravessam a experiência contemporânea do tempo guiada por processos de contínua otimização sem transformação real e que em grande medida derivam das intersecções entre digitalização e neoliberalização (Pereira, Araújo, 2022; Chun, 2016; Cesarino, 2022). Entre outros efeitos, a algoritmização do tempo e do futuro agenciada pelos sistemas de IA resulta numa *repetição (quase) sem diferença*, produzindo futuros que são *cópias atualizadas do passado ou, no máximo, do presente*.

Talvez por isso tais sistemas se integrem tão bem às práticas e projetos que atualizam o extrativismo, o colonialismo, o racismo, o sexismo, a xenofobia, e incidem mais violentamente sobre populações que estão sujeitas (justamente) à repetição de violências históricas, sobretudo aquelas ligadas às “falhas de origem” (Ricarte, no prelo) da civilização ocidental moderna: o capitalismo, o colonialismo e o patriarcado. Atentos a essa imbricação, Ricarte entende a IA hegemônica enquanto uma “máquina bio-necro-tecnopolítica que serve para manter a ordem capitalista, colonialista e patriarcal do mundo”²⁴⁰ (Ricarte, 2022, p. 727) e Parisi e Dixon-Román (2020) reconhecem nela um *colonialismo recursivo* através do qual a tecnopolítica do capitalismo racial é continuamente regenerada. Tratam-se, nesse sentido, de *tecnologias conservadoras* que se alinham à “reprodutibilidade técnica de estruturas sociais” (Bonaldo, op. cit., p. 7) e tornam improvável a ruptura com os padrões históricos incorporados em seus modelos.

Tal qual os indivíduos que no contexto da crise permanente do neoliberalismo e da digitalização são constantemente incitados a atualizarem-se — e é fundamental lembrar que a entrada massiva da IA nas dinâmicas de trabalho das mais distintas áreas se torna mais uma demanda de atualização para este sujeito — para, na prática, permanecerem praticamente os mesmos (Chun, 2016), tais sistemas sociotécnicos estão, em sua imensa maioria, prevendo para que o *status quo* permaneça (quase) o mesmo (Hong, 2022b). Como ironiza o teórico da tecnologia Benjamin Bratton no tuíte a seguir (Figura 12), no atual ecossistema digital estamos nos tornando apenas treinadores de *modelos futuros do passado*.

²⁴⁰ No original: “bio-necro-technopolitical machine that serves to maintain the capitalist, colonialist and patriarchal order of the world”.



Figura 12: Tuíte de Benjamin Bratton. Fonte: Twitter/X.

Sobretudo em sistemas algorítmicos baseados na previsão de riscos, essa lógica programática leva à automatização e amplificação de um passado discriminatório, como diversos pesquisadores e entidades da sociedade civil vêm apontando enfaticamente nos últimos anos (Cf. Eubanks, 2019; Chun, 2021; Birhane, 2021; O’Neil, 2021; Bender et al, 2021; Varon, Costanza-Chock, Gebru, 2024). Intimamente relacionados à digitalização e automação de políticas neoliberais e a um governo dos efeitos sobre as causas (Morozov, 2018), os usos da IA para a “otimização preditiva”²⁴¹ materializam a conversão dos problemas sociais em objetos de cálculo. Sua força epistêmica, como vimos, não reside na acurácia das predições sobre o futuro, mas na legitimação de ações antecipatórias no presente (Hong, 2022), tornando as implicações da performatividade algorítmica destes usos de outra ordem. O resultado, como define Crawford, “é um *ouroboros estatístico*: uma máquina de discriminação auto-reforçada que amplifica as desigualdades sociais sob o pretexto de neutralidade técnica.”²⁴² (Crawford, 2021, p. 131, grifo nosso)

O fato desses sistemas incidirem com mais violência sobre populações historicamente marginalizadas, intensificando padrões históricos de injustiça e discriminação, não é para nada fortuito. Afinal, tais sujeitos já eram vítimas da repetição de padrões históricos de opressão mesmo antes da digitalização. O uso de sistemas preditivos automatizados sobre essas populações, principalmente para tomada de decisões na esfera pública, significa condená-las ainda mais às amarras dessas repetições históricas cíclicas, agora balizadas pela retórica da gestão do risco, da segurança, da otimização e da neutralidade tecnológica. Ou seja, pela (suposta) transformação em técnica daquilo que é de ordem política.

²⁴¹ Para uma pesquisa específica sobre esse tipo de aplicações ver “Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms that Optimize Predictive Accuracy”: <https://predictive-optimization.cs.princeton.edu/>.

²⁴² No original: “is a statistical ouroboros: a self-reinforcing discrimination machine that amplifies social inequalities under the guise of technical neutrality.”

Nos ecossistemas informacionais mediados por algoritmos preditivos, das redes sociais às IAs Generativas, a repetição massiva produzida pela algoritmização tem como um de seus principais efeitos uma homogeneização cultural em larga escala ou, como propõe Mann (2024), uma *IA-formização* (*IA-formization*, em inglês) que tem tornado os ambientes digitais mais monolíticos e menos diversos. Tais ecossistemas tendem a se tornar sistemas conservadores e homogêneos baseados na *repetição amplificadora do similar*.

É claro que tal processo já estava em curso mesmo antes do ‘hype da IA’ e era facilmente perceptível sobretudo na ‘pasteurização’ da paisagem informacional das redes sociais. Contudo, a abordagem *all-in-one-model* (Varon, Costanza-Chock, Gebru, 2024) das atuais IAs Generativas e a escala e velocidade com que este paradigma está se expandido modificam radicalmente o escopo do problema. Beiguelman (2021), refletindo sobre essa questão no âmbito da cultura visual, afirma que a IA pode se tornar um ponto de vista tão hegemônico sobre nossos modos de olhar, que não tenhamos mais a possibilidade de ter um olhar que não dialogue com os padrões que constituem seus modelos. Diante disso, ela propõe a hipótese de que a homogeneização dos nossos modos de ver, perceber e figurar a realidade pelos procedimentos da inteligência artificial estariam gestando uma “eugenia maquínica” (Beiguelman, 2021, p. 118).

É curioso notar como o ‘*eterno retorno*’ do passado que assombra a cronopolítica algorítmica também convive com novas formas de falsificação deste, principalmente no campo das imagens com os *deepfakes* e os passados fictícios que estas imagens produzem (Beiguelman, op. cit.), seja na criação de ‘metaversos do passado’, ou no uso de IA para ‘reviver’ pessoas já falecidas²⁴³, como no polêmico caso brasileiro do clipe com Elis Regina para a Volks em 2023. Tais fenômenos evidenciam como os *usos algoritmizados do passado* podem produzir, simultaneamente, sua repetição contínua e novas formas de falsificação deste. À princípio paradoxais, essas duas vertentes, na verdade, atestam como os futuros maquínicos da IA se baseiam em *futuros do passado* que transformam a memória (ou mesmo a história) não em elaboração sobre os espaços de experiência coletivos (Koselleck, 2006), mas em mera mercadoria (Beiguelman, op. cit). A repetição recursiva do passado, seja ‘verdadeira’ ou ‘falsa’, evidencia como a crise contemporânea do futuro passa também por uma crise da

²⁴³ Cabe lembrar como a ‘greve de Hollywood’, organizada em 2023 pelo sindicatos de atores e roteiristas, tinha como uma de suas principais pautas a potencial exploração e uso não autorizado de suas imagens, vozes e performances por IAs, sem compensação justa ou consentimento.

memória, “como se não pudéssemos conviver com o passado e só fosse possível fazer sua cópia, não sua preservação pela memória” (ibid., p. 140).

As implicações da homogeneização maquínica ou *IA-formização* sobre a qual viemos discorrendo tendem a se tornar ainda mais agudas pelo fato dos modelos algorítmicos passarem a ser cada vez mais treinados com dados gerados por eles mesmos. Isso ocorre não somente por conta da demanda infinita e veloz por dados que a atual corrida dos grandes modelos de IA impõe, mas simplesmente porque os próprios dados disponíveis nos ecossistemas digitais serão cada vez mais gerados sinteticamente. Samuel (2023), refletindo sobre os efeitos da proliferação das IAs Generativas, principalmente na linguagem escrita, afirma que “em vez de nos preocuparmos com a perda de uma originalidade que talvez nunca tenhamos possuído, deveríamos falar sobre o risco dessa tecnologia corroer a ‘diversidade’ ou a ‘flexibilidade’ do pensamento — e substituí-la pela homogeneização.”²⁴⁴. Estaríamos testemunhando assim a realização do prognóstico de Flusser ao alertar que o “estágio redundante” é a tendência do funcionamento dos programas das máquinas cibernéticas: “o programa do aparelho é eliminar todos os ruídos, e isso significa que o seu *programa é tornar tudo redundante*” (Flusser, 2017, grifo nosso).

Privilegiando ações entre o presente e o futuro próximo, essas tecnologias parecem mais ‘inofensivas’ e neutras também pela microtemporalidades e fragmentações através das quais operam e são produzidas, contrastando com a própria escala dos modelos algorítmicos mais avançados, especialmente os grandes modelos de linguagem. Seja nas micropredições geradas pelos modelos (focadas no ‘próximo passo’, na ‘próxima palavra’, no ‘próximo pixel’, etc) ou no chamado microtrabalho envolvido na cadeia produtiva de tais sistemas, essa *múltipla fractalização temporal* através qual as maquinarias da IA operam ajuda a tornar mais opacos tanto os efeitos sistêmicos produzidos pela mediação algorítmica quanto a superestrutura extrativa (de dados, de trabalho, de recursos energéticos e minerais etc) que a viabiliza (Cf. Crawford, 2021). Afinal, como vimos especialmente no estudo de caso do TikTok, o uso de microtemporalidades tem sido fundamental para a potencialização e aceleração do extrativismo digital (ver item 3.3.1).

Como pontua Bonaldo em diálogo com Hong (2022) de forma assertiva: “*não é a eliminação da contingência, mas as técnicas de seu controle, que estão em jogo*” (Bonaldo, op. cit., p. 6,

²⁴⁴ No original: “Instead of worrying about the loss of an originality that perhaps we never possessed, we should talk about the risk of this technology eroding “diversity” or “flexibility” of thought — and replacing that with homogenization.”

grifo nosso). Assim, nas cronopolíticas da IA, o *sequestro do futuro* (Bruno, 2020) via predição algorítmica se dá tanto pela redução das possibilidades às probabilidades, quanto pela forma com que esses sistemas sociotécnicos tornam improvável a disrupção com relação aos padrões nele incorporados — enquanto a disrupção, principalmente sob a forma da Singularidade, é continuamente anunciada por seus produtores e promotores. Ao mesmo tempo, é evidente como a entrada massiva destas tecnologias nos mais variados âmbitos vêm causando rupturas históricas que são efetivas, como as que ocorrem no mundo do trabalho, por exemplo. O que estamos enfatizando é a relação existente entre a lógica operativa interna destas máquinas e a repetição mais ampla de continuidade e ampliação de desigualdades históricas.

Nesse sentido, os atuais sistemas de IA ‘nos dizem’ *como as coisas foram e como elas provavelmente serão* caso se mantenham como estão (e sabemos bem como estão). Eles pouco ajudam a imaginar como as coisas *poderiam ser* caso se instaure uma *ruptura* (epistêmica, política, social, cultural, etc) com os padrões dos passados datificados nos modelos algorítmicos. Apesar de que a IA Generativa começa a manifestar capacidade de produzir algo novo em domínios específicos, como a matemática²⁴⁵, este ainda é um grande desafio para a área.

Essa limitação, no entanto, não precisa ser um destino das tecnologias algorítmicas. Afinal, como pontuamos anteriormente, *transformar o contingente em inevitável* é uma das estratégias privilegiadas das Big Techs para controlar os rumos do desenvolvimento tecnológico nos mais diversos setores. Como pontua Reed ao refletir sobre as múltiplas e emaranhadas crises que enfrentamos atualmente,

[o] desenvolvimento tecnológico desenfreado é, em parte, cúmplice na amplificação dessas crises, mas o é principalmente por estar tão enraizado em diagramas sociopolíticos específicos, que impõem muito mais restrições delimitadas, por exemplo, àquilo que *os algoritmos fazem*, do que àquilo que *poderiam fazer*. (Reed, 2018, p. 7)

Para refletirmos um pouco mais sobre a defasagem entre *o que os algoritmos fazem* e *o que poderiam fazer* e concluirmos este tópico, retomaremos brevemente o ‘fio da cibernética’ e da filosofia da computação. Conforme vimos no tópico 3.2.1, os sistemas cibernéticos se baseiam numa estrutura espiralar em que cada *feedback looping* determina seu devir

²⁴⁵ Ver, por exemplo: “AI scientists make ‘exciting’ discovery using chatbots to solve maths problems”. <https://www.theguardian.com/science/2023/dec/14/ai-scientists-make-exciting-discovery-using-chatbots-to-solve-maths-problems>

parcialmente a partir dos movimentos circulares passados (Hui, 2019). O dispositivo cibernético da recursividade opera através de numa temporalidade circular que tende ao regresso infinito (Hayles, 2017). Ao mesmo tempo, como enfatizam autores como Hui (op. cit.) e Parisi (2017), em diálogo com as reflexões de Simondon (2008) sobre a tecnicidade dos objetos técnicos, um fator fundamental dos sistemas computacionais baseados na recursividade é que eles contém, em si mesmos, *margens de indeterminação* que lhes dotam de capacidades de autoaprendizagem e autoatualização ao longo do tempo, constituindo um aspecto fundamental de sua dinâmica evolutiva e plástica. A recursividade implica, assim, um modelo temporal e processual entrelaçado com a contingência, seja no sentido do modo como os sistemas cibernéticos transformam os erros e falhas em meios de aperfeiçoar seu aprendizado (Hui, 2019; Parisi, Dixon-Román, op. cit.), seja na indeterminação que é constitutiva das máquinas computacionais (Fazi, 2018). São justamente essas margens que fazem com que a repetição recursiva não seja igual à repetição mecânica. Para utilizar uma metáfora gráfica: enquanto uma se move em espiral, a outra se move em círculos coincidentes.

Beatrice Fazi, explorando uma “ontologia da contingência” como constitutiva dos meios computacionais e propondo uma alternativa às perspectivas que concebem a computação como simples reducionismo instrumental da complexidade do vivo através de um determinismo algorítmico, defende que “a computação está sempre exposta à indeterminação”²⁴⁶ (Fazi, 2018, p. 116). A partir das concepções de *incompletude* de Gödel e de *incomputabilidade* de Turing, fundacionais para a formalização teórica e axiomática da computação nos anos 30²⁴⁷, a autora propõe que a computação é um “*processo de determinar a indeterminação*”²⁴⁸ (ibid., p. 1, grifo nosso). Essa indeterminação, explica, não é a mesma

²⁴⁶ No original: “computation is always exposed to indeterminacy”.

²⁴⁷ “Em 1931, o lógico o lógico Kurt Gödel determinou a incompletude de sistemas axiomáticos formais, demonstrando que existem proposições que parecem ser verdadeiras quando vistas de fora do sistema, mas que não podem ser provadas ou refutadas dentro do sistema em questão. Em 1936, Alan Turing baseou-se na descoberta de Gödel para mostrar que algumas funções são incomputáveis, descrevendo assim os limites das máquinas de computação antes de qualquer máquina desse tipo ser construída. A formalização de Turing formulou o método computacional em termos de um procedimento (isto é, um algoritmo) que resolveria um problema em um número finito de etapas sequenciais. Esta formalização da computação, no entanto, também revelou a existência desafiadora de problemas incomputáveis – nomeadamente, problemas que não podem ser resolvidos através de meios algorítmicos porque os passos envolvidos no seu processamento são infinitos.” (Fazi, op. cit, p. 5, tradução nossa).

²⁴⁸ No original: “process of determining indeterminacy”.

do vivo, mas um tipo de *indeterminação lógico-matemática* interna ao processamento computacional.

Num texto em que recoloca o clássico problema de Turing sobre se as máquinas podem pensar (Turing, 1950) como “as máquinas podem pensar *algo novo?*” (Fazi, 2019, grifo nosso), a autora conclui que a possibilidade de conceber a criação do novo pelas máquinas computacionais, passa, antes de tudo, pela superação do paradigma que crê que elas devem (ou estão restritas a) simular o humano, o que denomina de *paradigma simulativo*. Segundo Fazi, “o paradigma simulativo obscurece o que é excitante e potencialmente criativo num modo de raciocínio que é puramente computacional e que é (...) dramaticamente estranho para nós”²⁴⁹ (ibid., p. 822). Assim, para ela, a questão sobre a produção do novo no contexto computacional deveria ser reformulada levando em consideração as especificidades das máquinas computacionais e sua autonomia ontológica e epistemológica.

As proposições de Fazi sobre a natureza contingente e o potencial criativo da computação nos ajudam a iluminar uma tensão que parece ser constituinte do modelo de racionalidade algorítmico. Por um lado, os algoritmos enquanto máquinas cibernéticas contém intrinsecamente uma parcela de indeterminação que é fundamental para sua lógica programática e suas potencialidades. Por outro, é inegável também que parte dos efeitos de seus usos, sobretudo em sua vertente hegemônica, consiste numa diminuição, justamente, da agência do indeterminado, do imprevisível, do improvável e do acaso nos fluxos de automatização nos quais eles agem. Dito de outro modo, se por um lado, o incalculável e o indeterminado são constitutivos dos mundos gerados e geridos pela datificação e algoritmização, por outro, estes mesmos mundos excluem (ou ao menos visam neutralizar a potência e a agência) aquilo que não pode ser transformado em cálculo (Mbembe, 2019b).

Uma das questões fundamentais que se desdobram a partir desta tensão é como as margens de indeterminação destes cada vez mais complexos sistemas técnicos têm sido (ou não) exploradas, por quem e com que finalidade. Afinal, o fato de existirem formal e tecnicamente, não é sinônimo de que tenham agência, sobretudo num contexto em que o desenvolvimento da IA tem sido pautado quase que unilateralmente pela maximização do lucro de grandes corporações de tecnologia. Nesse sentido, certamente um dos desafios cruciais da entrada radical da IA em nossas ecologias cognitivas planetárias (Hayles, 2017) é

²⁴⁹ No original: “the simulative paradigm obscures what is exciting and potentially creative in a mode of reasoning that is purely computational, and which is, insofar as it is purely computational, dramatically alien to us”.

como explorar as potencialidades epistemológicas e criativas específicas dos modos de existência algorítmicos sem que os futuros produzidos através da IA se tornem repetições recursivas de nossos passados-presentes colonialistas e extrativistas. Acompanhando Bonaldo, nos parece evidente que a possibilidade de uma computação que seja capaz de sair da mera repetição “não é uma tarefa esgotável por discussões técnicas” (Bonaldo, op. cit. p. 3). Afinal, muito mais do que uma impossibilidade tecnológica, trata-se de um impasse ético-político (ibid.). Por isso, como aponta o autor junto a Canclini, “a pergunta no que pensam os algoritmos é insuficiente se não discutirmos ao mesmo tempo por que são tão poucos aqueles que os fazem pensar e colhem seus resultados” (Canclini, 2021, p. 123),

[t]ão poucos aqueles que, controlando os “vetores da informação” (WARK, 2022), dominam de fato a sociedade global. Tão poucos aqueles que, monopolizando os dados e os parâmetros, pesos nas matrizes de análise, determinam o futuro naturalizando previsões desde um ‘*deus in machina*’. (Bonaldo, op. cit., p. 22, grifo nosso)

Levar a sério a indagação sobre o que condiciona as defasagens entre o que os *algoritmos fazem* e o que *poderiam fazer* (e quais defasagens são essas) nos parece um passo fundamental para uma disputa pelas possibilidades da tecnologia, por sua tecnodiversidade (Hui, 2020), e para um resgate da contingência no interior não só das máquinas computacionais mas das culturas, epistemologias e políticas algorítmicas. Afinal, como pontua Mbembe:

é sabido que, enquanto uma parte do tecnológico acontece efetivamente na própria potência da matéria, só existe tecnologia nos distintos modos de emprego dessa potência, ou seja, na forma como ela é posta em movimento por atores situados, ao longo de tempos e espaços sociais diversos. (Mbembe, 2021, p. 38)

4.3 Antropoceno e Tecnoceno: a dupla aceleração de nossa época

A aceleração tecnológica e aceleração climática compõem hoje a encruzilhada cronopolítica mais fundamental para a compreensão das relações entre espaços de experiência e horizontes de expectativa (Koselleck, 2006) deste tempo de urgências, catástrofes e riscos que calhou de ser o nosso (Cf. Haraway, 2016; Mbembe, 2021; Stengers, 2015). Elas conformam aquilo que Viveiros de Castro, numa palestra recente, chamou de *dupla aceleração de nossa época*, a saber:

a catástrofe cumulativa do Antropoceno, de um lado, com tudo o que isso implica de fracasso civilizatório, de ecocídio planetário e talvez de suicídio biológico, e, de outro lado, *o desenvolvimento tecnológico muito rápido dos dispositivos materiais de IA (...)*. (Viveiros de Castro, 2023, p. 1, grifo nosso)

Esse choque de acelerações’ pode também ser assim descrito:

A aceleração “intencional” da máquina capitalista, posta como solução de nossa presente miséria antropológica encontra-se em uma situação de contradição objetiva com uma outra aceleração nada intencional, o implacável processo de retroalimentação positiva das transformações ambientais deletérias para o *Umwelt* da espécie. (Danowski, Viveiros de Castro, 2017, p. 79).

As radicais transformações ocorridas nos últimos anos neste dois domínios e os prospectos sobre o futuro que elas suscitam são de um caráter completamente distinto das mudanças prometidas pelo moderno regime de tempo (Simon, 2021). O que viemos designando *futuros maquínicos* e aquilo que poderíamos chamar de *futuros antropocênicos* manifestam modalidades de futuro que desafiam (quando não tornam inócuos) os modos modernos de conhecer, organizar, conceitualizar, imaginar e se relacionar com o tempo. Como observa Turin, o próprio conceito moderno de História enquanto processo coletivo singular (Koselleck, 2006) “restrito a uma Humanidade que se temporaliza às custas do domínio e da exploração da Natureza vem se tornando cada vez mais opaco e fragilizado, perdendo sua capacidade de oferecer respostas e quadros de inteligibilidade diante desses novos problemas” (Turin, 2023, p. 706). Mesmo campos do conhecimento não restritos às ‘coisas humanas no tempo’, como a geologia, têm se mostrado incapazes de lidar com esse novo *tempo do tempo*, como o atesta a rejeição do reconhecimento oficial do Antropoceno como uma nova época na Escala de Tempo Geológico pelos “guardiões do tempo” da União Internacional de Ciências Geológicas (IUGS) em março de 2024²⁵⁰ (Esteves, 2024), revelando um descompasso não só entre a *ciência* e (aquilo que se costumava chamar de) *natureza*, mas também entre *ciência* e *sociedade*.

Esse descompasso é sintomático das confusões e desafios advindos da “súbita colisão dos Humanos com a Terra” (Danowski, Viveiros de Castro, 2017, p. 30) e do consequente desmoronamento da separação entre história humana e história natural (Chakrabarty, 2009) advindos com o Antropoceno.

²⁵⁰ A proposta de reconhecimento de uma nova época geológica iniciou-se em 2009, com a criação de um Grupo de Trabalho sobre o Antropoceno (*Anthropocene Working Group – AWG*) dentro da Comissão Internacional de Estratigrafia da União Internacional de Ciências Geológicas (IUGS). Para ser reconhecido oficialmente, o Antropoceno precisaria ser registrável geologicamente por meio de um estrátotipo global (GSSP, ou “golden spike”), ou seja, uma marca inequívoca de sua existência nas camadas da Terra. O AWG sugeriu que o Antropoceno começaria por volta de 1950, marcado pela presença de isótopos radioativos de testes nucleares nos sedimentos e gelo, além de outras marcas como plásticos e mudanças atmosféricas. Após 14 anos de estudos, em outubro de 2023, o grupo formalizou a proposta de que o Antropoceno sucederia o Holoceno, iniciado há aproximadamente 11.700 anos. No entanto, em 2024, a proposta foi rejeitada por 12 votos a 4. Entre os argumentos contrários, destacou-se a percepção de que o evento escolhido como marco inicial era demasiado restrito, recente e inadequado para definir uma nova época geológica.

Nesse *presente denso* (Haraway, 2016), em que nos encontramos emaranhados em processos em muitos sentidos inaugurais, o próprio tempo “parece estar, não apenas acelerando, mas mudando qualitativamente ‘tempo todo’” (Danowski, Viveiros de Castro, op. cit., p. 23) resultando numa “instabilidade metatemporal” (ibid.) que nos confronta com a urgência desse tempo ‘fora do eixo’. Como observa Mbembe,

[p]or mais que as velocidades estejam em constante disparada e as distâncias incessantemente sendo conquistadas, o tempo concreto, o da carne do mundo e da sua respiração e o do Sol que envelhece, não é mais extensível ao infinito. *Estamos em plena era da combustão do mundo*. Portanto, é à urgência que estamos confrontados. (Mbembe, 2021, p. 27, grifo nosso)

A urgência e o incômodo que nos atravessam sintomatizam, entre outras coisas, a sensação de *perda de controle* em relação a tais fenômenos. Isso porque o duplo vetor da aceleração contemporânea deriva de processos que, apesar de obviamente terem sido causados pelo ser humano — mais precisamente, por aqueles que nunca reconheceram todo humano como Humano —, escapam de várias maneiras ao seu controle. Isso não significa que as pretensiosas ambições da engenharia da natureza, da sociedade e do humano (Simon, Tamm, op. cit.) movidas pela “*flecha-(assassina)-do Tempo do Tecno-Heróico*” (Le Guin, 2021, p. 23) estejam superadas, como os últimos avanços da IA e a ideologia tecnólatra que os sustentam deixam nítido e sobre as quais nos deteremos adiante.

“Imprevisibilidade, incompreensibilidade, sensação de pânico diante da perda do controle, senão mesmo perda da esperança: eis o que são certamente desafios inéditos para a orgulhosa segurança intelectual da modernidade” (Danowski, Viveiros de Castro, op. cit., p.). A sensação de perda de controle sobre um futuro que se torna ameaçador e imprevisível está diretamente relacionada ao fato de que se tratam de processos movidos ou, mais precisamente, acelerados, por *escalas e temporalidades sobre-humanas*. Por sua incomensurabilidade, eles podem ser considerados aquilo que Timothy Morton (2013) chama de *hiperobjetos*, um fenômeno e/ou entidade que é tão grandioso em sua escala e temporalidade que excede nossa capacidade de percepção. Seus impactos sistêmicos e complexos ocorrem no Sistema Terra, produzindo experiências que colapsam o humano como centro da organização espaço-temporal e suscitando um angustiante sentimento de *inevitabilidade*, seja acerca da aceleração climática ou tecnológica.

É evidente, no entanto, que a aceleração das estruturas e experiências temporais não é um fenômeno recente. Estudos como os de Hartmut Rosa (2019) e Judy Wajcman (2015) mostram que a aceleração técnica e social, sob a lógica do progresso, constitui uma

característica fundamental da Modernidade. Os autores exploram as relações entre sociedade moderna e aceleração, evidenciando como os padrões temporais modernos estão intrinsicamente ligados ao uso das tecnologias e à mercantilização do tempo promovida pelo capitalismo industrial. Além disso, os debates sobre aceleração e Modernidade normalmente remetem à clássica análise de Max sobre o capitalismo e sua necessidade constante de acelerar o capital (Wajcman, op. cit.). Ou seja, quanto mais rapidamente o dinheiro puder ser convertido na produção de bens e serviços, maior será a capacidade de expansão e valorização do capital. Afinal, quando tempo é dinheiro, a velocidade se torna um bem inquestionável. Nesse processo, as inovações tecnológicas desempenham um papel central, pois os avanços na transmissão de informações, no transporte de mercadorias e na mobilidade dos corpos reduzem tanto os custos quanto o tempo necessário para a circulação do capital em escala global — fenômeno que Marx descreveu como a “aniquilação do espaço pelo tempo” (ibid.). Sob essa perspectiva, nossa atual “sociedade de alta-velocidade” (ibid.), marcada pelo papel central da digitalização, pode ser vista como uma intensificação do movimento histórico de aceleração do capital.

Contudo, apesar dessa inegável continuidade histórica, nos parece que os tipos de agentes técnicos envolvidos na aceleração contemporânea são de outra natureza em relação àqueles que ‘aceleraram’ a Modernidade²⁵¹. Isso porque, como mencionamos acima, eles são movidos por temporalidades que ultrapassam nossas capacidades cognitivas e sensoriais. Se, na Modernidade, a aceleração esteve associada à maximização da eficiência produtiva e à compressão espacial promovida por inovações como o transporte ferroviário e as telecomunicações, na contemporaneidade, a aceleração digital desdobra-se em ritmos e escalas que escapam à apreensão humana. Algoritmos financeiros, modelos preditivos de IA e redes automatizadas de comunicação funcionam em velocidades incomensuráveis, tornando o próprio tempo um elemento gerenciado de forma abstrata e desvinculada da percepção subjetiva.

Para Franco Berardi, essa aceleração dos fluxos de informação para além dos limites da subjetividade humana e a infiltração da velocidade da *infoesfera* sobre a *piscoesfera* tem relação direta com a crise do futuro que atravessamos (Berardi, 2019, 2020). Nesse sentido, a atual impossibilidade de imaginar futuros alternativos seria também resultado desta

²⁵¹ Aqui também podemos pensar num contraponto entre máquina externa e máquina interna.

“aniquilação do tempo” (Celis Bueno e Schutz, 2021b) provocada pela necessidade intrínseca de aceleração do capitalismo.

É fundamental lembrar também que o duplo vetor aceleracionista que estamos explorando está estreita e materialmente entrelaçado, ao ponto de pesquisadores como Costa (2021) proporem colocar a ênfase na denominação de nossa época histórica justamente no desenvolvimento tecnológico e em sua capacidade de afetar de modo material todo o planeta. Tendo como marco inicial a ameaça atômica, para a autora, o Tecnoceno designa

a época em que, mediante o desenvolvimento de tecnologias de alta complexidade e altíssimo risco, deixamos pegadas no mundo que expõem não só as populações de hoje, mas também as gerações futuras, de nossa espécie e de outras espécies, nos próximos milênios.²⁵² (Costa, 2021, p. 10)

Ao longo deste capítulo utilizaremos o termo Tecnoceno reconhecendo-o como um possível modo de qualificar o Antropoceno, tal qual propõe Costa (op. cit.), mas sem recorrermos a uma substituição. Deste modo, utilizaremos ambos os termos, Antropoceno e Tecnoceno, como significantes entrelaçados, mas que permitem dar contornos mais precisos, respectivamente, à questão da aceleração climática e tecnológica que nos interessam de sobremaneira.

O entrelaçamento material da dupla aceleração que marca nossa época é especialmente evidente no âmbito da IA, seja por conta dos novos riscos sistêmicos que ela desencadeia — ainda que, como argumentamos adiante, tais riscos vem sendo em grande medida capitalizados pelos financiadores do setor para alimentar imaginários sobre um ‘*apocalipse da IA*’ — ou por conta dos altíssimos custos ambientais envolvidos no funcionamento de suas infraestruturas, especialmente de seus *datacenters*. Em 2024, obteve bastante repercussão um estudo realizado pelo *Washington Post* em parceria com a Universidade da Califórnia que concluiu que o ChatGPT consome uma garrafa d’água e 0,14 kWh para gerar um email de 100 palavras (Verma, Tan, 2024). A ‘solução’ proposta por tecnoboys como Sam Altman para suprir a demanda energética insaciável dos grandes modelos é nada menos do que fusão nuclear (Paddison, 2024).²⁵³

²⁵² No original: “la época en la que, mediante la puesta en marcha de tecnologías de alta complejidad y altísimo riesgo, dejamos huellas en el mundo que exponen no solo a las poblaciones de hoy, sino a las generaciones futuras, de nuestra especie y de otras especies, en los próximos milenios.”

²⁵³ Ao mesmo tempo, é inegável a importância dos modelos algorítmicos para a Ciência do Clima na simulação de cenários futuros. Como afirma Chun (2021), tais modelos evidenciam como a predição de futuros prováveis pode ser utilizada não para que sejam aceitos como inevitáveis, mas justamente para que não se realizem.

Uma expressão corriqueira utilizada para descrever esses novos futuros, especialmente os tecnológicos, parece denotar de forma sutil a percepção de que a humanidade já não está no ‘no comando do tempo’: ‘*o futuro chegou*’, dizem por toda parte as manchetes que relatam a mais nova proeza das ‘super IAs’. Ou seja, não somos mais nós que “criamos a eterna velocidade onipresente” e avançamos corajosamente em direção ao futuro, como nas palavras futuristas de Marinetti, mas é ele que, guiado agora por outros agentes, frutos autonomizados daquele mesmo progresso (para o ‘bem’ e para o ‘mal’), vem até nós. “O futuro está vindo em nossa direção a toda velocidade”, como diz Stengers (2024, p. 151).

Numa via, é como se o progresso técnico tivesse acelerado a tal ponto que se rompeu ‘de nós’, e, em seguida, dobrou-se ‘sobre nós’. Noutra, a brutal “intrusão de Gaia” (Stengers, 2015) não se refere mais a uma “barbárie por vir” (ibid.), mas em plena realização. Neste sentido, os futuros ‘*AI-driven*’ e ‘*Climate-driven*’ podem ser entendidos enquanto futuros pós-humanos ou, talvez para usarmos uma denominação menos teleológica, *futuros outros-que-humanos*. Os processos que os movem estão na origem de uma “crise de autoestima” (Viveiros de Castro, op. cit.) pela qual a ainda hegemônica civilização euro-americana e seu protagonista, o *Homem Branco Transcendental*, estão passando. Uma crise que deriva da “percepção de que [seu protagonista] se tornou incapaz de compreender plenamente o que, no entanto, é capaz de fazer: da bomba atômica e seus efeitos ‘inconcebíveis’ às máquinas de IA” (ibid., p. 3). Isso porque, se por um lado, tanto Antropoceno quanto Tecnoceno são frutos diretos da arrogância do *anthropos*, por outro, seus efeitos lhe retiram privilégios onto-epistemológicos que detinha até então em relação àqueles que Viveiros de Castro chama de os “outros habitantes da tríplice fronteira antropológica”, a saber, *animais e máquinas*²⁵⁴. Esse triângulo formado por humanos, animais e máquinas²⁵⁵, cujos vértices podem ser relacionados a três tipos de inteligência (respectivamente, a *cultural*, a *natural* e a *artificial*), para o autor, continua a definir os contornos das metafísicas

²⁵⁴ Importante destacar que o mesmo privilégio foi condição de possibilidade para a concretização de tais processos. Afinal, como pontuam Bruno, Pereira e Faltay (2023, p. 237), “grande parte dos desastres em curso foram produzidos a partir do arrogante pedestal do excepcionalismo humano (ao qual nem todos os humanos tinham acesso, uma vez que o lugar de domínio do humano frente à natureza tem sido acompanhado de toda sorte de hierarquias e discriminações entre humanos).”

²⁵⁵ Como salienta o autor, todos esses termos são “sinédoques para categorias internamente diferenciadas” (p.9). De modo que “*animais* = bichos, plantas, fungos, amebas, procariotas etc; *espíritos* = almas, espectros, divindades et; *humanos* = a espécie darwiniana / a condição arendtiana; *máquinas* = relógios → máquinas a vapor → máquinas eletrônicas → computadores e aplicativos de IA.” (ibid.)

contemporâneas. “No centro de gravidade desse triângulo”, afirma, “está a Terra.” (ibid., p. 9)

Apesar de que essa tríplice fronteira antropológica ruiu há muito em campos como a cibernética, matriz fundamental do modelo de racionalidade algorítmico, como argumentamos amplamente nesta investigação²⁵⁶, e em toda a produção teórica alinhada ao pós-humanismo, como o paradigmático *Manifesto Ciborgue* de Haraway (2019)²⁵⁷, parece ter sido somente após a entrada em cena dos eloquentes grandes modelos de linguagem e a aceleração da corrida pela Inteligência Artificial Geral (junto da expectativa e do temor que se tornem — finalmente! — autoconscientes) que a “confusão de fronteiras” (Haraway, op. cit.) causada pelos *habitantes maquínicos* desse triângulo foram sentidas de um modo mais amplo e radical. Sintomático da ameaça ao excepcionalismo humano e da *ferida narcísica*²⁵⁸ (Malabou, 2019) que essa mutação encarna foi o forte retorno da pergunta por “aquilo que realmente nos faz humanos” ou no que consiste a “verdadeira inteligência” (Cf. Chomsky, Roberts, Watumull, 2023) e permita denunciar a falsidade das outras e nos reassegurar uma segurança onto-epistemológica em relação a esses agora ameaçadores cognoscentes (ora porque demasiado estúpidos, ora porque demasiado espertos), como comentamos na análise sobre o ChatGPT.

Essa confusão na tríplice fronteira relaciona-se àquilo que Berardi (2019) descreve como a infiltração da lógica do *bios* nas máquinas e da lógica da *tekné* para a vida humana e social, tornando ambíguas as distinções entre uns e outros. Como já vislumbrava Haraway em 1985,

[a]s máquinas do final do século XX tornaram completamente ambígua a diferença entre o natural e o artificial, entre a mente e o corpo, entre aquilo que se autocria e aquilo que é externamente criado, podendo-se dizer o mesmo de muitas outras distinções que se costumavam aplicar aos organismos e às máquinas. *Nossas máquinas são perturbadoramente vivas e nós mesmos perturbadoramente inertes*” (Haraway, [1985] 2019, p. 162, grifo nosso)

No supracitado texto, Viveiros de Castro propõe a tese de que, progressivamente desde o século XVII, no Ocidente, “as máquinas vieram ocupar o lugar que antes os espíritos

²⁵⁶ Lembremos, por exemplo, que o seminal livro de Wiener, publicado em 1948, se chama *Cibernética ou controle e comunicação no animal e na máquina*.

²⁵⁷ No manifesto, Haraway descreve as três quebras de fronteira cruciais que o ciborgue encarna: a fronteira entre o humano e o animal, a fronteira entre o organismo e máquina; a fronteira entre o físico e o não físico.

²⁵⁸ Malabou (2019) propõe que a IA pode ser considerada uma *quarta ferida narcísica*: “Primeiro Copérnico, seguido por Darwin, depois a psicanálise, e agora o quarto golpe: a captura da inteligência por sua própria simulação, excedendo-a e transcendendo-a.” (p. 164, tradução nossa)

ocupavam nas cosmologias extramodernas” (op. cit., p. 10) e o que chama de “sistemas de IA neuromórfica” são os representantes mais bem-sucedidos dessa *substituição animista*. Ao invés do “*the ghost in the machine*”, como na fórmula de Ryle (1949) para descrever o dualismo cartesiano alma/corpo, agora a fórmula poderia ser descrita como “*the ghost as the machine*”, ou “*the machine as the ghost*”, propõe.

Mbembe (2021), em seu último livro, no qual caracteriza nossa época como a de um *Brutalismo* fruto da confluência de múltiplos processos de *fraturamento, fissuração e extração*, entre eles a aceleração tecnológica e a transição para uma civilização computacional, também associa a ascensão das tecnologias digitais a um “espetacular retorno do animismo” (p. 29). Para o autor, esse retorno seria fruto da “potencial fusão da mercadoria, da tecnologia e da religião” (p. 100). Esse novo animismo, afirma Mbembe, “se confunde com a razão eletrônica e algorítmica, que é tanto seu meio quanto seu envelope, e até mesmo seu motor” (p. 34).

A atualização do animismo identificada por Viveiros de Castro e Mbembe, em última instância, remete às próprias relações entre *técnica e magia* ou *técnica e religião* e cujas ressonâncias, sabe-se bem, nunca desapareceram do imaginário popular sobre a tecnologia. Como nos mostrou Simondon (2008), tais relações encontram-se na gênese do objeto técnico. Nesse sentido, essa espécie de ‘animismo maquínico’ do qual a IA é hoje protagonista pode ser visto como a *revitalização de um laço originário entre os modos de existência técnico e religioso através da magia*. Em seu clássico *O modo de existência dos objetos técnicos*, Simondon (op. cit.) propõe que o pensamento técnico e o pensamento religioso, dos quais derivaram, respectivamente, a *tecnicidade* e a *religiosidade*, teriam se desprendido a partir do que chama de *processo de defasagem da unidade mágica primitiva*, sob a forma de figura/objetivação (a tecnicidade) e fundo/subjetivação (a religiosidade). Em suas palavras:

Tecnicidade e religiosidade não são formas degradadas da magia, nem sobreviventes de magia; elas provêm do desdobramento do complexo mágico primitivo, reticulação do meio humano original, em figura e fundo. Por meio desse par, técnica e religião se tornam herdeiras da magia, e não cada uma por si mesma. A religião não é mais mágica do que a técnica; é a fase subjetiva do resultado do desdobramento, enquanto a técnica é a fase objetiva desse mesmo desdobramento. Técnica e religião são contemporâneas uma da outra e são, consideradas separadamente, mais pobres do que a magia da qual emergem. (Simondon, 2008, p. 191, tradução minha).

A relação primordial entre os modos de existência *mágico, técnico e religioso* teorizada por Simondon ajuda a explicar por que, em nosso *Tempo do Fim* (Anders, 2007), as escatologias do Antropoceno e do Tecnoceno começam a se sobrepor com tamanha força, como veremos adiante.

4.4 Habitando o cronotopo das ameaças e promessas finais: mitologias do fim

*O fim do mundo é um tema aparentemente interminável
— pelo menos, é claro, até que ele aconteça.*
Déborah Danowski e Eduardo Viveiros de Castro

Simon (2019, 2021), opondo-se às teses presentistas sobre o contemporâneo e insistindo na recuperação da categoria de futuro por meio de uma filosofia da história que não se limite às coisas exclusivamente humanas, argumenta que o tipo de percepção de mudança subjacente às visões do futuro do pós-guerra pelas sociedades ocidentais, especialmente as mobilizadas atualmente nos campos tecnológico e ecológico, compartilham expectativas de uma mudança que gira em torno de uma alteração radical, ou mesmo da extinção, do ser humano tal como o conhecemos. Tratam-se de mudanças existenciais profundas que ele qualifica como *sem precedentes*. Esse *futuro sem precedentes* implica um “futuro [que] deixa de ser feito da mesma matéria que o passado, [que] torna-se radicalmente *outro*, não-nosso, um tempo que exige nossa desaparecimento para aparecer” (Danowski, Viveiros de Castro, op. cit. 45). Segundo Simon, esses prospectos de *mudanças sem precedentes* relacionam-se à expectativa de um “evento singular e disruptivo” (Simon, 2021, p. 151), um “ponto de não retorno que é antecipado como um evento singular após o qual assume-se que os seres humanos perderão o controle sobre o domínio de suas próprias condições”²⁵⁹ (Simon, 2019, p. 7).

O exemplo mais paradigmático do tipo de evento disruptivo descrito por Simon é a hipótese da *Singularidade Tecnológica*²⁶⁰. Ainda que o afã de superação do humano permeia toda a história da IA (para muitos, inclusive, é seu *leitmotiv*), a tese de que a evolução acelerada da capacidade de inteligência dos computadores produzirá entidades com inteligência superior à humana que superarão a capacidade conjunta da massa cinzenta do planeta²⁶¹ foi popularizada principalmente por Vernor Vinge (1993) num texto intitulado *The Coming Technological*

²⁵⁹ No original: “This point of no return is the anticipated singular event, after which human beings are assumed to lose control over engineering their own condition”.

²⁶⁰ O termo Singularidade provém da física onde designa fenômenos tão extremos que as equações não são mais capazes de descrevê-los, como os buracos negros.

²⁶¹ Ainda que estejam estreitamente relacionados, os conceitos de Singularidade e Inteligência Artificial Geral (a chamada AGI) são distintos. A AGI refere-se a uma forma de inteligência artificial que é capaz de realizar qualquer tarefa cognitiva que um ser humano possa realizar. Ao contrário das IAs atuais, que são especializadas em tarefas específicas (IAs Fracas ou Estreitas), a AGI teria a capacidade de aprender e se adaptar a uma ampla gama de tarefas, possuindo compreensão e inteligência comparáveis às de um ser humano. Em resumo, a AGI é vista como um passo necessário para alcançar a Singularidade.

Singularity e, posteriormente, por Ray Kurzweil (2005) no livro *The Singularity is Near*. Para Vinge, trata-se de “uma mudança comparável à ascensão da vida humana na Terra” (ibid.). Outro popular defensor das teses da Singularidade é o historiador Yuval Harari (2016), para quem o avanço tecnológico estaria nos conduzindo a um novo estágio, protagonizado por um *Homo Deus*, dotado de habilidades como a imortalidade, a onisciência e o controle sobre o ambiente.

Tal como a volta de Cristo nas profecias milenaristas — uma ‘coincidência de topos’ para nada fortuita —, desde que começou a ser anunciada, a Singularidade sempre está (ou estaria) muito próxima. Em tons simultaneamente proféticos e pretensamente científicos, marca identitária dessa ‘trupe’, Vinge (op. cit.) previu que o evento divisor ocorreria nos (então) próximos 30 anos. No supracitado livro, Kurzweil (op. cit.) afirmava que ele ocorreria em 2025, mas numa entrevista recente a Joe Rogan, afirmou que será em 2029²⁶². Os profetas singularitários crêem que após o “Arrebatamento cibernético súbito” (Danowski, Viveiros de Castro, op. cit., p. 69) entraremos numa era pós-humana (ou, em seu léxico, *transhumana*) de fusão entre inteligência biológica e maquina que viabilizará, entre outras conquistas, a superação do nosso grande ‘empecilho’, a mortalidade²⁶³.

Segundo os singularitários californianos, a evolução acelerada até o Ponto Ômega, vista como inevitável e único destino possível da IA, seguirá uma *trajetória exponencial*. Nessas previsões, as concepções de um *progresso tecnológico linear*, que marcou o período keynesiano-fordista, são substituídas por concepções de um *progresso exponencial* que tem como modelo a *Lei de Moore*²⁶⁴ (Kurzweil, 2004), resultando num crescimento condensado e acelerado. Por conta dessa taxa explosiva de desenvolvimento, segundo Kurzweil (op. cit.), o crescimento tecnológico no século XXI será equivalente a 20 mil anos de progresso na velocidade atual.

A Lei de Moore é uma matriz epistemológica fundamental para os ‘tecnoboys do Vale’ e suas

²⁶² A entrevista foi realizada em março de 2024 e está disponível em: https://www.youtube.com/watch?v=w4vrOUau2iY&ab_channel=PowerfulJRE

²⁶³ Numa entrevista à Revista Piauí, Kurzweil afirmou não ter “dúvida que caminharemos para a imortalidade, seja por meio de processos que revertam o envelhecimento das células biológicas, ou mesmo por meio da transferência de nosso conteúdo cerebral para um meio físico além do corpo, um novo hardware. Um robô construído à nossa imagem e semelhança é um dos modelos.” (Piauí, 2010)

²⁶⁴ Segundo tal lei, formulada por Gordon Moore em 1965 baseando-se em suas próprias observações, um dos fundadores da Intel, o número de transistores em um circuito integrado, e por consequência a capacidade de processamento dos chips, dobra aproximadamente a cada dois anos, o que implica um crescimento exponencial na capacidade de processamento dos chips sem um aumento proporcional no custo.

visões (delirantes) de futuro e sociedade. Num texto intitulado *Moore Law for Everything*, Sam Altman (2021), CEO da OpenAI, defende que a lei deve ser aplicada às políticas públicas para “distribuir de forma justa parte da riqueza futura” em prol de um cínico “capitalismo para todos”. Chocante até para quem sempre esteve ciente do tecnosolucionismo promovido pelas elites tecnológicas, Altman afirma que “[a] melhor forma de aumentar a riqueza social é diminuir o custo dos bens”. Essa redução, claro, será viabilizada pela introdução de uma IA superpoderosa nas cadeias produtivas, uma vez que, segundo ele, a mão-de-obra é o principal obstáculo para o barateamento destas cadeias. A proposta de Altman é ‘simples’: uma versão aceleracionista da fórmula liberal de que primeiro é preciso fazer crescer o bolo — agora com a ajuda da IA! — para depois dividi-lo. A conclusão de sua proposta ‘descomplicada’ é que esse modelo de governo resultaria, além de tudo, num vantajoso *tecnosolucionismo populista*:

Um grande futuro não é complicado: precisamos de tecnologia para criar mais riqueza e de políticas para distribuí-la de forma justa. Tudo o que for necessário será barato e todos terão dinheiro suficiente para poder pagar. Como este sistema será enormemente popular, os decisores políticos que o adotarem cedo serão recompensados: eles próprios tornar-se-ão enormemente populares (Altman, 2021).

Uma das instituições do Vale do Silício mais influentes atualmente na reverberação das narrativas futuristas relacionadas à IA e à tecnologia de forma geral é o *Future of Life Institute* (FLI), *think tank* fundado em 2014 e que tem entre seus conselheiros Elon Musk, Stephen Hawking e até Morgan Freeman. Declarando-se com nada menos do que a missão de “preservar o futuro da vida”, a instituição ganhou repercussão mundial ao liderar a *Carta Aberta pela pausa de grandes experimentos com IA*, publicada em março de 2023 e assinada por mais de 1000 personalidades e pesquisadores (entre eles, Harari e, claro, o próprio Musk).

Um dos significantes fundamentais no léxico do pretensioso *Future of Life Institute* são os *riscos existenciais em larga escala* (ou apenas *x-risks*, em inglês). O termo, cunhado pelo filósofo e professor da Universidade de Oxford Nick Bostrom, figura altamente influente nas futurologias californianas, descreve os riscos originados de atividades humanas capazes de provocar “a aniquilação da vida inteligente originada na Terra ou a redução permanente e drástica do seu potencial para um desenvolvimento futuro desejável”²⁶⁵ (Bostrom, 2009), como a inteligência artificial, a biotecnologia, as armas nucleares e as mudanças climáticas. A partir do lançamento do ChatGPT, a noção, antes circunscrita a nichos específicos, se

²⁶⁵ No original: “the annihilation of Earth-originating intelligent life or the permanent and drastic curtailment of its potential for future desirable development”.

transformou em um dos principais motes para pautar o debate público sobre as implicações do uso da IA. Uma segunda carta, lançada em maio daquele ano pelo *Center for AI Safety*, intitulada *Declaração sobre Risco da IA* e assinada por especialistas e outras figuras influentes do meio, afirmava que “[a] mitigação do risco de extinção provocado pela IA deve ser uma prioridade global, juntamente com outros riscos à escala social, como pandemias e guerra nuclear.”²⁶⁶

Além dos *riscos existenciais*, outras ideologias que moldam os imaginários tecnofuturistas promovidos pelo Instituto são o *altruísmo efetivo* e o *longoprazismo* (respectivamente, *effective altruism* e *longtermism*, em inglês), espécies de ‘rebrandings’ do caldo ideológico místico-neoliberal que sempre fez parte do Vale do Silício (Cf. Barbrook, Cameron 1995), agora mais sintonizado com *zeitgeist* distópico da segunda década do século XXI. O primeiro pode ser descrito como uma abordagem (que se enuncia como) filosófica e ética que defende que influenciar positivamente o futuro a longo prazo deve ser uma prioridade moral do nosso tempo e é liderada por nomes como William MacAskill (2022) e Toby Ord (2021). Essa perspectiva sugere evitar a concentração em ‘urgências imediatas’, direcionando a atenção para questões que afetarão significativamente o *futuro da humanidade em séculos ou centenas de milhares de anos*.

Num texto intitulado *The case for strong longtermism*, em co-autoria com Hilary Greaves, MacAskill (2021), prevendo nosso futuro enquanto espécie, afirma que, se durarmos quanto uma espécie mamífera típica normalmente dura, ainda temos 200.000 anos pela frente. “Se a saga da humanidade fosse um romance, ainda estaríamos na primeira página”²⁶⁷, afirmam (Greaves, MacAskill, op. cit., p. 2). Os autores propõem então que, “para fins de avaliações de ações, podemos, em primeira instância, muitas vezes simplesmente ignorar todos os efeitos contidos nos primeiros 100 (ou até mesmo 1000) anos, focando principalmente nos efeitos no futuro a longo prazo”²⁶⁸ (ibid., p. 1). Para os longoprazistas, é claro, o desenvolvimento tecnológico é um dos principais meios de garantir ‘nosso futuro triunfante’ e aumentar a probabilidade de realizar ‘nosso potencial’ (Geburu, Torres, 2024), atributo considerado inerente à nossa ‘brilhante’ espécie. Isso inclui toda sorte de IAs superpoderosas, colônias espaciais, geoengenharia, simulações digitais e, claro, a velha obsessão de driblar a

²⁶⁶ Disponível em: <https://www.safe.ai/work/statement-on-ai-risk#open-letter>

²⁶⁷ No original: “If humanity’s saga were a novel, we would be on the very first page.”

²⁶⁸ No original: “for the purposes of evaluating actions, we can in the first instance often simply ignore all the effects contained in the first 100 (or even 1000) years, focusing primarily on the further-future effects.”

mortalidade com tecnologias como a criogenia (Gaglioni, 2023).

O longoprazismo é uma das principais causas do chamado *altruísmo efetivo*, projeto que defende “o uso de evidências e da razão para descobrir como beneficiar os outros tanto quanto possível, e agir com base nisso”²⁶⁹ (MacAskill, 2017, p. 1) priorizando supostamente questões globais. Suas raízes remontam ao ensaio *Famine, Affluence, and Morality*, publicado por Peter Singer em 197, que estabelece as bases da proposta e argumenta, entre outras coisas, que pessoas ricas têm o dever moral de doar recursos para causas humanitárias (Singer, [1971] 2015). A comunidade em torno do altruísmo efetivo conta atualmente com cerca de 7.000 membros²⁷⁰ e mobilizou a doação de milhões de dólares a instituições filantrópicas alinhadas à causa nos últimos anos.

Apesar de tentador, nos parece que mais do que rebater factualmente propostas e hipóteses como as da *Singularidade* (Vinge, Kurzweil), da *Lei de Moore para Tudo* (Altman), dos *Riscos Existenciais em Larga Escala* (Bostrom), do *Longoprazismo* e do *Altruísmo Eficaz* (MacAskill) e apontar a inabalável arrogância e cinismo dos (sempre tão) *white man saviours* envolvidos nesse conjunto ideológico, elas deveriam ser levadas a sério como aquilo que realmente são: *mitologias tecnoteológicas* (Danowski, Viveiros de Castro, op. cit.) que integram as *fabulações de fim do mundo* do nosso tempo de tantos fins mas ainda incertos recomeços. Como pontuam Danowski e Viveiros de Castro, “o regime semiótico do mito, indiferente à verdade ou à falsidade empírica dos conteúdos, instaura-se sempre que a relação entre os humanos como tais e suas condições mais gerais de existência se impõe como problema para a razão.” (op. cit., p. 21).

De modo geral, essas mitologias tecnoteológicas manifestam visões de uma “humanidade sem mundo” baseadas na “ideia de uma autofabricação do homem do futuro e de seu ambiente pela eugénica e pela síntese tecnológica de uma nova Natureza” (Danowski, Viveiros de Castro, op. cit., p. 69). São também espécies de “evangelhos do reencantamento capitalista” (ibid., p. 73) pregados por tecnobilionários que querem ser vistos como a vanguarda de uma forma ‘mais esclarecida’ de capitalismo capaz de direcionar as forças produtivas da hipermodernidade para gerar “uma ordem ecopolítica fundada no acesso universal (ou tal é a promessa) da humanidade à nova abundância material” (ibid., p. 73).

²⁶⁹ No original: “using evidence and reason to figure out how to benefit others as much as possible, and taking action on that basis.”

²⁷⁰ Dados da Wikipedia.

Elas nutrem fantasias que transmutam a *perspectiva da extinção* numa *perspectiva* (ou, como diriam seus mentores, numa oportunidade) *de crescimento infinito* e de evolução para um estágio pós-humano. À possibilidade de um *mundo-sem-nós* aberta pela emergência climática, tais fabulações parecem responder com um *nós-sem-mundo* — ainda que ‘nós’ e ‘mundo’ possuam acepções completamente distintas em cada formulação. Não é para nada fortuito o fato de que a maioria dos bilionários que estão ‘fugindo para o espaço’ estão, em sua maioria, diretamente envolvidos no desenvolvimento e/ou financiamento da Inteligência Artificial. Trata-se de uma realização demasiado literal do diagnóstico de Latour (2020) de que as atuais classes dirigentes não pretendem mais liderar, mas se refugiar fora do mundo. Por isso, a dimensão simbólica de tais narrativas só pode ser plenamente compreendida no contexto (histórico, político, cultural, afetivo) da crise ecológica. Eis mais uma face do imbricamento que viemos enfatizando ao longo deste capítulo entre a aceleração tecnológica e a aceleração climática. Essa dimensão evidencia por que, com os últimos avanços da IA, *as escatologias do Antropoceno* começam a se confundir com *as escatologias do Tecnoceno*, resultando num ‘entrelaçamento escatológico’ especialmente evidente nas narrativas sobre os *riscos existenciais em larga escala*, como argumentamos melhor adiante.

Segundo Haraway (2018), as tecnomitologias do Vale do Silício são inseparáveis das figuras cristãs da salvação. Ao analisar as figurações da tecnociência na virada do milênio, principalmente nos Estados Unidos, a autora afirma que a relação com o tempo aí presente é indissociável das histórias seculares de salvação do realismo cristão e suas figuras cheias de promessas e ameaças. Como se sabe, nas figurações do Apocalipse, as promessas de salvação andam sempre lado a lado das ameaças finais. Deste modo, para a autora, “[a] história da salvação secular depende do poder das imagens e da temporalidade das ameaças e promessas últimas para conter a heteroglossia e o fluxo de eventos”²⁷¹ (ibid., p. 10). Salvação e catástrofe são, assim, históricos “companheiros na novela da tecnociência”²⁷² (ibid., p. 8) que vivem juntos no “*cronotopo das promessas e ameaças finais*”²⁷³ (ibid., p. 41). Para Haraway, a tecnociência é “um discurso milenar sobre começos e fins, primeiras e últimas coisas, sofrimento e progresso, figura e realização” (ibid., p. 10). De uma perspectiva milenarista, afirma ela,

²⁷¹ No original: “Secular salvation history depends on the power of images and the temporality of ultimate threats and promises to contain the heteroglossia and flux of events.”

²⁷² No original: “bedfellows in the soap opera of technoscience”.

²⁷³ No original: “the chronotope of ultimate threats and promises”.

a crença no avanço do desastre é na verdade parte de uma confiança na salvação, seja ela esperada por revelações sagradas ou profanas, por meio de revolução, avanços científicos dramáticos ou arrebatamento religioso. (...). Essa promessa justifica o status sagrado dos cientistas, mesmo, ou especialmente, fora de seus domínios de expertise prática. De fato, a promessa da tecnociência é, argumentavelmente, seu principal peso social. A deslumbrante promessa sempre foi o lado oposto da pose enganosamente sóbria da racionalidade científica e do progresso moderno dentro da cultura da ausência de cultura. Se a obtenção de energia limpa ilimitada através do átomo pacífico, a inteligência artificial ultrapassando o meramente humano, um escudo impenetrável contra o inimigo interno ou externo, ou a prevenção do envelhecimento se materializa é vastamente menos importante do que viver sempre no cronotopo das promessas incríveis. Em relação a esses sonhos, a impossibilidade da materialização ordinária é intrínseca à potência da promessa. O desastre alimenta a esperança radiante e o desespero sem fundo, e eu, pelo menos, estou saciado. Pagamos caro por viver dentro do cronotopo de ameaças e promessas finais.²⁷⁴ (Haraway, 2004, p. 41).

Nesse sentido, é importante salientar que as visões e imaginários sobre o futuro promovidas pelos ‘tecnoevangelistas da IA’, importam muito mais enquanto *promessas de disrupção* do que como disrupções de fato. Afinal, é a contínua renovação de tais promessas que é fundamental para produzir os efeitos (materiais e imaginários) destas tecnologias e legitimar seus poderes — seja para salvar ou destruir a humanidade. Como observa Hong (2022, p. 378) sobre os tecnofuturos: “o que está sendo negociado aqui não são conquistas concretas ou estados finais, mas o poder performativo da própria promessa”²⁷⁵. Por isso, seja acerca do velho projeto de um colonialismo cósmico, da chegada de uma super IA que pode simultaneamente destruir a humanidade e garantir seu *upgrade* para a próxima fase do ‘jogo da vida’, ou mesmo acerca da precisão das predições algorítmicas e das expectativas de transformação a elas associadas, é importante lembrar, como bem argumenta Haraway no trecho acima, que *a potência de tais promessas depende da impossibilidade de sua realização ordinária*. É justamente essa impossibilidade que faz com que esses futuros maquínicos nunca cessem de chegar (mesmo quando enunciados como a causa derradeira do Fim), uma vez que se renovam indefinidamente na atualização performativa de suas próprias ameaças e promessas.

²⁷⁴ No original: “belief in advancing disaster is actually part of a trust in salvation, whether deliverance is expected by sacred or profane revelations, through revolution, dramatic scientific breakthroughs, or religious rapture. That promise justifies the sacred status of scientists, even, or especially, outside their domains of practical expertise. Indeed, the promise of technoscience is, arguably, its principal social weight. Dazzling promise has always been the underside of the deceptively sober pose of scientific rationality and modern progress within the culture of no culture. Whether unlimited clean energy through the peaceful atom, artificial intelligence surpassing the merely human, an impenetrable shield from the enemy within or without, or the prevention of aging ever materializes is vastly less important than always living in the time zone of amazing promises. In relation to such dreams, the impossibility of ordinary materialization is intrinsic to the potency of the promise. Disaster feeds radiant hope and bottomless despair, and I, for one, am satiated. We pay dearly for living within the chronotope of ultimate threats and promises.”

²⁷⁵ No original: “What is being traded here are not concrete achievements or end states but the performative power of the promise itself.”

Analisando uma dualidade similar àquela identificada por Haraway sobre a tecnociência, a escritora de ficção científica Ursula Le Guin em sua *Teoria da bolsa de ficção*, vaticina que “se a ficção científica é a mitologia da tecnologia moderna, então o seu mito é trágico (Le Guin, op. cit., p. 23). “Tecnologia”, ou “ciência moderna” (...), segue ela,

é uma tarefa heróica, herculeana e prometeica, concebida como triunfo, logo como tragédia em última instância. A ficção incorporando esse mito será, e sempre foi, *triumfante* (O homem conquista a terra, o espaço, alienígenas, a morte, o futuro etc.) e *trágica* (apocalipse, holocausto, antes ou agora). (ibid.)

Contudo, ainda que as figuras-gêmeas da promessa e da ameaça, da salvação e da catástrofe, do triunfo e da tragédia sejam constitutivas das mitologias tecnocientíficas, como apontam Haraway e Le Guin, e os imaginários sobre riscos existenciais associados à tecnologia e especialmente à IA sempre permearam a ficção científica (de *2001 Uma Odisseia no Espaço*, passando por *Blade Runner*, o *Exterminador do Futuro*, à recente *Westworld...*), nos parece que a chegada do *Tempo do Fim* (Anders, 2007) modifica e complexifica os sentidos dessa dupla natureza.

Nas supracitadas hipóteses sobre riscos existenciais em larga escala é impressionante, mas sintomático, como as mudanças climáticas são consideradas uma variável com menor potencial de ‘catástrofe existencial’ do que aquele desencadeado, por exemplo, pela Inteligência Artificial. Em *The Precipice*, Toby Ord (2020), parceiro de Bostrom, considera que há apenas 1 em 1.000 chances das mudanças climáticas causarem uma catástrofe existencial nos próximos 100 anos (mesmo percentual para uma guerra nuclear), comparados com 1 em 10 chances de máquinas superinteligentes o fazerem (Ord, 2020) (Figura 13).

<i>Existential catastrophe via</i>	<i>Chance within next 100 years</i>
Asteroid or comet impact	~ 1 in 1,000,000
Supervolcanic eruption	~ 1 in 10,000
Stellar explosion	~ 1 in 1,000,000,000
Total natural risk	~ 1 in 10,000
Nuclear war	~ 1 in 1,000
Climate change	~ 1 in 1,000
Other environmental damage	~ 1 in 1,000
'Naturally' arising pandemics	~ 1 in 10,000
Engineered pandemics	~ 1 in 30
Unaligned artificial intelligence	~ 1 in 10
Unforeseen anthropogenic risks	~ 1 in 30
Other anthropogenic risks	~ 1 in 50
Total anthropogenic risk	~ 1 in 6
Total existential risk	~ 1 in 6

Figura 13: Tabela do livro *The Precipice* com estimativas sobre ‘catástrofes existenciais’ para os próximos 100 anos. Fonte: Ord (2020).

Num texto intitulado *O Futuro da Humanidade* no qual desenvolve o que chama de “previsões realistas” sobre nossos cenários futuros a longo prazo, Bostrom (2009) afirma claramente que “os riscos existenciais mais graves deste século derivam de desenvolvimentos tecnológicos esperados”²⁷⁶ (p. 11), como armas nucleares, agentes patogênicos projetados, a nanotecnologia molecular e máquinas superinteligentes. Seguindo a estrutura catastrofista-salvacionista destas mitologias, “as mesmas tecnologias que irão representar estes riscos também nos ajudarão a mitigar alguns riscos”²⁷⁷ (ibid.). No mesmo texto, o autor afirma que quanto maior a escala de tempo com que se considera os cenários futuros, “menor será a probabilidade da civilização tecnológica permanecer dentro da zona que designamos por ‘a condição humana’”²⁷⁸ (p. 24). Ou seja, do ponto de vista ‘longoprazista’ de Bostrom, o século atual e os próximos serão uma fase crítica para a humanidade, mas se conseguirmos ultrapassar esse breve ‘momento de crise’, a esperança de vida da civilização se tornará extremamente elevada. Entre outros motivos porque, a estas alturas, o empreendimento colonial cósmico já estaria em pleno funcionamento, diminuindo nosso risco de extinção

²⁷⁶ No original: “The most severe existential risks of this century derive from expected technological developments.”

²⁷⁷ No original: “The same technologies that will pose these risks will also help us to mitigate some risks.”

²⁷⁸ No original: “the less likely it is that technological civilization will remain within the zone we termed “the human condition”.

enquanto espécie. Ou seja, na *big picture* da ‘virtuosa’ humanidade, mesmo eventos como o Antropoceno podem ser apenas ‘pequenos troços’ em nossa história caso um seletivo grupo de humanos (e pode-se imaginar bem qual seria) conseguisse sobreviver e reconstruir a civilização com a ajuda das máquinas.

Nesses imaginários e narrativas tecnofuturistas — talvez as verdadeiras alucinações da IA —, marcadamente eugenistas²⁷⁹, os riscos existenciais em larga escala associados à tecnologia funcionam retoricamente como espécies de ‘*metaversos do fim*’ que promovem uma forma sorradeira e cínica de negacionismo climático. Através desses metaversos, a elite tecnológica global vem sequestrando o debate não só sobre o futuro da IA, mas sobre o futuro da humanidade e do planeta, deslocando o foco da discussão e da ação política daquilo que é *preocupantemente possível* (e já violentamente presente) para o *fascinantemente improvável*; uma ironia com o saber probabilístico que orienta a racionalidade algorítmica, diga-se de passagem. Isso não significa, é claro, que não existam de fato riscos sistêmicos associados à IA, como mencionamos em diversos momentos desta tese. Contudo, como apontam Bruno, Pereira e Faltay (2023), envolver os últimos desenvolvimentos da IA

em exercícios de futurologia voltados para perigos futuros em nada plausíveis, desvia o debate e escrutínio públicos sobre os danos passados e presentes desse modelo. A diferença das futurologias apocalípticas, esses danos já são reais, produzindo efeitos que atingem principalmente populações historicamente marginalizadas: pretas, periféricas e pobres. *Focar nossa lente analítica no presente revela que as aplicações de maior impacto da IA tendem a ser aquelas que permeiam instituições e práticas de maneira muito menos visível e cujos efeitos incidem na vida das pessoas de maneira muito mais corriqueira*, seja no uso de sistemas biométricos em espaços públicos e privados ou na automatização de decisões em setores como o trabalho, crédito financeiro, segurança pública, habitação, justiça, educação etc. (p. 239, grifo nosso)

Nesse sentido, as *tecnomitológicas do Fim* criadas pelos financiadores e promotores do setor envolvendo um ‘apocalipse da IA’ visam apenas inflar os superpoderes destas máquinas através de uma temporalidade messiânica e apocalíptica, distorcendo a percepção pública dos problemas que realmente ameaçam o futuro da humanidade e do planeta, incluindo o modo como essas tecnologias têm acentuado o capitalismo predatório que as produz e sustenta e a consequente catástrofe ecológica. Como observa Malabou, “ênfaticar novos perigos tecnológicos permite a hipocrisia daqueles que estão manipulando essa mesma tecnologia.

²⁷⁹ As pesquisadoras Timnit Gebru e Émile Torres (2024) vêm descrevendo esse conjunto de ideologias, junto a outras sobre as quais são discorreremos aqui, sob o acrônimo TESCREAL (*transhumanism, Extropianism, singularitarianism, cosmism, Rationalism, Effective Altruism, and longtermism*). Elas destacam o caráter escatológico e o ‘fundo eugenista’ destas ideologias, uma vez que todas endossam o uso de tecnologias emergentes, sobretudo da IA, para ‘melhorar’ radicalmente a humanidade e criar uma nova espécie pós-humana.

Não passa de conversa de bombeiros piromaníacos.”²⁸⁰ (2019, p. 152). Por isso, lembram Danowski e Viveiros de Castro, os discursos sobre o ‘fim do mundo’ dependem sempre de perspectivas situadas que determinam quem está definindo tal fim e de que mundo, afinal, se trata.

[o] ‘fim do mundo’ só tem um sentido determinado nestes discursos — só se torna ele próprio pensável como possível —, se se determina simultaneamente *para quem* este mundo que termina é *mundo*, quem é o mundano ou o ‘mundanizado’ que *define o fim*. O mundo, em suma, é uma perspectiva objetiva. (op. cit., p. 37)

²⁸⁰ No original: “Emphasizing novel technological dangers enables the hypocrisy of those who are manipulating that very technology.”

Considerações finais

O Tempo do Fim Versus o Fim do Tempo: Portanto, por sua natureza mesma, essa era é uma “suspensão”, e nosso “modo de ser” nessa era deve ser definido como “ainda não sendo inexistentes”, “ainda não exatamente sendo inexistentes”. Assim, a questão moral básica de épocas anteriores deve ser reformulada radicalmente: ao invés de perguntar “Como devemos viver”, devemos agora perguntar “Iremos viver?”. Para nós, que somos “ainda não inexistentes” nessa Era de Suspensão, só há uma resposta: embora a qualquer momento O Tempo do Fim possa se converter n’O Fim do Tempo, devemos fazer tudo a nosso alcance para tornar O Tempo Final infundável.
Gunther Anders

Ao longo desta investigação, nos guiamos pela hipótese central, desdobrada em múltiplas camadas, de que, em virtude da centralidade do conhecimento e da intervenção sobre a dimensão temporal para a racionalidade algorítmica — um *saber-poder preditivo* que tem como alvo privilegiado “o que poderia advir” (Rouvroy, Berns, 2015, p. 54) —, os algoritmos da IA podem ser considerados *tecnologias de governo do tempo*. Seus efeitos mais poderosos, marcadamente performativos, residem no modo com que vêm, progressivamente, modulando nossos *campos de possibilidades*. Seja na regulação personalizada do espectro visível e legível nas plataformas digitais, na automatização de interpretações e decisões nas mais variadas esferas, nas políticas da classificação (Crawford, 2021) que ordenam os arquivos de saber onde as IAs são treinadas, nos mundos sintéticos e cada vez mais abundantes fabricados pelos modelos generativos, vem sendo gestada uma *algoritmização do tempo* que, como vimos, caracteriza-se principalmente pela *redução das possibilidades às probabilidades* via lógicas estatísticas e preditivas.

Essa *modulação do possível através das malhas do provável* que o governo algorítmico introduz e através do qual produz seus efeitos mais poderosos está atada a uma gestão calculada da *incerteza* ou do *acaso* que visa, em grande medida, controlar capturar uma indeterminação que é constitutiva das potências de individuação dos mais variados modos de existência no tempo (Cf. Simondon, 2009, Barad, 2007, 2015, 2020), do ser à mais ínfima matéria. Como argumenta Karen Barad, de uma perspectiva filosófica alinhada à física quântica, há uma “indeterminação ontológica, uma abertura radical, uma infinidade de possibilidades”²⁸¹ (Barad, 2015, p. 401) que habita o cerne mais ínfimo da matéria e que é a

²⁸¹ No original: “Ontological indeterminacy, a radical openness, an infinity of possibilities”.

condição para a possibilidade de todas as estruturas em suas reconfigurações dinâmicas de in/estabilidade. Para a autora, “a matéria, em sua materialização iterativa/recursiva, é um jogo dinâmico de in/determinação”²⁸² (ibid.). Nesse entender,

[m]esmo as menores partículas de matéria são uma enorme multidão. Cada ‘indivíduo’ é composto de possíveis histórias de intra-ações virtuais com todos os outros. A indeterminação é um desfazer/refazer da identidade que abala as próprias fundações do não/ser²⁸³ (Barad, 2020).

Para Rouvroy, é precisamente essa *in/determinação* radical e ontológica do ser apontada por Barad que o capitalismo digital e o governo algorítmico neutralizam e expropriam (Rouvroy, Athanasiadou, Klumbyté, 2022). O diagnóstico da autora se aproxima daquilo que Bruno (2020) chamou de *sequestro do futuro* ao refletir sobre os efeitos das previsões e ações em tempo real sobre os comportamentos e arquiteturas de escolha que marcam a condução de condutas nas plataformas digitais. Ainda que nem todos os seus usos estejam voltados para este fim, o que as *tecnologias algorítmicas de governo do tempo* materializadas na IA capturam e expropriam (seguidamente transformando em mais-valia) são, em última instância, as potências de individuação, transformação e abertura que compõem as virtualidades disso que usualmente se chama de futuro — e que ao longo deste trabalho privilegiamos chamar de possível. Essa *reserva de devir* a partir da qual os processos de individuação se alimentam e se desdobram (Simondon, 2009), modulando o que somos e, principalmente, *o que podemos vir a ser*. Tal instância, como aponta Rouvroy, deveria ser uma reserva para a imaginação, criação, fabulação coletiva, um espaço de abertura ao que ‘não está mais’ ou ‘não está presente ainda’, um local de investimento em prol do comum ao invés de consumo (Rouvroy, Athanasiadou, Klumbyté, op. cit.). Simultaneamente, esse *controle ou sequestro da indeterminação, do futuro ou da contingência* — como se queira chamá-lo, afinal — aponta para uma das tensões que subjaz aos mundos gerados e geridos pela datificação e algoritmização, que discutimos no capítulo 4 e que retomamos brevemente adiante.

Antes disso, contudo, cabe pontuar novamente que a regulação do possível, esse “conjunto de ações sobre ações possíveis” (Foucault, 1995, p. 243), como demonstrou Foucault, é inerente ao próprio exercício do poder. Ou seja, *todo governo é, no limite, um ‘governo do possível’*.

²⁸² No original: “Matter in its iterative materialization is a dynamic play of in/determinacy”.

²⁸³ “even the smallest bits of matter are an enormous multitude. Each ‘individual’ is made up of all possible histories of virtual intra-actions with all Others. Indeterminacy is an un/doing of identity that unsettles the very foundations of non/being.”

A novidade do conjunto de fenômenos e objetos alvo de nossas indagações (e tentativas de respostas) — ou, nos termos caros ao filósofo, a *ruptura histórica* — é que os algoritmos se convertam nos agentes centrais desse governo do possível na contemporaneidade: de nossas oportunidades, decisões, interpretações, visualidades, textualidades, subjetividades, etc.

Esse modelo de racionalidade ‘algoritmocêntrico’ no qual os algoritmos ocupam lugar privilegiado sobre os modos de conhecer, gerir, decidir e intervir sobre a realidade e os sujeitos, como vimos no Capítulo 1, integra um amplo e profundo deslocamento epistêmico, iniciado no pós-guerra, marcado pela passagem do domínio de um *modelo de racionalidade iluminista baseado na reflexividade crítica* para um *modelo de racionalidade baseado em regras algorítmicas* (Bruno, 2021; Daston, 2022; Halpern, 2022). Um modelo de racionalidade que encontra-se intimamente ligado à *matriz da cibernética*, seja em suas premissas epistemológicas ou na estrutura temporal recursiva que introduz, e que atravessou de diversos modos nossa investigação. Essa passagem histórica poderia ser sintetizada como uma inflexão do predomínio de uma *razão reflexiva* para uma *razão sintética* e já sinalizava a *perda do excepcionalismo humano* que as atuais máquinas da IA tornam patente. Disso decorrem várias consequências, como o fato desse modelo de racionalidade, ancorado numa razão sintética que se afasta da subjetividade reflexiva moderna e, portanto, da *consciência*, operar, primordialmente, por meio de *processos cognitivos não-conscientes* (Hayles, 2017). Como vimos, é em grande medida por meio de uma espécie de ‘sincronização’ ou ressonância ao nível do cognitivo não-consciente que os agenciamentos maquínico-humanos que compõem os atuais ecossistemas digitais, materializados principalmente em usuários e algoritmos, operam e produzem seus principais efeitos. Inspirando-nos nas proposições de Hayles (op. cit.) sugerimos que os algoritmos são *cognoscentes técnicos* que performam processos cognitivos não-conscientes. Nos algoritmos baseados em aprendizado de máquina, herdeiros da vertente conexionista da IA, esses processos são, primordialmente, o reconhecimento de padrões e correlações por meio de métodos indutivos de raciocínio.

Nesse sentido, reforçando uma das questões que propusemos ao refletir sobre os deslocamentos epistemológicos em jogo com a pervasividade dos sistemas algorítmicos sobre nossos modos de saber e de ser, talvez um dos principais questionamentos não seja *quando as máquinas se tornarão conscientes* — evento que povoa há décadas os imaginários sobre a IA, sempre anunciado como iminente, mas que mesmo no atual e impressionante estágio de desenvolvimento dessas máquinas, segue longe de concretizar-se —, mas quais as implicações (ontológicas, epistemológicas, políticas, culturais, subjetivas, etc) da crescente

delegação e automação cognitivas em distintas esferas a um *tipo de racionalidade que não se funda na consciência*. Refletir sobre a algoritmização da vida a partir da perspectiva de uma mudança de modelo de racionalidade dominante, do ponto de vista teórico-metodológico, também nos permite atentar para deslocamentos históricos mais amplos e profundos, evitando a armadilha de pensar tais transformações somente segundo a temporalidade da otimização e da inovação que as apresenta sempre como a ‘grande e derradeira disrupção’ da vez.

Além disso, a intensa interpenetração, recursividade e distribuição de processos cognitivos entre agentes técnicos e humanos que marca nossa atual *ecologia cognitiva planetária* (Hayles, 2017) é uma ocasião propícia para questionar o que se torna a cognição em tais contextos sociotécnicos sem reduzi-la à oposição ou competição entre humanos e máquinas (Bruno, Pereira e Faltay, 2023) ou num retorno à defesa (indefensável para os problemas próprios do nosso tempo) do excepcionalismo humano. Essa lógica competitiva atravessou as últimas décadas de desenvolvimento da IA, fomentando desde confrontos públicos²⁸⁴ até os imaginários sobre a Singularidade Tecnológica e o ‘apocalipse da IA’ — temas que exploramos no capítulo dedicado aos futuros maquínicos. Narrativas que permanecem tributárias da fórmula ficcional que Le Guin (2021) chama de *história como batalha*.

E se, como propõem Bruno, Pereira e Faltay (op. cit., p. 237), “ao invés de nutrir os fascínios e temores que a lógica competitiva suscita, aproveitarmos o ‘abalo existencial’ desencadeado pela perda de nosso excepcionalismo ontológico e epistemológico para começar a pensar a cognição ou mesmo a inteligência a partir da *relacionalidade*, da *coprodução* e da *interpelação mútua entre agentes heterogêneos?*” (p. 237). Recolocar o problema a partir dessa perspectiva, sugerem, inspirados em proposições como as de Barad (2007) e Haraway (2023), é uma aposta numa *ontologia relacional*. Assim, não se trata somente de compreender a cognição e a inteligência como processos relacionais, “mas também de entender que a relacionalidade modifica os próprios entes nela envolvidos” (ibid.). Como observa há muito Haraway, atenta aos hibridismos e aos *emaranhados fazedores de mundos*: “os parceiros não precedem o encontro”²⁸⁵ (2008, p. 4).

Dessa forma, mais do que a substituição do humano pela máquina — embora refletir à sério sobre as consequências sociais da intensa automação em curso também passa,

²⁸⁴ Como as antológicas vitórias do Deep Blue sobre Kasparov no xadrez em 1996 e do AlphaGo sobre Lee Sedol no Go em 2016.

²⁸⁵ No original: “The partners do not precede the meeting”.

indiscutivelmente, por esta questão —, uma das questões cruciais da crescente presença e importância dos algoritmos da IA nas ecologias cognitivas contemporâneas é “como essa relação máquina-humana, que tende a ser cada vez mais íntima e complexa, transforma o que somos e o que historicamente se entende como o humano” (Bruno, Pereira e Faltay, op. cit. p. 238). Nesse sentido, é fundamental compreender que o humano em relação ao qual essas máquinas são ora espelho, ora antítese (a depender da posição do interpretante) não está nunca dado de antemão, mas é ele mesmo continuamente *feito e refeito* ao longo da história “das ferramentas e tecnologias com as quais concebemos, agimos e transformamos o mundo e a nós mesmas” (ibid., p. 238).

Contudo, um olhar minimamente crítico ao *status quo* das tecnologias algorítmicas verá como as potencialidades dessa ontologia relacional — a partir das quais poderíamos *devir-com* (Haraway, 2008, 2016) máquinas que não sejam *nem senhoras nem servas*, mas, *companheiras* (Beiguelman, 2023) com quem o encontro resulte em *alteridades significativas* (Haraway, 2008) — encontram-se quase totalmente esvaziadas no atual modelo hegemônico de desenvolvimento da IA. Monopolizado pelos interesses comerciais das gigantes da tecnologia e sustentado por um *extrativismo multiescalar* de recursos materiais e humanos que alimenta o voraz metabolismo do modelo ‘AI driven’ (simultaneamente econômico, epistemológico e político), esse modelo segue movido (e de modo cada vez mais acelerado) por aquilo que Le Guin (op. cit., p. 23) chama de “flecha-(assassina)-do-Tempo do Tecno-Heróico”.

É justamente essa matriz hegemônica *preditivo-aceleracionista* da IA — baseada na produção de previsões estatísticas via aprendizado de máquina e movida pela otimização ininterrupta e pela crença de que ‘*big is better*’ (Varoquaux, Luccioni, Whittaker, 2024) — que está estrategicamente interessada nos usos dessas ferramentas como tecnologias de governo do tempo, em reduzir o possível ao provável, em tornar as trajetórias dos sujeitos e fenômenos mais previsíveis e em sequestrar o futuro e a contingência para transformá-los num território especulativo de bens negociáveis (Bruno, 2020; Hong, Szpunar, 2019). Como vimos especialmente no estudo de caso do TikTok, trata-se também de um paradigma tecnológico marcado pela infiltração das lógicas da financeirização não só nos modelos de negócios das plataformas digitais, mas em seus regimes de temporalidade. Esse espraiamento vem produzindo aquilo que designamos de *ambientes digitais de alta frequência*: ambientes dominados por temporalidades fragmentadas e aceleradas que maximizam ainda mais o extrativismo que sustenta essas arquiteturas algorítmicas.

Sugerimos no último capítulo que as práticas e imaginários da Inteligência Artificial na contemporaneidade vêm produzindo uma modalidade específica de futuro histórico, que propomos chamar de *futuros maquínicos*. Simultaneamente materiais e simbólicos (manifestos tanto objetivamente nas previsões algorítmicas quanto nos discursos, promessas, imaginários e ideologias que permeiam o campo da IA e também produzem os efeitos performativos destas tecnologias), os futuros maquínicos agenciados pela IA integram um projeto no qual a tecnologia se converte na ferramenta e cronotopo privilegiado para produzir e imaginar o futuro, seja em seu potencial de *salvação* ou *destruição*.

Como enfatizamos, esses futuros maquínicos são marcados por (ou mesmo se sustentam através de) uma série de tensões, contradições e ambiguidades. Num primeiro nível porque podem ser considerados correlatos de uma *crise do futuro* na esfera social (Koselleck, 2006; Fisher, 2019, 2022; Berardi, 2019). Sintetizamos essa crise como o deslocamento da crença em um *futuro como projeto utópico*, que marcou o século XX, para um *futuro como programa maquínico* no século XXI, que tem como agentes privilegiados os algoritmos da IA.

Essas tensões, contradições e ambiguidades também atravessam as próprias dimensões em que os futuros maquínicos se manifestam. Por um lado, a lógica programática da previsão algorítmica — baseada no reconhecimento de padrões, em correlações, probabilidades e na estrutura temporal da recursividade — transforma o futuro numa *repetição recursiva* que dá continuidade e amplifica (ao mesmo tempo em que despolitiza e neutraliza) os padrões de desigualdade históricos datificados nos modelos algorítmicos. Deste ponto de vista, esses sistemas são aquilo que Oliveira (2023) chamou de “oráculos do passado”, tornando improvável a ruptura com tais padrões históricos. Ao nível cultural, o excesso de espirais de reforço dos mesmos padrões nos ecossistemas digitais tem produzido uma *IA-formização* (Mann, 2024) que já é claramente perceptível, por exemplo, nos *outputs* dos modelos generativos de texto e imagem e nas redes sociais de forma mais ampla. Essa homoneização maquínica em larga escala — que, importante enfatizar, responde muito bem às demandas por produção incessante de conteúdo que o modelo de negócios das plataformas digitais requer — indica que a perda de diversidade em nossos ecossistemas midiáticos deveria preocupar mais do que uma purista e anacrônica preocupação pela perda da originalidade.

Ao mesmo tempo, observamos como subjaz ao programa e à própria tecnopolítica das máquinas computacionais baseadas na recursividade uma *tensão entre repetição e*

indeterminação. Isso porque, numa via, a indeterminação é constitutiva de seus modos de existência e o que lhes garante abertura para o contínuo aperfeiçoamento (Simondon, 2008; Hui, 2019), fundando o que Fazi (2018) chama de uma *ontologia da contingência*. Noutra via, os mundo gerados e geridos pela datificação e algoritmização operam tanto uma redução da indeterminação, como argumentamos no começo destas considerações, quanto excluem (ontológica, epistemológica e politicamente) aquilo que não pode ser computável por esse regime de conhecimento (Mbembe, 2019b; Hong, 2022). Disso deriva o paradoxo de que estas máquinas cibernéticas preditivas *operam através do indeterminado e do incalculável e simultaneamente os reduzem* dos fluxos de automatização nos quais são inseridas.

Já ao nível dos discursos e imaginários sobre a Inteligência Artificial, promovidos principalmente pelos ‘tecnoboys do Vale’ e expressos em narrativas e ideologias como a Singularidade Tecnológica, os riscos existenciais em larga escala, o aceleracionismo, o transhumanismo, o longoprazismo e o altruísmo eficaz (Cf. Gebru, Torres, 2024), o futuro se manifesta como uma contínua *promessa de ruptura*. Tais discursos e imaginários, fortemente potencializados após o ‘abalo existencial’ desencadeado pela chegada do ChatGPT, exaltam e produzem os ‘superpoderes da IA’ através de uma temporalidade messiânica e apocalíptica, seja para a salvação da humanidade ou para sua derradeira catástrofe. Apesar de autoras como Haraway (2004) e Le Guin (2021) enfatizarem que a ambivalência das promessas e ameaças são constitutivas das narrativas da tecnociência, argumentamos que nosso tempo, marcado pelo cruzamento da aceleração tecnológica e climática (Viveiros de Castro, 2023), faz proliferar as *mitologias do fim* (Danowski, Viveiros de Castro, 2017) relacionadas a esses dois domínios. É nesse contexto cultural, histórico e político que ganham tração os imaginários sobre um suposto ‘*Apocalipse da IA*’, segundo os quais corremos grandes riscos existenciais enquanto espécie de sermos destruídos (mas também transmutados em pós-humanos) por máquinas superinteligentes e superpoderosas. *Tecnomitologias* que manifestam uma forma sorrateira de negacionismo sobre as verdadeiras origens de nosso colapso civilizacional.

Supor que é mais provável que nosso Fim (assim como nosso recomeço) será fruto da agência isolada de máquinas superpoderosas e não das consequências do capitalismo fóssil e predatório através do qual o ideal de progresso da civilização moderna ocidental se desenvolveu — e que, claro, é a condição de possibilidade para a produção dessas mesmas

máquinas — é como dizer, parafraseando amargamente o lema que animou os movimentos sociais no começo do século, que ‘*um outro fim do mundo é possível*’. Nesse sentido, poderíamos dizer que os futuros maquínicos agenciados pela IA produzem um *duplo sequestro do futuro*: através da performatividade e dos estreitamentos do possível que as predições algorítmicas produzem e através dos imaginários e narrativas sobre o futuro que seus ‘tecnoevangelistas’ promovem, regulando também nesse sentido nossos campos de possibilidade e, sobretudo, nosso senso de urgência diante dos problemas do presente. Talvez o realismo capitalista de Fisher baseado na fórmula de que “é mais fácil imaginar o fim do mundo do que o fim do capitalismo” (Fisher, 2019) esteja se convertendo num realismo tecnocrático (não menos capitalista, é claro) no qual, dado o ‘andar distópico da carruagem’, *é impossível não imaginar o fim do mundo*, mas nunca pelos motivos que estão realmente nos levando ao colapso e cujo enfrentamento requer a invenção, regeneração, negociação e sustentação de modos menos arrogantes de (co)existência — dos quais a tecnologia pode fazer parte caso sejamos capazes de concebê-la enquanto uma companheira entre outras ao invés de uma arma para a destruição ou para a salvação.

Como pondera com razão Viveiros de Castro (2023, p. 17-18): “*as máquinas de IA não virão nos salvar [e eu acrescentaria, ou nos destruir]; mas tampouco irão embora*”. O que nos resta então, caso nos recusemos a habitar o cronotopo das ameaças e promessas finais (Haraway, 2004) e a assumir as posturas invertidas gêmeas do *tecnosolucionismo* ou do *tecnocatastrofismo*? A tarefa de elaborar respostas consistentes para essa questão certamente transcende os limites dessa tese. Nosso foco nessa investigação esteve muito mais centrado num certo diagnóstico sobre um modelo dominante de IA (mais especificamente, em compreender sobretudo seus aspectos epistemológicos e cronopolíticos) do que em traçar linhas tecnopolíticas de fuga. Trata-se, contudo, de uma pergunta-desafio que vem se tornado inevitável nos espaços de reflexão crítica sobre a tecnologia e que gostaríamos de deixar reverberando. Assim, esboçamos a seguir algumas breves pistas que podem apontar respostas parciais e provisórias para essa questão e nos lembrar que aquilo que vem sendo apresentado como o destino inevitável da tecnologia, é — novamente ele — contingente.

Um primeiro aspecto consiste na aposta em projetos de *tecnologias situadas* (Haraway, 2016), que reconheçam suas perspectivas parciais, ancoradas a contextos, problemas e conhecimentos específicos e localizados. Tecnologias derivadas de debate público e de práticas colaborativas, orientadas pela *tecnodiversidade* (Hui, 2020), pelo *bem comum*

(Coding Rights, 2024) e por uma *ética para o mundo majoritário*²⁸⁶ (Ricaurte, 2022). Um contraponto à abordagem ‘*one-model-for-everything*’ (Varon, Costanza-Chock, Gebru, 2024) e ao suposto de que ‘*big is better*’ (Varoquaux, Luccioni, Whittaker, 2024) hoje materializados sobretudo nos grandes modelos fundacionais de IA Generativa monopolizados pelas Big Techs. Como assinala um documento elaborado a partir do evento *IA para o bem comum: fortalecendo sistemas alternativos*²⁸⁷, “a ideia de um modelo tecnológico universal que funciona para todos é uma falácia, um argumento fictício que alimenta uma monocultura tecnológica cultivada em modelos de monopólio” (Coding Rights, 2024). No documento também sugere-se que: “a tecnologia precisa ser distribuída, para que as pessoas que a desenvolvem possam permanecer conectadas às suas raízes e trabalhar em questões que compreendem e com padrões comunitários”. Para tanto, é preciso questionar localmente a partir de debates e negociações coletivas e multidisciplinares, sobretudo quando se tratam de algoritmos de uso público: qual a finalidade de um sistema algorítmico específico? Que problema ele visa resolver? Que outras alternativas podem ser consideradas para tal fim? Quais seus possíveis impactos (sociais, culturais, ambientais)? Com que atores, dados, recursos minerais, energéticos, financeiros, humanos, etc, ele será desenvolvido? É fundamental lembrar também, como vêm pautando principalmente os movimentos pelo banimento do reconhecimento facial no Brasil²⁸⁸ e no mundo, que “*nem toda tecnologia deve existir*”. Se uma tecnologia deve ou não ser desenvolvida é uma discussão que precisa ser feita.” (Coding Rights, op. cit., grifo do autor). Contígua a essa discussão e mais alinhada às questões que mobilizaram essa tese, é importante que se pondere nesses debates que parcela do incalculável, do imprevisível, do imponderável que compõe inerentemente a realidade deve assim permanecer. Ao mesmo tempo, como fazer com que tais dimensões não percam agência diante da voracidade da racionalidade algorítmica e do modo privilegiado com que simultaneamente cria e exclui mundos é um problema com o qual permanecer.

²⁸⁶ Ricaurte utiliza o termo “mundo majoritário” para referir-se aos países onde reside a maior parte da população. O termo foi cunhado pelo fotojornalista bengali Shahidul Alam como uma alternativa a termos como “terceiro mundo” ou “países em desenvolvimento” e que expressam, segundo ela, uma visão linear e ocidental de desenvolvimento, que é imposta ao resto do mundo.

²⁸⁷ O evento foi realizado em junho de 2024 no Rio de Janeiro e contou com a presença de várias organizações e lideranças relacionadas à tecnopolítica, entre elas, Timnit Gebru. Para saber mais sobre o evento e as recomendações elaboradas a partir dele, ver: <https://codingrights.org/library-item/ai-e-outras-taticas-para-fomentar-ecossistemas-locais-descentralizados-para-o-desenvolvimento-de-tecnologia-situada/>

²⁸⁸ Para uma sistematização das iniciativas no Brasil, ver: <https://lavits.org/iniciativas-que-apoiam-o-banimento-do-reconhecimento-facial-avancam-no-brasil/>

Os estudos feministas sobre a objetividade na ciência e tecnologia nos ensinam há muito que todo saber é necessariamente um saber parcial, uma perspectiva *desde algum lugar* (Haraway, 1995). O que poderia desdobrar-se na máxima de que *toda tecnologia é uma tecnologia parcial, uma perspectiva desde algum lugar*. Como mostra Haraway, as pretensões totalizantes, onipotentes e transcendentais só foram permitidas historicamente àqueles que podem ver desde lugar nenhum e “cujo Olho produz, apropria e ordena toda a diferença” (ibid., p. 27). Segundo a autora, conhecimentos não localizáveis são também aqueles que não são convocados a prestar contas sobre suas responsabilidades. Contudo, como também alerta ela, referindo-se à ciência mas totalmente aplicável à tecnologia: é preciso ir além tanto da mera denúncia do viés quanto de uma romantização e/ou apropriação das perspectivas periféricas e ditas subjugadas pautadas por visões relativistas e construcionistas sobre a realidade. “As perspectivas dos subjugados não são posições ‘inocentes’. Ao contrário, elas são preferidas porque, em princípio, são as que têm menor probabilidade de permitir a negação do núcleo crítico e interpretativo de todo conhecimento” (ibid., p. 23), porque “parecem prometer explicações mais adequadas, firmes, objetivas, transformadoras do mundo” (ibid.). Mas, ‘ver desde baixo’ requer tanta habilidade quanto ‘ver desde cima’. Assim, não basta mostrar que ‘tudo é socialmente construído’, logo, enviesado. É preciso, como ela propõe, insistir e produzir *explicações melhores* sobre “um mundo que possa ser parcialmente compartilhado e amistoso em relação a projetos terrestres de liberdade finita, abundância material adequada, sofrimento reduzido e felicidade limitada” (ibid., p. 16). Tais explicações melhores podem incluir aquelas elaboradas a partir de dados e algoritmos, desde que vinculadas a outra tecnopolítica que não a que hoje monopoliza o desenvolvimento tecnológico.

Na busca desejante por esse mundo parcialmente compartilhado, é fundamental exaltar o papel e o potencial das *práticas e poéticas artísticas* como exercícios privilegiados de tecnologias situadas, para a criação de cosmotécnicas (Hui, 2020) que tensionem as lógicas, projetos e valores hegemônicos, assim como para a produção de outros imaginários sobre a tecnologia e sobre os futuros maquínicos, derivando em outros modos de individuação com/através da tecnologia. Esses territórios artísticos e experimentais oferecem condições férteis para tensionar e refazer os enredamentos entre humanos, máquinas e seus futuros compartilhados. Trata-se de um âmbito em que a possibilidade de práticas de companheirismo com as máquinas (Beiguelman, 2021), de conexões inventivas e parciais com alteridades significativas (Haraway, 2008, 2016), mais protegidas da profunda assimetria

de poder que marca essas conexões em nossas práticas digitais cotidianas, têm maior probabilidade de prosperar — ainda que não haja garantias.

O último conjunto de pistas propositivas que gostaríamos de destacar reside na necessária construção de uma *outra cronopolítica para a IA*. Esse esforço pode se desdobrar em múltiplas direções, mas enfatizamos aqui duas. A primeira, mais evidente, mas não por isso mais simples de ser posta em prática: trata-se da necessidade de desenvolver tecnologias que não estejam pautadas pelos imperativos do *regime de temporalidade aceleracionista*. Esse regime vem ditando não somente a velocidade exponencial com que novos modelos, funcionalidades e empresas são lançados e rapidamente superados no mercado (onde predominam as métricas como única medida de relevância), mas também transforma os ecossistemas algorítmicos naquilo que chamamos de *ambientes digitais de alta frequência*.

Ao nível da produção de subjetividades, esse regime inviabiliza, ou ao menos evita intensamente, a *hesitação*, a *dúvida*, a *ambiguidade*, a *reflexividade*. Essas atitudes dependem, justamente, de uma pausa, uma brecha, um intervalo, em suma, de uma suspensão dos fluxos da automatização acelerada para que possam surgir. Isabelle Stengers, ao defender a desaceleração como condição para a construção coletiva de *outra ciência*, alerta que:

a rapidez requer e produz insensibilidade a tudo o que poderia desacelerar: fricções, atritos, hesitações que fazem sentir que não estamos sozinhos no mundo. Desacelerar significa nos tornarmos novamente capazes de aprender, de encontrar e de reconhecer o que nos une e nos mantém unidos, de pensar, imaginar e, no mesmo processo, criar junto a outros, vínculos que não sejam de captura. (Stengers, 2023, p. 117-118)

Abrir linhas de fuga para novos possíveis da tecnologia passa também por criarmos, como sugere Bruno (2023) inspirada por Stengers, *capacidades de hesitar juntos* — uma postura que se contrapõe radicalmente aos imperativos de otimização acelerada do tecnosolucionismo. Além disso, como apontamos no Capítulo 4, a incapacidade de imaginar futuros alternativos na contemporaneidade passa também pela aceleração do tempo causada pela infiltração da velocidade da infoesfera capitalista sobre a piscoesfera (Berardi, 2019, 2020; Celis Bueno, Schultz, 2021;). Ou seja, para imaginarmos, criarmos e fabularmos futuros alternativos genuinamente novos será necessário também habitar outras temporalidades.

Uma segunda proposição cronopolítica pode ser sintetizada na ideia de *elaborar outros usos do passado por meio das IAs*, explorando abordagens que não estejam alinhadas à simples automatização e reprodução acrítica dos passados datificados nos modelos algorítmicos. Se as

IAs são “oráculos do passado” (Oliveira, op. cit.), por que não investir em suas potencialidades não para prever futuros fadados à repetição, mas como *ferramentas de diagnóstico* que revelem padrões históricos e epistemológicos inalcançáveis à percepção humana e aos métodos tradicionais das ciências? Afinal, como ficou evidente com as novas formas de negacionismo e revisionismo histórico que emergiram com a extrema direita nos últimos anos (e para os quais, inclusive, a IA passa a ser uma ferramenta muito útil), *o passado está permanentemente em disputa tanto quanto o futuro*. Ou ainda, como sugere Chun, que utilizemos mais as previsões algorítmicas como a Ciência do Clima: “não para que aceitemos fatalisticamente o futuro que eles prevêm, mas sim para que façamos tudo o que for necessário para evitar que esse futuro ocorra” (Chun, 2021, p. 257).

Contudo, para realizar a tarefa ético-política da nossa geração — ou seja, para fazer com que “O Tempo do Fim” não se transforme n“O Fim do Tempo” (Anders,2013), para “adiar o fim do mundo” (Krenak, 2019) — será preciso, como propõe Alyne Costa (no prelo), ir além das meras previsões. É preciso, defende ela, “*imprever o mundo*”, ou, mais precisamente, “*imprever o fim do mundo*” (ibid.). Ao refletir sobre o colapso ecológico e todas as outras falhas que se acumulam a ele, a autora reconhece a importância da ciência climática para antecipar o futuro via previsões, mas afirma que elas são insuficientes. “Primeiro porque nós nunca poderemos antecipar tudo; se esperarmos saber mais antes de agir, será tarde demais. E segundo porque, para ser capaz de prevenir uma catástrofe, é preciso, antes, que *acreditemos* na possibilidade de que ela ocorra.” (ibid.). Para isso, ela advoga a importância de um outro tipo de verdade, uma que vá além da mera extrapolação de tendências do presente para o futuro; uma verdade que, ironicamente, pode ser encontrada na *ficção* e em seu potencial para imprever o mundo. Nas palavras da autora:

[t]alvez nossas melhores chances residam em apelar para um outro tipo de verdade, cuja solidez provém da possibilidade mesma de que as coisas possam ser diferentes: aquela que é contada via o caminho tortuoso e peculiar da ficção, que associa o mundo ao inesperado e imprevisível, que ensina a acreditar naquilo cuja efetuação não podemos prever e que justamente, por esse acreditar, nos torna capazes mesmo de imprever. Trata-se, portanto, de uma verdade que *não mais se opõe à imaginação*, mas se alia a ela para explorar âmbitos até então não visitados da realidade. (Costa, no prelo).

Diante das catástrofes que se acumulam, afirma ela, “não temos outra alternativa que não *reabilitar o imprevisível*” (ibid., grifo nosso). Precisamos “apostar no inaudito contra a certeza de que nada novo pode se passar que nutre tanto o negacionismo quanto o derrotismo” e dar lugar a “histórias provenientes de mundos que nunca tiveram a audácia de decretar a realidade como previsível” (ibid.). Nesse sentido, acreditar no mundo nesse

momento histórico *fraturado e fraturante*, assim como tornar “*O Tempo Final infundável*” (Anders, 2013, grifo nosso), exigirá de nós — os “ainda não exatamente inexistentes” (ibid.) — a capacidade de simultaneamente reconhecer a catástrofe e *apostar no imprevisível* e (complementando Costa) *no improvável*. Manter abertas estas brechas cronopolíticas, que se tornam cada vez menos sensíveis, imagináveis ou mesmo práticas sob o domínio da racionalidade algorítmica no capitalismo digital, é fundamental para dar agência aos possíveis que o habitam e que escapam das previsões probabilísticas e das demandas por certezas que esse modelo de mundo impõem.

Referências

- ABID, Abubakar; FAROOQI, Maheen; ZOU, James. Large language models associate Muslims with violence. *Nat Mach Intell*, vol. 3, pp. 461–463, junho de 2021.
- ABIDIN, Crystal. Mapeando celebridades da Internet no TikTok: Explorando Economias da Atenção e Trabalhos de Visibilidade. *Pauta Geral*, v. 8, n. 2, pp. 1-50, 2021.
- AGAMBEN, Giorgio. Por uma teoria do poder destituente. [Texto traduzido a partir de palestra pública proferida em Atenas]. Tradução Luhuna Carvalho. *Revista Punkto*, ed. 7, maio 2015. Disponível em: <https://www.revistapunkto.com/2015/05/por-uma-teoria-da-potencia-destituente.html>. Acesso em: 06 ago. 2024.
- AIROLDI, Massimo. *Machine Habitus: Towards a sociology of algorithms*. Cambridge: Polity Press, 2022.
- ALVES, André; LIEDKE, Lucas. Empuxo à nostalgia: o que a insistência pelo saudosismo tem a dizer sobre o nosso tempo?. *Vibes*, 14 jun. 2023. Disponível em: <https://floatvibes.substack.com/p/vibes-empuxo-a-nostalgia>. Acesso em: 3 maio 2024.
- ALTMAN, Sam. *Moore's Law for Everything*. 16 mar. 2021. Disponível em: <https://moores.samaltman.com/>. Acesso em: 6 maio 2024.
- AMADEU, Sérgio; SOUZA, Joyce; CASSINO, João F. *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal*. São Paulo: Autonomia Literária, 2021.
- AMOORE, Louise. Introduction: Thinking with Algorithms: Cognition and Computation in the Work of Katherine Hayles. *Theory, Culture and Society*, vol. 36, n. 2, pp. 1-14, 2019.
- AMOORE, Louise; PIOTUKH, Volha. Interview with N. Katherine Hayles. *Theory, Culture & Society*, v. 36, n. 2, pp. 145–155, 2019.
- AMOORE, Louise. *The politics of possibility: risk and security beyond probability*. Durham: Duke University Press, 2013.
- ANDERS, Günther. *Le temps de la fin*. Paris: L'Herme, 2007.
- _____. Teses para a Era Atômica. *Sopro Panfleto Político-Cultural*, n. 87, abril 2013 [1962]. Disponível em: <https://culturaebarbarie.org/sopro/outros/anders.html>. Acesso em: 7 maio 2024.
- ANDERSON, Chris. The End of Theory: the data deluge makes the scientific method obsolete. *Wired*, 23 jun. 2008. Disponível em: <https://www.wired.com/2008/06/pb-theory/>. Acesso em: 4 fev. 2022.
- ANDERSON, Ian; GIL, Santiago; GIBSON, Clay; WOLF, Scott; SHAPIRO, Will; SERERCI, Oguz; GREENBERG, David. ‘Just The Way You Are’: Music Listening and Personality. *Spotify R&D Research*, 15 dez. 2020. Disponível em:

- <https://research.atspotify.com/just-the-way-you-are-music-listening-and-personality/>. Acesso em: 8 fev. 2022.
- ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. Machine Bias. *Pro Publica*, 23 maio 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 8 fev. 2022.
- APPADURAI, Arjun. *Banking on words: the failure of language in the age of derivative finance*. Chicago: The University of Chicago Press, 2016.
- APPRICH, Clemens. Introduction. In: APPRICH, Clemens; CHUN, Wendy; CRAMER, Florian; STEYERL, Hito. *Pattern Discrimination*. Minneapolis, Lüneburg: University of Minnesota Press, meson press, 2018.
- ARADAU, Claudia.; BLANKE, Thomas. *Algorithmic reason: the new government of self and other*. Oxford, New York: Oxford University Press, 2022.
- ARAUJO, Willian F.; KARHAWI, Issaaf. “Todo mundo pode ser famoso com o algoritmo do TikTok”: imaginários e saberes sobre eficiência algorítmica e potência viral. Anais do 46º Congresso Brasileiro de Ciências da Comunicação, setembro de 2023. Disponível em: https://sistemas.intercom.org.br/pdf/link_aceite/nacional/11/0816202323241564dd84cf04065.pdf. Acesso em: 25 ago. 2024.
- ARBIX, Glauco. IA: medo das máquinas ou dos desacertos humanos? *Revista Brasileira*, fase X, ano II, n. 115, abril/maio/junho, 2023, pp. 8-15.
- ARIELY, Dan. *Previsivelmente Irracional: as forças invisíveis que nos levam a tomar decisões erradas*. Rio de Janeiro: Editora Sextante, 2020.
- AUSTIN, John. *How to do things with words*. Oxford: Clarendon Press, 1962.
- AVANESSIAN, Armen. *Aceleracionismo: estrategias para una transición hacia el postcapitalismo*. Buenos Aires: Caja Negra, 2017.
- BARAD, Karen. What Is the Measure of Nothingness? Infinity, Virtuality, Justice. *Infrasonica*, Sonic Realism, # wave 1, abril 2020. Disponível em: <https://www.infrasonica.org/en/wave-1/what-is-the-measure-of-nothingness>. Acesso em: ago. 2024.
- _____. TransMaterialities: Trans*/Matter/Realities and Queer Political Imaginings. *GLQ: A Journal of Lesbian and Gay Studies*, 21(2-3), pp. 387-422, 2015.
- _____. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Durham: Duke University Press, 2007.
- BARDINI, Pedro. Por que cientistas ainda não acham que estamos no antropoceno. *Nexo Jornal*, 8 mar. 2024. Disponível em: <https://www.nexojornal.com.br/expresso/2024/03/08/por-que-cientistas-ainda-nao-acham-que-estamos-no-antropoceno>. Acesso em: 5 maio 2024.

BARBROOK, Richard; CAMERON, Andy. The Californian ideology. *Mute*, v. 1, n. 3, 1995. Disponível em: <https://www.metamute.org/editorial/articles/californian-ideology>. Acesso em: 7 ago. 2024.

BARLOW, John P. *A Declaration of the Independence of Cyberspace*. 1996. Disponível em: <https://www.eff.org/pt-br/cyberspace-independence>. Acesso em: 30 abril 2024.

BATESON, Gregory. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. Chicago: University of Chicago Press, 1972.

Trecho de: Leticia Cesarino. “O mundo do avesso (Coleção Exit)”. Apple Books.

BECK, Ulrich. *Sociedade de risco: Rumo a uma outra modernidade*. São Paulo: Editora 34, 2011.

BEIGUELMAN, Giselle. *Políticas da Imagem: vigilância e resistência na dadosfera*. São Paulo: Ubu Editora, 2021.

_____. Máquinas companheiras. *Morel*, n. 7, pp.. 76-86, 2023.

BENBOUZID, Bilel; CARDON, Dominique. Machines à prédire. *Réseaux*, vol. 211, n. 5, pp. 9-33, maio 2018.

BENDER, Emily M.; GEBRU, Timnit Gebru; MCMILLAN-MAJOR, Angelina; SHMITCHELL, Shmargaret. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610, 623, março de 2021. Disponível em: <https://dl.acm.org/doi/10.1145/3442188.3445922>.

BENJAMIN, Walter. *Magia e Técnica, arte e política: ensaios sobre literatura e história da cultura*. São Paulo: Brasiliense, 2012.

BENTES, Anna C. Franco. *Da Madison Avenue ao Vale do Silício: ciências comportamentais do engajamento, tecnologias de influência e economia da atenção*. Tese (Doutorado em Comunicação e Cultura) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2022.

BENTES, Anna. Dobras #47 // O Design Comportamental e a Economia da Atenção. *Blog do MediaLab.UFRJ*, maio 2021. Disponível em: <https://medialabufrj.net/blog/2021/05/dobras-47-o-design-comportamental-e-a-economia-da-atencao-enganchar-engajar-e-influenciar/>. Acesso em: 4 fev. 2022.

BERARDI, Franco B. *Asfixia: capitalismo financeiro e a insurreição da linguagem*. São Paulo: Ubu Editora, 2020.

_____. *Depois do Futuro*. São Paulo: Ubu Editora, 2019.

_____. *Futurabilidad: la era de la impotencia y el horizonte de posibilidad*. Buenos Aires: Caja Negra, 2019b.

- BERGSON, Henri. *Matéria e Memória: ensaio sobre a relação do corpo com o espírito*. 4. ed. São Paulo: Editora Martins Fontes, 2011.
- BERNSTEIN, Peter L. *Desafio aos Deuses: a fascinante história do risco*. Rio de Janeiro: Elsevier, 2011.
- BEVERUNGEN, Armin. Algorithmic Trading, Artificial Intelligence and the Politics of Cognition. In: SUDMANN, Andreas. *The Democratization of Artificial Intelligence: Net Politics in the Era of Learning Algorithms*. Bielefeld: transcript Verlag, 2019.
- BIRHANE, Abeba. Colonização Algorítmica da África. In: SILVA, Tarcízio. *Comunidades, algoritmos e ativismos digitais: olhares afrodiaspóricos*. São Paulo: Editora LiteraRUA, 2020.
- _____. The Impossibility of Automating Ambiguity. *Artificial Life*, vol. 27, n. 1, pp. 1-18, 2021.
- BHANDARI, Aparajita, BIMO, Sara Bimo. Why's Everyone on TikTok Now? The Algorithmized Self and the Future of Self-Making on Social Media. *Social Media + Society* Janeiro-março 2022, pp. 1-11.
- BONALDO, Rodrigo B. As palavras e os tokens: Projeção vetorial aplicada ao estudo da semântica dos tempos históricos. *Revista de Teoria da História*, v. 27, n. 1, p. 7–50, 2024. Disponível em: <https://revistas.ufg.br/teoria/article/view/79370>. Acesso em: 29 ago. 2024.
- _____. História mais do que humana: descrevendo o futuro como atualização repetidora da Inteligência Artificial. *História (São Paulo)*, v.42, e2023037, 2023. Disponível em: <https://doi.org/10.1590/1980-4369e2023037>. Acesso em 7 ago. 2024.
- BORGES, Jorge Luis. *História da Eternidade*. São Paulo: Companhia das Letras, [1936] 2010.
- BOSTROM, Nick. The Future of Humanity. In: OLSEN, Jan-Kyrre B., SELINGER, Evan, RIIS, Soren (eds.). *New Waves in Philosophy of Technology*. New York: Palgrave MacMillan, 2009.
- BRIDLE, James. *A nova idade das trevas: A tecnologia e o fim do futuro*. São Paulo: Todavia, 2019. *E-book*.
- BROWN, Wendy. *Nas Ruínas do Neoliberalismo: A Ascensão da Política Antidemocrática no Ocidente*. São Paulo. Editora Politéia, 2019.
- _____. *Undoing the Demos: neoliberalism's stealth revolution*. New York: Zone Books, 2015.

BRUNO, Fernanda. Tecnopolítica, racionalidade algorítmica e mundo como laboratório. In: GROHMANN, Rafael. *Os laboratórios do trabalho digital: entrevistas*. São Paulo: Boitempo, 2021.

_____. Arquiteturas algorítmicas e negacionismo: a pandemia, o comum, o futuro. In: DUARTE, Luisa; GORGULHO, Victor (orgs.). *No Tremor do Mundo: ensaios e entrevistas à luz da pandemia*. Rio de Janeiro: Cobogó, 2020. E-book.

_____. A economia psíquica dos algoritmos: quando o laboratório é o mundo. *Nexo*. Jun. 2018. Disponível em: <https://www.nexojornal.com.br/ensaio/2018/A-economia-ps%C3%ADquica-dos-algoritmos-quando-o-laborat%C3%B3rio-%C3%A9-o-mundo>. Acesso em: 4 fev. 2022.

_____. *Máquinas de ver, Modos de Ser: vigilância, tecnologia e subjetividade*. Porto Alegre: Sulina, 2013.

_____. Contramanual para câmeras inteligentes: vigilância, tecnologia e percepção. *Galáxia*, n. 24, pp. 47- 63, dez. 2012.

BRUNO, Fernanda; BENTES, Anna; FALTAY, Paulo. Economia Psíquica dos Algoritmos e Laboratório de Plataforma: Mercado, Ciência e Modulação do Comportamento. *Revista Famecos*, v. 26, n. 3, 2019.

BRUNO, Fernanda; RODRÍGUEZ, Pablo M. The Dividual: Digital Practices and Biotechnologies. *Theory, Culture and Society*, set. 2021.

BRUNO, Fernanda; PEREIRA, Paula; FALTAY, Paulo; BENTES, Anna; ANTOUN, Mariana; MARRARY, Moisés; ROCHA, Natássia ; PIO, Debora; STRECKER, Helena. *'Tudo por conta própria': aplicativos de autocuidado psicológico e emocional*. [Relatório de pesquisa]. 2020. Disponível em: https://medialabufRJ.net/wp-content/uploads/2020/05/Relatorio_PsiApps_MediaLabUF-RJ-1.pdf. Acesso em: 15 abril 2022.

BRUNO, Fernanda; PEREIRA, Paula C.; FALTAY, Paulo. Inteligência artificial e saúde: ressituar o problema. *Reciis – Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, Rio de Janeiro, v. 17, n. 2, p. 235-242, 2023.

BRUNO, Fernanda; PEREIRA, Paula; FALTAY, Paulo. Introdução. In: BRUNO, Fernanda; PEREIRA, Paula; FALTAY, Paulo. *#Fail: pensar e fazer mundos a partir de suas falhas e ruínas (no prelo)*.

BUCHER, Taina. A technicity of attention: how software ‘makes sense’. *Culture machine*, v. 13, 2012.

_____. *If...Then: algorithmic power and politics*. New York : Oxford University Press, 2018.

_____. Imaginários e políticas dos algoritmos: entrevista com Taina Bucher. *DigiLabour*, 12 jul. 2019. Disponível em:

<https://digilabour.com.br/2020/07/12/imaginarios-e-politicas-dos-algoritmos-entrevista-com-taina-bucher/>. Acesso em: 4 fev. 2022.

- BUTLER, Judith. *Problemas de gênero: feminismo e subversão da identidade*. Rio de Janeiro: Civilização Brasileira, 2003.
- CALIMAN, Luciana V.; ALMEIDA, Rogério G. Entrevista com Ian Hacking (por Paul Kennedy e David Cayley). *Psicologia e Sociedade*, vol. 21, n.3, pp. 465-470, 2009.
- CALIMAN, Luciana. Os regimes da atenção na subjetividade contemporânea. *Arq. bras. psicol.* [online]. vol.64, n.1, pp. 2-17, 2012.
- CALLON, Michel. *The Laws of the Markets*. Wiley-Blackwell, 1998.
- CALUDE, Cristian; LONGO, Giuseppe. The Deluge of Spurious Correlations in Big Data. *Foundations of Science*. vol. 22., pp. 595–612, mar. 2016.
- CANCLINI, Néstor G. *Cidadãos substituídos por algoritmos*. São Paulo: Edusp, 2021.
- CARDON, Dominique. *Algorithmic Personalization: Sociological and Ethical Issues*. [Palestra proferida na conferência Legal Challenges of the Data Economy], mar. 2019. Disponível em: <https://www.law.uchicago.edu/recordings/dominique-cardon-algorithmic-personalization-sociological-and-ethical-issues>. Acesso em: 15 abril 2022.
- _____. The power of algorithms. *Pouvoirs*, vol. 4, n. 1, pp. 63-7, 2018.
- CARDON, Dominique; COINTET, Jean-Philippe; MAZIERES, Antoine. Neurons spike back: The Invention of Inductive Machines and the Artificial Intelligence Controversy. *Réseaux*, vol. 5, n. 211, pp. 173-220, 2018.
- CARDOSO, Paula. O Cadastro Positivo e o Ranking do Homem Endividado. *Blog do MediaLab.UFRJ*, maio 2019. Disponível em: <https://medialabufrj.net/blog/2019/05/dobras-32-o-cadastro-positivo-e-o-ranking-do-homem-endividado/>. Acesso em: 4 fev. 2022.
- CASTRO, Edgardo. *El Vocabulario de Michel Foucault: un recorrido alfabético por sus temas, conceptos y autores*. Bernal: Universidad Nacional de Quilmes, 2004.
- CELIS BUENO, Claudio; SCHULTZ, María Jesús. Extractivismo de datos. *Imaginación maquina*, 2021. Disponível em: <http://imaginacionmaquinica.cl/extractivismo-de-datos>. Acesso em: 1 maio 2022.
- _____. Imaginación, política, algoritmos. *Imaginación maquina*, 2021. Disponível em: <https://imaginacionmaquinica.cl/post-350#post-350-footnote-30>. Acesso em: 13 ago. 2024.
- CELIS BUENO, Claudio. *The Attention Economy: Labour, Time and Power in Cognitive Capitalism*. London: Rowman & Littlefield, 2017.

_____. La economía de la atención: del ciber-tiempo al tiempo cinematográfico. *Revista Hipertextos*, vol. 8, n. 14, pp. 59-71, 2020.

CEPELEWICZ, Jordana. To Make Sense of the Present, Brains May Predict the Future. *Quanta Magazine*, 10 jul. 2018. Disponível em: <https://www.quantamagazine.org/to-make-sense-of-the-present-brains-may-predict-the-future-20180710/>. Acesso em: 4 fev. 2022.

CESARINO, Leticia. Pós-verdade e a crise do sistema de peritos: uma explicação cibernética. *Ilha Revista de Antropologia*, Florianópolis, v. 23, n. 1, p. 73–96, 2021. DOI: 10.5007/2175-8034.2021.e75630. Disponível em: <https://periodicos.ufsc.br/index.php/ilha/article/view/75630>. Acesso em: 7 ago. 2024.

_____. *O Mundo do Avesso: verdade e política na era digital*. São Paulo: Ubu, 2022.

CHAKRABARTY, Dipesh. The Climate of History: Four Theses. *Critical Inquiry*, The University of Chicago Press, vol. 35, n. 2, pp. 97-222, 2009. Disponível em: <https://www.journals.uchicago.edu/doi/10.1086/596640>. Acesso em: 7 ago. 2024.

CHAMAYOU, Grégoire. *Teoria do Drone*. São Paulo: Ubu, 2015.

CHENEY-LIPPOLD, J. *We are Data: Algorithms and the Making of Our Digital Selves*. New York: New York University Press, 2017.

CHOMSKY, Noam; ROBERTS, Ian; WATUMULL, Jeffrey. A falsa promessa do ChatGPT. *Folha de São Paulo*, 10 mar. 2023. Disponível em: <https://www1.folha.uol.com.br/tec/2023/03/a-falsa-promessa-do-chatgpt.shtml>. Acesso em: 5 maio 2024.

CHUN, Wendy H. K. *Updating to Remain the Same: habitual new media*. Cambridge, London: MIT Press, 2016.

_____. Querying Homophily. In: APPRICH, Clemens; CHUN, Wendy H. K.; CRAMER, Florian; STEYERL, Hito. *Pattern Discrimination*. Lüneburg, Minneapolis: meson press, University of Minnesota Press, 2018.

_____. *Discriminating data : correlation, neighborhoods, and the new politics of recognition*. Cambridge: The MIT Press, 2021.

CODING RIGHTS. “IA para o bem comum: fortalecendo sistemas alternativos”, recomendações do encontro com Timnit Gebru. *Site da Coding Rights*, 4 de julho de 2024. Disponível em: <https://codingrights.org/library-item/ai-e-outras-taticas-para-fomentar-ecossistemas-locais-descentralizados-para-o-desenvolvimento-de-tecnologia-situada/>. Acesso em: 31 ago. 2024.

COMAROFF, Jean; COMAROFF, John. Millennial capitalism: first thoughts on a second coming. *Public Culture*, v. 12, n. 2, pp. 291-343, 2000.

- CORRÊA, Murilo D. C. Chapação maquínica, alucinação estatística: pensar como o ChatGPT. Rede Universidade Nômade (blog), 23 de março de 2023. Disponível em: <https://uninomade.net/chapacao-maquinica-alucinacao-estatistica-pensar-como-o-chatgpt/>. Acesso em: 20 ago. 2024.
- COULDRY, Nick; MEJÍAS, Ulises A. *The Costs of Connection: how data is colonizing human life and appropriating it for capitalism*. Redwood City: Stanford University Press, 2019.
- CORTIZ, Diogo. *Diogo Cortiz*, 15 de fevereiro de 2023. Entendendo as alucinações do ChatGPT. Disponível em: <https://diogocortiz.com.br/entendendo-as-alucinacoes-do-chatgpt/>. Acesso em: 29 ago. 2024.
- COSTA, Alyne. Acreditar no mundo, imprever o fim do mundo. In: BRUNO, Fernanda; PEREIRA, Paula; FALTAY, Paulo. *#Fail: fazer mundos a partir de suas falhas e ruínas. No prelo*.
- COSTA, Flavia. *Algoritmos, biohackers y nuevas formas de vida*. Buenos Aires: Taurus, 2021.
- CUPPLES, Sarah. Frictionless design, frictionless racism. *Medium*, 1 de fevereiro de 2021. Disponível em: <https://uxdesign.cc/frictionless-racism-1097022d07f8>. Acesso em: 25 ago. 2024.
- CRARY, Jonathan. *Capitalismo tardio e os fins do sono*. São Paulo: Ubu, 2016.
- CRAWFORD, Kate. *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*. New Haven, London: Yale University Press, 2021.
- DANOWSKI, Deborah; VIVEIROS DE CASTRO, Eduardo. *O passado ainda está por vir*. São Paulo: n-1 edições, 2023.
- DANOWSKI, Deborah; VIVEIROS DE CASTRO, Eduardo. *Há mundo por vir? Ensaio sobre os medos e os fins*. Desterro (Florianópolis): Cultura e Barbárie, Instituto Socioambiental, 2017.
- DARPA. DoD News Briefing - Secretary Rumsfeld and Gen. Myers. *News Transcript*, 12 fev. 2002. Disponível em: <https://web.archive.org/web/20160406235718/http://archive.defense.gov/Transcripts/Transcript.aspx?TranscriptID=2636>. Acesso em: 9 jul. 2024.
- DASTON, Lorraine. *Rules: A Short History of What We Live By*. Princeton: Princeton University Press, 2022
- DAVIS, William. How statistics lost their power – and why we should fear what comes next. *The Guardian*, 19 jan. 2017. Disponível em: <https://www.theguardian.com/politics/2017/jan/19/crisis-of-statistics-big-data-democracy>. Acesso em: 15 abril 2022.

- DELEUZE, Gilles. Post-scriptum sobre as sociedades de controle. In: DELEUZE, Gilles. *Conversações: 1972-1990*. São Paulo: Editora 34, 1992.
- DELEUZE, Gilles. Prefácio: As andaduras do tempo. In: DELEUZE, Gilles. *Dois Regimes de Loucos: Textos e entrevistas (1975-1995)*. São Paulo: Editora 34, 2016.
- _____. *Foucault*. São Paulo: Editora Brasiliense, 2005.
- DIDI-HUBERMAN, Georges. *Atlas - Cómo llevar el mundo a cuestras?* [Catálogo da Exposição do Museu Nacional Rainha Sofia]. Madrid: Museu Nacional Rainha Sofia, 2010.
- DIETER, Michael; GAUTHIER, David. On the Politics of Chrono-Design: Capture, Time and the Interface. *Theory, Culture & Society*, vol. 36, n. 2, pp. 61–87, mar. 2019.
- DOMINGOS, Pedro. *O Algoritmo Mestre*. São Paulo: Novatec Editora, 2017.
- DORNELLES, D.S.; BONALDO, R.B. História e distopia: três abordagens teóricas (presentismo, atualismo e um futuro sem precedentes). *Revista Aedos*, [S. l.], v. 13, n. 30, p. 21-41. Disponível em: <https://seer.ufrgs.br/index.php/aedos/article/view/105283>. Acesso em: 7 ago. 2024.
- DRESSEL, Julia; FARID, Hany. The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, vol. 4, n. 1., jan. 2018. Disponível em: <https://www.science.org/doi/10.1126/sciadv.aao5580>. Acesso em: 15 abril 2022.
- EYAL, Nir. *Hooked: como construir produtos e serviços formadores de hábitos*. Editora AlfaCon: Cascavel, 2020.
- ESPOSITO, Elena. *The Future of Futures: The Time of Money in Financing and Society*. Cheltenham, Northampton: Edward Elgar Publishing, 2011.
- ESTEVES, Bernardo. O Antropoceno já era: Os guardiões da linha do tempo rejeitaram a criação de uma nova época geológica. *Piauí*, ed. 211, abril de 2024. Disponível em: <https://piaui.folha.uol.com.br/materia/o-antropoceno-ja-era/>. Acesso em: 15 dez. 2024.
- EUBANKS, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: Picador, 2019.
- FALTAY, Paulo. *Máquinas paranoides e sujeitos influenciáveis: conspiração, conhecimento e subjetividade em redes algorítmicas*, 2020. Tese (Doutorado em Comunicação e Cultura) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2020.
- FARMAN, Abou. Re-enchantment cosmologies: mastery and obsolescence in an intelligent universe. *Anthropological Quarterly*, vol. 85, n. 4, pp. 1069-1088, 2012. JSTOR, The George Washington University Institute for Ethnographic Research. Disponível em: <https://www.jstor.org/stable/41857290>. Acesso em: 7 ago. 2024.
- FAUSTINO, Deivison; LIPPOLD, Walter. *Colonialismo Digital: por uma Crítica Hacker-fanoniana*. São Paulo: Boitempo, 2023.

FAZI, Beatrice. Can a machine think (anything new)? Automation beyond simulation. *AI & Society*, v. 34, n. 813, p. 813-824, 2019. Disponível em: <https://doi.org/10.1007/s00146-018-0821-0>. Acesso em: 7 ago. 2024.

_____. *Contingent Computation: Abstraction, experience, and indeterminacy in computational aesthetics*. London, New York: Rowman Littlefield, 2018.

FERDINAND, Malcom. *Uma ecologia decolonial: pensar a partir do mundo caribenho*. São Paulo: Ubu Editora, 2022.

FERREIRA DA SILVA, Denise. *A dívida impagável: Uma crítica feminista, racial e anticolonial do capitalismo*. Rio de Janeiro: Zahar, 2024.

_____. *Homo Modernus: Para uma ideia global de raça*. Coleção Encruzilhada. Rio de Janeiro: Cobogó, 2022.

FINN, Ed. *What Algorithms want*. London: MIT Press, 2017.

FISHER, Mark. *Fantasma da minha vida: escritos sobre depressão, assombrologia e futuros perdidos*. São Paulo: Autonomia Literária, 2022.

_____. *Realismo Capitalista: no hay alternativa?*. Buenos Aires: Caja Negra, 2019.

FOUCAULT, Michel. *Vigiar e punir: nascimento da prisão*. Petrópolis: Editora Vozes, [1975] 2014.

_____. *Segurança, Território e População: curso dado no College de France (1977-1978)*. São Paulo: Martins Fontes, 2008.

_____. *O nascimento da biopolítica: curso dado no College de France (1978-79)*. São Paulo: Martins Fontes, 2008b.

_____. *Ditos & Escritos IV: estratégia, poder-saber*. Rio de Janeiro: Editora Forense Universitária, 2006.

_____. *As palavras e as coisas: uma arqueologia das ciências humanas*. 8ª ed. São Paulo: Martins Fontes, [1966] 2000.

_____. O que são as luzes? In: *Ditos e escritos II – arqueologia das ciências e história dos sistemas de pensamento*, pp. 335-351. Rio de Janeiro: Forense Universitária, 2000b.

_____. O Sujeito e o Poder. In: RABINOW, Paul, DREYFUS, Hubert L. *Michel Foucault: uma trajetória filosófica (para além do estruturalismo e da hermenêutica)*. Rio de Janeiro: Forense Universitária, 1995.

FLUSSER, Vilém. *Pós-História: vinte instantâneos e um modo de usar*. São Paulo: É Realizações, 2019.

- _____. *O último juízo: gerações II: castigo e penitência*. São Paulo: É Realizações, 2017.
- FUKUYAMA, Francis. *O Fim da História e o Último Homem*. Rio de Janeiro: Rocco, 2015 [1989].
- GAGLIONI, César. O que é ‘longoprazismo’. E quais as críticas a ele. *Nexo Jornal*, 22 ago.2022. Disponível em: <https://www.nexojornal.com.br/expresso/2022/08/22/o-que-e-longoprazismo-e-quais-as-criticas-a-ele>. Acesso em: 6 maio 2024.
- GEBRU, T., & Torres, Émile P. The TESCREAL bundle: Eugenics and the promise of utopia through artificial general intelligence. *First Monday*, vol. 29, n.4, abril 2024. Disponível em: <https://doi.org/10.5210/fm.v29i4.13636>. Acesso em: 7 ago. 2024.
- GIBSON, William. *Neuromancer*. São Paulo: Editora Aleph, [1984] 2016.
- GILLESPIE, Tarleton. Algorithm. In: PETERS, Benjamin (ed.). *Digital Keywords: a vocabulary of information society and culture*. Princeton, Oxford: Princeton University Press, 2016.
- GITELMAN, Lisa; JACKSON, Virginia. Introduction. In: GITELMAN, Lisa. *“Raw Data” is an Oxymoron*. Cambridge, London: MIT Press, 2013.
- GREAVES, Hilary, MACASKILL, William. *The case for strong longtermism*. Global Priorities Institute, junho 2021. Disponível em: <https://globalprioritiesinstitute.org/wp-content/uploads/The-Case-for-Strong-Longtermism-GPI-Working-Paper-June-2021-2-2.pdf>. Acesso em: 6 maio 2024.
- GREENWALD, Michelle. Audience, Algorithm And Virality: Why TikTok Will Continue To Shape Culture In 2021. *Forbes*, 1 de abril de 2021. Disponível em: <https://www.forbes.com/sites/michellegreenwald/2021/04/01/audience-algorithm-and-virality-why-tiktok-will-continue-to-shape-culture-in-2021/?sh=6a1c6f0c2af7>. Acesso em: 23 ago. 2024.
- GROYS, Boris. From Writing to Prompting: AI as Zeitgeist-Machine. *E-Flux Notes*, 10 de agosto de 2023. Disponível em: <https://www.e-flux.com/notes/553214/from-writing-to-prompting-ai-as-zeitgeist-machine>. Acesso em: 29 ago. 2024.
- GUYER, Jane. Prophecy and the near Future: Thoughts on Macroeconomic, Evangelical, and Punctuated Time. *American Ethnologist*, v. 34, n. 3, pp. 409-21, 2007.
- HACKING, Ian. *The Emergence of Probability: a philosophical study of early ideas about probability, induction and statistical inference*. Cambridge: Cambridge University Press, 2006 [1975].
- _____. *The Taming of Chance*. Cambridge: Cambridge University Press, 1990.
- HALPERN, Orit. The Future Will Not Be Calculated: Neural Nets, Neoliberalism, and Reactionary Politics. *Critical Inquiry*, v. 48, n. 2, pp. 334-359, 2022.

HAN, Byung-Chul. *Psicopolítica: o neoliberalismo e as novas técnicas de poder*. Belo Horizonte: Âyiné, 2018.

HARAWAY, Donna. Manifesto ciborgue: ciência, tecnologia e feminismo-socialista no final do século XX. In: BUARQUE DE HOLLANDA, Heloisa. *Pensamento feminista: conceitos fundamentais*. Rio de Janeiro: Bazar do Tempo, [1985] 2019.

_____. *Modest_Witness@Second_Millennium: FemaleMan©_Meets_OncoMouseTM*. New York: Routledge, 2018

_____. *When species meet*. Minneapolis, London: University of Minesota Press, 2008.

_____. *Staying with the Trouble: making kin in the Chthulucene*. Durham, London: Duke University Press, 2016.

HARARI, Yuval N. *Homo Deus: Uma breve história do amanhã*. São Paulo: Companhia das Letras, 2016.

HARTOG, François. *Regimes de Historicidade: presentismo e experiências do tempo*. Belo Horizonte: Autêntica, 2021.

HAYLES, N. Katherine. *Unthought: the power of cognitive nonconscious*. Chicago: University of Chicago Press, 2017. *E-book*.

_____. *My Mother was a Computer: Digital Subjects and Literary Texts*. Chicago: University of Chicago Press, 2005

_____. *How we became posthuman: virtual bodies in cybernetics, literature and informatics*. Chicago: University Chicago Press, 1999.

HEALY, Kieran. The Performativity of Networks. *European Journal of Sociology/Archives Européennes de Sociologie*, vol. 56, n. 2, pp. 175–205, 2015.

HEIDEGGER, Martin. *O fim da filosofia e a tarefa do pensamento*. Trad. Ernildo Estein. São Paulo: Abril Cultural, 1979.

HELMOND, Anne. The platformization of the web: Making web data platform ready. *Social Media + Society*, v. 1, n. 2, 2015.

HERN, Alex. Cambridge Analytica: how did it turn clicks into votes?. *The Guardian*, 6 de maio de 2018. Disponível em: <https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie>. Acesso em: 8 fev. 2022.

_____. How TikTok’s algorithm made it a success: ‘It pushes the boundaries’. *The Guardian*, 24 de outubro de 2022. Disponível em: <https://www.theguardian.com/technology/2022/oct/23/tiktok-rise-algorithm-popularity>. Acesso em: 15 out. 2023.

HILDEBRANDT, Mireille. Defining profiling: a new type of knowledge? *In:* HILDEBRANDT, Mireille; GUTWIRTH, Serge. (Orgs). *Profiling the European citizen. Cross-disciplinary perspectives*. Dordrecht: Springer Science, 2008.

HONG, Sun-ha. PREDICTIONS WITHOUT FUTURES. *History and Theory*, v. 61, n. 3, p. 371–390, 24 ago. 2022. Disponível em: <https://doi.org/10.1111/hith.12269>. Acesso em: 10 ago. 2024.

_____. Prediction as Extraction of Discretion. *In:* Conference on Fairness, Accountability, and Transparency (FAccT '22), June 2022, Seoul, Republic of Korea. *Proceedings [...]*. New York, NY, USA: ACM- Association for Computing Machinery, 2022b. Disponível em: <https://doi.org/10.1145/3531146.3533155>. Acesso em: 10 ago. 2024.

HONG, Sun-ha; SZPUNAR, Piotr M. The Futures of Anticipatory Reason: Contingency and Speculation in the Sting Operation. *Security Dialogue*, v. 50, n. 4, p. 314–330, 24 jun. JSTOR, 2019. Disponível em: <https://www.jstor.org/stable/26753663>. Acesso em: 10 ago. 2024.

HÖLSCHER, Lucian. *El Descubrimiento del Futuro*. Madrid: Siglo XXI, 2014.

HUI, Yuk. Problems of Temporality in the Digital Epoch. *In:* VOLMAR, Axel, STINE, Kyle (eds.). *Media Infrastructures and the Politics of Digital Time: Essays on Hardwired Temporalities*. Amsterdam: Amsterdam University Press, 2021.

_____. *Tecnodiversidade*. São Paulo: Ubu Editora, 2020.

_____. *Recursivity and contingency*. Lanham: Rowman & Littlefield International, 2019.

INTRONA, Lucas D. Algorithms, Governance, and Governmentality: On Governing Academic Writing. *Science, Technology, & Human Values*, vol. 41, n. 1, pp. 17–49, 2016.

_____. The algorithmic choreography of the impressionable subject. *In:* SEYFERT, Robert, ROBERGE, Jonathan (eds.). *Algorithmic Cultures: essays on meaning, performance and new technologies*. Abindgon, New York: Routledge, 2016a.

JAMESON, Fredric. *Pós-modernismo: a lógica cultural do capitalismo tardio*. São Paulo: Ática, 2002.

JOHN-MATHEWS, Jean-Marie; CARDON, Dominique; BALAGUÉ, Christine. From Reality to World: A Critical Perspective on AI Fairness. *Journal of Business Ethics*, vol. 178, pp. 945–959, 2022. Disponível em: <https://doi.org/10.1007/s10551-022-05055-8>. Acesso em: 10 ago. 2024.

- JOHO, Jess. TikTok's algorithms knew I was bi before I did. I'm not the only one. *Mashable*, 8 de setembro de 2022. Disponível em: <https://mashable.com/article/bisexuality-queer-tiktok>. Acesso em: 20 out. 2023.
- JONES, Phil. *Work without the Worker: labour in the age of platform capitalism*. Verso Books: 2021.
- KANASHIRO, Marta M. ChatGPT: o que *não* está se debatendo. *Outras Palavras*, 24 de abril de 2023. Disponível em: <https://outraspalavras.net/tecnologiaemdisputa/chatgpt-o-que-nao-esta-se-debatendo/>. Acesso em: 31 ago. 2024.
- KOSELECK, Reinhart. *Futuro Passado: contribuição à semântica dos tempos históricos*. Rio de Janeiro: Contraponto, Editora PUC-Rio, 2006.
- _____. *Estratos do Tempo: estudos sobre história*. Rio de Janeiro: Contraponto, Editora PUC-Rio, 2014.
- KOSINSKI, Michal; STILLWELL, David; GRAEPEL, Thore Graepel. Private Traits and Attributes Are Predictable from Digital Records of Human Behavior. *Proceedings of the National Academy of Sciences*, vol. 110, n. 15, pp. 5802–5805, abril 2013.
- KOSINSKI, Michal; WANG, Yilun. Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation From Facial Image. *Journal of Personality and Social Psychology*, vol. 114, n. 2, pp. 246–257, fev. 2018.
- KRENAK, Ailton. *Ideias para adiar o fim do mundo*. São Paulo: Companhia das Letras, 2019.
- KRUG, Steve. *Não me Faça Pensar: Atualizado*. Rio de Janeiro: Alta Books, 2014.
- KURZWEIL, Ray. *The singularity is near: when humans transcend biology*. New York: Penguin, 2005.
- _____. The Law of Accelerating Returns. In: TEUSCHER, C. *Alan Turing: Life and Legacy of a Great Thinker*. Springer: Berlin, Heidelberg, 2004.
- LAPOUJADE, David. Desprogramar o futuro. In: NOVAES, Adauto (org). *Mutações: o futuro não é mais o que era*. São Paulo: Edições SESC São Paulo, 2015.
- LATOURE, Bruno. *Onde Aterrizar? como se orientar politicamente no antropoceno*. Bazar do Tempo, 2020.
- _____. *Jamais fomos modernos: ensaio de antropologia simétrica*. Tradução de Carlos Irineu da Costa. São Paulo: Editora 34, 2019.
- _____. *Reagregando o Social: uma introdução à teoria do Ator-Rede*. Salvador/Bauru: Edufba/Edusc, 2012.
- LAZZARATO, Maurizio. *Signos, máquinas, subjetividades*. São Paulo: Edições SESC; n-1 edições, 2014.
- _____. *O governo do homem endividado*. São Paulo: n-1 edições, 2017.

- LAW, John; SINGLETON, Vicky. Performing technology's stories: on social constructivism, performance and performativity. *Technology and Culture*, v. 41, n. 4, p. 765-775, out. 2000.
- LE GUIN, Ursula K. *A teoria da bolsa da ficção*. São Paulo: n-1 edições, 2021.
- LEVY, Pierre. *A inteligência coletiva: Por uma antropologia do ciberespaço*. São Paulo: Edições Loyola, 1998.
- LYOTARD, Jean-François. *A condição pós-moderna*. 12ª ed. Rio de Janeiro: José Olympio, [1979] 2009.
- LUHMANN, Niklas. *Sistemas sociais: esboço de uma teoria geral*. Petrópolis: Editora Vozes, [1984] 2016.
- MACKENZIE, Adrian. The production of prediction: What does machine learning want?. *European Journal of Cultural Studies*, vol. 18, n. 4-5, pp. 429-445, 2015.
- _____. The Performativity of Code: Software and Cultures of Circulation. *Theory, Culture & Society*, vol. 22, n. 1, pp. 71-92, 2005.
- MACKENZIE, Donald. *An engine Not a Camera: how financial models shape markets*. Cambridge: MIT Press, 2008.
- MACASKILL, William. Effective Altruism: Introduction. *Essays on Philosophy*, vol. 18, n. 1, *What We Owe the Future*. New York: Basic Books, 2022.
- MALABOU, Catherine. *Morphing intelligence: from iq measurement to artificial brains*. New York : Columbia University Press, 2019.
- MANN, Hamilton. AI Homogenization Is Shaping The World. *Forbes*, 5 de março de 2024. Disponível em: <https://www.forbes.com/sites/hamiltonmann/2024/03/05/the-ai-homogenization-is-shaping-the-world/>. Acesso em: 29 ago. 2024.
- MAYBIN, Simon. Busting the attention span myth. BBC News, 10 de março de 2017. Disponível em: <https://www.bbc.com/news/health-38896790>. Acesso em: 27 ago. 2024.
- MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. Boston: Houghton Mifflin Harcourt, 2013.
- MARCUS, Gary. How come GPT can seem so brilliant one minute and so breathtakingly dumb the next?. *Marcus on AI*, 1 de dezembro de 2022. Disponível em: <https://garymarcus.substack.com/p/how-come-gpt-can-seem-so-brilliant>. Acesso em: 29 ago. 2024.
- MARKELJ, Jernej, BUENO, Claudio C. Machinic agency and datafication: Labour and value after anthropocentrism. *Convergence*, vol. 30, n.3, pp. 1058-1075, 2024.

- MASSUMI, Brian. *Ontopower: war, powers, and the state of perception*. Durham: Duke University Press, 2005.
- MBEMBE, Achille. *Brutalismo*. São Paulo: editora n-1, 2021.
- _____. Futures of Life and Futures of Reason. *Public Cultures*, vol. 33, n. 1, pp. 11-33, janeiro de 2021 [2021b].
- _____. Thoughts on the planetary: An interview with Achille Mbembe. *New Frame*, 17, janeiro, 2019b. Disponível em: <https://www.newframe.com/thoughts-on-the-planetary-an-interview-with-achille-mbembe/>. Acesso em: 4 fev. 2022.
- _____. Bodies as Borders. *The European South*, n. 4, pp. 5–18, 2019.
- _____. *Crítica da Razão Negra*. São Paulo: n-1 edições, 2018.
- MCQUILLAN, Dan. Algorithmic paranoia and the convivial alternative. *Big Data & Society*, nov. 2016.
- MCLUHAN, Marshall. *Os meios de comunicação como extensão do homem*. São Paulo: Cultrix, 1974.
- MCSPADDEN, Kevin. You Now Have a Shorter Attention Span Than a Goldfish. *Time*, 14 de maio de 2015. Disponível em: <https://time.com/3858309/attention-spans-goldfish/>.
- MENDES, Margarida. Colonialismo Molecular. In: MENDES, Margarida (ed.). *Matter Fictions*. Berlim: Sternberg Press, 2017.
- MERTON, Robert. *Sociologia: teoria e estrutura*. São Paulo: Editora Mestre Jou, 1968.
- MESQUITA, Israel. Affordance: O óbvio do óbvio. *Medium UX Collective*, 25 set.2018. Disponível em: <https://brasil.uxdesign.cc/affordance-o-%C3%B3bvio-do-%C3%B3bvio-e91761f4403b>. Acesso em: 10 jul. 2024.
- MICROSOFT RESEARCH. Sparks of Artificial General Intelligence: Early experiments with GPT-4 [Relatório de pesquisa]. março de 2023. Disponível em: <https://www.microsoft.com/en-us/research/publication/sparks-of-artificial-general-intelligence-early-experiments-with-gpt-4/>. Acesso em: 20 de ago. de 2024.
- MIKE, Ananny; CRAWFORD, Kate. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, vol. 20, n. 3, pp. 973-989, dez. 2016. Disponível em: <https://doi.org/10.1177/1461444816676645>. Acesso em: 11 ago. 2024.
- MINOIS, Georges. *História do Futuro: dos profetas à prospectiva*. São Paulo: Editora Unesp, 2015. *E-book*.

MONAHAN, Torin. Algorithmic Fetishism. *Surveillance & Society*, v. 16, n. 1, pp. 1–5, abr. 2018.

MORAES, Alana. Tecno-governança extrativista e a nova colonialidade cibernética: como habitar a incerteza?. In: BRUNO, Fernanda; PEREIRA, Paula; FALTAY, Paulo. *#Fail: fazer mundos a partir de suas falhas e ruínas. No prelo*.

MURARI, Lucas; ANDUEZA, Nicholas; CARDOSO, Paula. Máquinas, visualidades, relações - da inteligência artificial à artificialidade da inteligência: entrevista com Giselle Beiguelman. *Revista Eco-Pós*, vol. 27, n. 1, 2024.

MOROZOV, Evgeny. *Big Tech: a ascensão dos dados e a morte da política*. São Paulo: Ubu Editora, 2018.

MORTON, Timothy. *Hyperobjects: Philosophy and Ecology after the End of the World*. Minneapolis: University of Minnesota Press, 2013.

NALIN, Carolina. “Não podemos repetir o que fizemos com as redes sociais”, diz pesquisadora sobre inteligência artificial. *O Globo*, 5 de abril de 2023. Disponível em: <<https://oglobo.globo.com/economia/tecnologia/noticia/2023/04/nao-podemos-repetir-o-que-fizemos-com-as-redes-sociais-diz-pesquisadora-sobre-inteligencia-artificial.ghtml>>. Acesso em: 29 ago. 2024.

NEWPORT, Cal. TikTok and the Fall of the Social-Media Giants. *The New Yorker*, 28 de julho de 2022. Disponível em: <https://www.newyorker.com/culture/cultural-comment/tiktok-and-the-fall-of-the-social-media-giants>. Acesso em: 15 out. 2023.

NOBLE, Safiya U. *Algoritmos da Opressão*. Santo André: Editora Rua do Sabão, 2021.

NOVAES, Adauto (org). *Mutações: o futuro não é mais o que era*. São Paulo: Edições SESC São Paulo, 2015.

OCHIGAME, Rodrigo. The invention of ‘Ethical AI’ How Big Tech Manipulates Academia to Avoid Regulation. *The Intercept*, 20 de dezembro de 2019. Disponível em: <https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/>. Acesso em: 29 ago. 2024.

OLIVEIRA, Luiz A. Luiz Alberto Oliveira: Entrevista a Rosiska Darcy de Oliveira. *Revista Brasileira*, fase X, ano II, n. 115, pp. 16-26, abril/maio/junho, 2023.

O’NEIL, Cathy. *Algoritmos de Destruição em Massa*. Santo André: Editora Rua do Sabão, 2021.

OPEN AI. *GPT-4 Technical Report* [Relatório de pesquisa]. 2023. Disponível em: <https://cdn.openai.com/papers/gpt-4.pdf>. Acesso em: 25 ago. 2024.

_____. *Language Models are Few-Shot Learners*. 2020. Disponível em: <https://arxiv.org/abs/2005.14165>. Acesso em: 20 ago. 2024.

- ORGAZ, Cristina. 'TikTok foi feito para ser viciante': o homem que investigou as entranhas do aplicativo. *BBC News Brasil*, 3 de dezembro de 2020. Disponível em: <https://www.bbc.com/portuguese/geral-55173900>. Acesso em: 20 de out. 2023.
- ORD, Toby. *The Precipice: Existential Risk and The Future of Humanity*. New York: Hachette Books, 2020.
- PADDISON, Laura. Sam Altman diz que fusão nuclear é solução para demanda de energia de IA. *CNN Brasil*, 26 de março de 2024. Disponível em: <https://www.cnnbrasil.com.br/tecnologia/sam-altman-diz-que-fusao-nuclear-e-solucao-para-demanda-de-energia-de-ia/>
- PARISER, Eli. *O filtro invisível: O que a internet está escondendo de você*. Rio de Janeiro: Zahar, 2012.
- PARISI, Luciana. Reprograming Decisionism. *E-flux journal*, n. 85, 2017. Disponível em: <https://www.e-flux.com/journal/85/155472/reprogramming-decisionism/>. Acesso em: 4 fev. 2022.
- PARISI, Luciana; DIXON-ROMÁN, Ezequiel. Recursive Colonialism and Cosmo-Computation. *Social Text Journal* (online), 24 nov. 2020. Disponível em: https://socialtextjournal.org/periscope_article/recursive-colonialism-and-cosmo-computation/. Acesso em: 5 maio 2024.
- PARRA, Henrique; MORAES, Alana; MALHÃO, Rafael. Resumo do Seminário Temático “Tecnopolíticas, Cosmopolíticas: conflitualidades, modos de saber e tecnologias face ao plantationceno”. In: VIII Reunião de Antropologia da Ciência e Tecnologia (REACT), 2021. Disponível em: [PDF VIII ReACT - STs aprovados.pdf](#). Acesso em: 11 ago. 2024.
- PASQUINELLI, Matteo. *The Eye of the Master*. Londres: Verso Books, 2023.
- PASQUINELLI, Matteo. “How to Make a Class: Hayek’s Neoliberalism and the Origins of Connectionism”. *Qui Parle*, v. 30, n. 1, junho de 2021.
- _____. How a Machine Learns and Fails: a grammar of error for Artificial Intelligence. *Spheres*, n. 5, pp. 1-17, nov. 2019.
- _____. Machines that Morph Logic: Neural Networks and the Distorted Automation of Intelligence as Statistical Inference. *Site 1: Logic Gate: The Politics of the Artifactual Mind*, 2017. Disponível em: <http://www.glass-bead.org/article/machines-that-morph-logic/?lang=enview>. Acesso em: 4 fev. 2022.
- _____. Arcana Mathematica Imperii: The Evolution of Western Computational Norms. In: HLAVAJOVA, Maria, SHEIKH, Simon (eds.). *Former West: Art and the Contemporary after 1989*. Cambridge: The MIT Press, pp. 281-295, 2016.
- _____. *Anomaly Detection: the mathematization of the abnormal in the metadata society*. [Painel apresentado no Transmediale Festival] 2015. Disponível em:

<http://www.academia.edu/download/37642065/Pasquinelli_Anomaly_Detection.pdf>. Acesso em: 4 fev. 2022.

_____. Capitalismo maquínico e mais-valia de rede: Notas sobre a economia política da máquina de Turing. *Lugar Comum*, n. 39, pp. 13-36, 2013.

_____. Google's PageRank Algorithm: a diagram of cognitive capitalism and the rentier of the common intellect. In: BECKER, F., STALDER, F. (eds). *Deep Search*. London: Transaction Publishers, 2009.

PASQUINELLI, Matteo; JOLER, Vlader. O Manifesto Nooscópio: Inteligência Artificial como Instrumento de Extrativismo do Conhecimento. *Blog da Lavits*, maio de 2020. Disponível em: <https://lavits.org/o-manifesto-nooscopio-inteligencia-artificial-como-instrumento-de-extrativismo-do-conhecimento/?lang=pt>. Acesso em: 4 fev. 2022.

PEIRCE, Charles. Deduction, Induction, and Hypothesis. *Popular Science Monthly*, vol. 13, 1878, pp. 470-482. Disponível em: <https://cspeirce.omeka.net/items/show/7>.

PERDOMO, Juan; ZRNIC, Tijana; MENDLER-DÜNNER, Celestine; HARDT, Moritz. *Performative Prediction*. arXiv, 2021. Disponível em: <https://arxiv.org/abs/2002.06673>. Acesso em: 8 fev. 2022.

PEREIRA, Paula. // “Como trabalham os brasileiros que treinam inteligências artificiais?” Paula Cardoso comenta matéria do Intercept Brasil. *Blog do MediaLab.UFRJ*, 21 de junho de 2023. Disponível em: <https://medialabufrj.net/blog/2023/06/como-trabalham-os-brasileiros-que-treinam-inteligencias-artificiais-paula-cardoso-comenta-materia-do-intercept-brasil/>. Acesso em: 15 de dez. 2023.

_____. #Dobras 38 // Precariado algorítmico: o trabalho humano fantasma nas maquinarias da inteligência artificial. *Blog do MediaLab.UFRJ*, set. 2019. Disponível em: <https://medialabufrj.net/blog/2019/09/dobras-38-precariado-algoritmico-o-trabalho-humano-fantasma-nas-maquinarias-da-inteligencia-artificial/>. Acesso em: 15 abril 2022.

PEREIRA, Mateus; ARAÚJO, Valdei. *Atualismo 1.0: como a ideia de atualização mudou o século XXI*. Vitória: Editora Mil Fontes, 2022.

_____. O passado como distração: Modos de vestir a história no neopopulismo brasileiro. *Revista De Teoria Da História*, vol. 25, n. 2, pp. 70–88, 2022b.

PERRIGO, Billy. Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic. *Time*, 18 de janeiro de 2023. Disponível em: <https://time.com/6247678/openai-chatgpt-kenya-workers/>. Acesso em: 27 ago. 2024.

PETRONIO, Rodrigo. O tempo que vem: historicidade e escatologia em Vilém Flusser. In: *Pós-História: Vinte instantâneos e um modo de usar*. São Paulo: É Realizações, 2019.

- PIAUI. Ray Kurzweil e o mundo que nos espera: Uma entrevista com o inventor e futurólogo americano. *Piauí*, ed. 43, abril 2010. Disponível em: <https://piaui.folha.uol.com.br/materia/ray-kurzweil-e-o-mundo-que-nos-espera/>. Acesso em: 6 maio 2024.
- PICKERING, Andrew. *The Mangle of Practice: science, society, and becoming*. Durham: Duke University Press, 1995.
- QUIJANO, Aníbal. Coloniality and Modernity/Rationality. *Cultural Studies*, vol. 21, n. 2-3, pp. 168-178, 2007.
- RAUNIG, Gerald. *Dividuum: machinic capitalism and molecular revolution*. MIT Press, 2016.
- REED, Patricia. *Xenofilia e desnaturalização computacional*. São Paulo: Zazie Edições, 2018.
- RICAURTE, Paola. Ethics for the majority world: AI and the question of violence at scale. *Media, Culture & Society*, vol. 44, issue 4, pp. 726-745, maio 2022. Disponível em: <https://doi.org/10.1177/01634437221099612>. Acesso em: 11 ago 2024.
- _____. Data Epistemologies, Coloniality of Power, and Resistance. *Television & New Media*, vol. 20, n. 4, pp. 350- 365, março 2019.
- RODRÍGUEZ, Pablo M. *Las Palabras en Las Cosas: saber, poder y subjetivación entre algoritmos y biomoléculas*. Buenos Aires: Cactus, 2019.
- ROSA, Harmut. *Aceleração: a transformação das estruturas temporais na Modernidade*. São Paulo: Editora Unesp, 2019.
- ROUVROY, Antoinette. El uso de los datos masivos para gobernar. *Viento Sur*, 2 de fevereiro de 2021. Disponível em: <https://vientosur.info/el-uso-de-los-datos-masivos-para-gobernar/>. Acesso em: 4 fev. 2022.
- _____. Entrevista com Antoinette Rouvroy: Governamentalidade Algorítmica e a Morte da Política. *Revista de Filosofia Moderna e Contemporânea, Brasília*, v. 8, n. 3, pp. 15-28, dezembro de 2020.
- _____. Algorithmic governmentality: radicalisation and immune strategy of capitalism and neoliberalism?. *La Deleuziana*, n. 3, pp. 30-36, 2016.
- _____. The end(s) of critique: data behaviourism versus due process. In: HILDEBRANDT, Mireille, VRIES, Katja de. *Privacy, Due Process and the Computational Turn*. London: Routledge, 2013.
- ROUVROY, Antoinette; BERNS, Thomas. Governamentalidade Algorítmica e perspectivas de emancipação: o díspar como condição de individuação pela relação? *Revista ECO Pós*, vol. 18, n. 2, pp. 36-56, out. 2015. Disponível em: <https://doi.org/10.29146/eco-pos.v18i2.2662>. Acesso em: 11 ago. 2024.

- ROUVROY Antoinette; BERNS Thomas. Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps « numériques »..., *Multitudes*, 2010/1, n.40, pp. 88-103. DOI: 10.3917/mult.040.0088. Disponível em: <https://www.cairn.info/revue-multitudes-2010-1-page-88.htm>. Acesso em: 11 ago. 2024.
- ROUVROY, Antoinette; STIEGLER, Bernard. The Digital Regime of Truth: From Algorithmic Governmentality to a New Rule of Law. *La Deleuziana*, n. 3, pp. 6-27, 2016.
- ROUVROY, Antoinette; ATHANASIADOU, Lila; KLUMBYTÉ, Goda . Re-Imagining a 'We' Beyond the Gathering of Reductions: Propositions for the Three Ecologies. [Entrevista]. *FOOTPRINT*, The Epiphylogenetic Turn and Architecture: In (Tertiary) Memory of Bernard Stiegler, pp. 121–134, Spring/Summer 2022. Disponível em: <https://journals.open.tudelft.nl/footprint/article/view/5933>. Acesso em: 11 ago. 2024.
- RYLE, Gilbert. *The Concept of Mind*. London: Hutchinson, 1949.
- SADOWSKI, Jathan. When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, vol. 6, n. 1, pp. 1-12, 2019.
- SAMUEL, Sigal. What happens when ChatGPT starts to feed on its own writing?. *Vox*, 10 de abril de 2023. Disponível em: <https://www.vox.com/future-perfect/23674696/chatgpt-ai-creativity-originality-homogenization>. Acesso em: 29 ago. 2024.
- SAP. *Your Predictive Journey*. [Material comercial], 2017.
- SCHAUL, Kevin; CHEN, Szu Y.; TIKU, Nitasha. Inside the secret list of websites that make AI like ChatGPT sound smart. *The Washington Post*, 19 de abril de 2023. Disponível em: <https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning/>. Acesso em: 27 ago. 2024.
- SEEVER, Nick. Captivating algorithms: recommender systems as traps. *Journal of Material Culture*, v. 24, n. 4, pp. 421-436, dez. 2018.
- SEYFERT, Robert; ROBERGE, Jonathan. What are algorithmic cultures? In: SEYFERT, Robert; ROBERGE, Jonathan (eds.). *Algorithmic Cultures: essays on meaning, performance and new technologies*. Abingdon, New York: Routledge, 2016.
- SHANAHAN, Murray. Talking about Large Language Models. *Communications of the ACM*, vol. 67, n. 2, pp. 68-79, jan. 2024. Disponível em: <https://dl.acm.org/doi/10.1145/3624724>. Acesso em: 20 ago. 2024.
- SILES, Ignacio. *Living with algorithms: Agency and user culture in Costa Rica*. Cambridge, MA: MIT Press, 2023.
- SILES, Ignacio; GÓMEZ-CRUZ, Edgar; RICAURTE, Paola . Fluid agency in relation to algorithms: Tensions, mediations, and transversalities. *Convergence*, vol.30, issue 3, pp.1025-1040, maio 2023. Disponível em:

<https://journals.sagepub.com/doi/abs/10.1177/13548565231174586>. Acesso em: 11 ago. 2024.

SILES, Ignacio; VALERIO-ALFARO, Luciana; MELÉNDEZ-MORAN, Ariana. Learning to like TikTok... and not: Algorithm Awareness as Process. *New Media & Society*, v. 0, n. 0, dezembro de 2022, pp. 1-17. Disponível em: https://www.researchgate.net/publication/366231844_Learning_to_like_TikTok_and_not_Algorithm_awareness_as_process. Acesso em: 11 ago. 2024.

SILVA, Tarcízio. Colonialidade difusa no aprendizado de máquina: camadas de opacidade algorítmica na Imagenet. In: AMADEU, Sérgio; SOUZA, Joyce; CASSINO, João F. *Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal*. São Paulo: Autonomia Literária, 2021.

_____. Visão Computacional e Racismo Algorítmico: branquitude e opacidade no aprendizado de máquina. *Revista ABPN*, v. 12, pp. 428-448, 2020.

SIMON, Zoltán Boldizsár. A Transformação do Tempo Histórico: Temporalidades Processual e Eventual. *Revista de Teoria da História*, Goiânia, v. 24, n. 1, pp. 139–155, 2021. Disponível em: <https://revistas.ufg.br/teoria/article/view/69676>. Acesso em: 11 ago. 2024.

_____. *History in Times of Unprecedented Change: A Theory for the 21st Century*. London, New York: Bloomsbury Academic, 2019.

SIMON, Zoltán Boldizsár; TAMM, Marek. HISTORICAL FUTURES. *History and Theory*, v. 60, n. 1, pp. 3–22, mar. 2021. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1111/hith.12190>. Acesso em: 11 ago. 2024.

SIMONDON, Gilbert. *La individuación a la luz de las nociones de forma e información*. Buenos Aires: Ediciones La Cebra e Editorial Cactus, 2009.

_____. *El modo de existencia de los objetos técnicos*. Buenos Aires: Prometeo Libros, 2008.

SINGER, Peter. *Famine, Affluence, and Morality*. Oxford: Oxford University Press, [1971], 2015.

SMITH, Ben. How TikTok Reads Your Mind. *The New York Times*, 5 dez. 2021. Disponível em: <https://www.nytimes.com/2021/12/05/business/media/tiktok-algorithm.html>. Acesso em: 10 jul. 2024.

SRNICEK, Nick. *Platform Capitalism*. Malden: Polity Press, 2017.

STENGERS, Isabelle. Uma outra ciência é possível. Rio de Janeiro: Bazar do Tempo, 2023.

_____. *No Tempo das Catástrofes*. São Paulo: Cosac Naify, 2015.

STENGERS, Isabelle; DEBAISE, Didier. The Insistence of Possibles: towards a speculative pragmatism. *Parse: Journal*, n. 7, p. 12-19, 2017. Disponível em:

<http://parsejournal.com/article/the-insistence-of-possibles%E2%80%A8-towards-a-speculative-pragmatism/>. Acesso em: 4 fev. 2022.

STEYERL, Hito. Mean Images. *New Left Review*, vol. 140-141, 2023. Disponível em: <https://newleftreview.org/issues/ii140/articles/hito-steyerl-mean-images>. Acesso em: 29 de ago. 2024.

_____. A Sea of Data: Pattern Recognition and Corporate Animism (Forked Version). In: APPRICH, Clemens; CHUN, Wendy H. K.; CRAMER, Florian; STEYERL, Hito. *Pattern Discrimination*. Lüneburg, Minneapolis: meson press, University of Minnesota Press, 2018.

_____. How To Kill People: A Problem of Design. *E-flux Architecture*, dezembro de 2016. Disponível em: <https://www.e-flux.com/architecture/superhumanity/68653/how-to-kill-people-a-problem-of-design/>. Acesso em: 25 ago. 2024.

STOKEL-WALKER, Chris. TikTok Wants Longer Videos—Whether You Like It or Not. *Wired*, 21 de fevereiro de 2022. Disponível em: <https://www.wired.com/story/tiktok-wants-longer-videos-like-not/>.

STRECKER, Helena. Dobras #53 // “Você É O Que Você Ouve”: Spotify E A Promessa Algorítmica De Nos Conhecer Melhor Do Que Nós Mesmos. *Blog do MediaLab.UFRJ*, maio de 2022. Disponível em: <https://medialabufRJ.net/blog/2022/05/dobras-53-voce-e-o-que-voce-ouve-spotify-e-a-promessa-algoritmica-de-nos-conhecer-melhor-do-que-nos-mesmos/>. Acesso em: 22 ago. 2024.

THALER, Richard H.; SUNSTEIN, Cass R. *Nudge: como tomar melhores decisões sobre saúde, dinheiro e felicidade*. Rio de Janeiro: Objetiva, 2019.

THE CENTRE FOR ATTENTION STUDIES. Do we have your attention? How people focus and live in the modern information environment. [Relatório de pesquisa]. Fevereiro de 2022. Disponível em: <https://www.kcl.ac.uk/policy-institute/assets/how-people-focus-and-live-in-the-modern-information-environment.pdf>. Acesso em: 20 out.

TIKTOK. Thanks a billion!. *TikTok NewsRoom*, 27 set. 2021. Disponível em: <https://newsroom.tiktok.com/en-us/1-billion-people-on-tiktok>. Acesso em: 10 jul. 2024.

_____. How TikTok recommends videos #ForYou. *TikTok Newsroom*, 18 jul. 2020. Disponível em: <https://newsroom.tiktok.com/en-us/how-tiktok-recommends-videos-for-you>. Acesso em: 10 jul. 2024.

TIQQUN. *The Cybernetic Hypothesis*. 2010. Disponível em: <https://theanarchistlibrary.org/library/tiqqun-the-cybernetic-hypothesis.pdf>. Acesso em: 4 fev. 2022.

TURING, Alan M. Computing machinery and intelligence. *Mind*, vol. LIX, issue 236, out. 1950, pp. 433–460. Disponível em: <https://doi.org/10.1093/mind/LIX.236.433>. Acesso em: 11 ago. 2024.

- TURIN, Rodrigo. Antropoceno e futuros presentes: entre regime climático e regimes de historicidade potenciais. *Topoi*, vol. 24, n. 4, pp. 703–724, 2023.
- VARIAN, Hal. Beyond Big Data. *Business Economics*, vol. 49, n. 1, pp. 27-31, março de 2014.
- VAROQUAUX, Gaël, LUCCIONI, Alexandra S., WHITTAKER, Meredith. *Hype, Sustainability, and the Price of the Bigger-is-Better Paradigm in AI*. Disponível em: <https://arxiv.org/abs/2409.14160>. Acesso em: 10 jan. 2025.
- VARON, Joana; COSTANZA-CHOCK, Sasha; GEBRU, Timnit. *Fostering a Federated AI Commons ecosystem* [Policy paper]. 2024. Disponível em: https://codingrights.org/docs/Federated_AI_Commons_ecosystem_T20Policybriefing.pdf. Acesso em: 20 ago. 2024.
- VASWANI, Ashish et al. *Attention is all you need*. Conference on Neural Information Processing Systems. 2017. Disponível em: <https://arxiv.org/abs/1706.03762>. Acesso em: 20 ago. 2024.
- VERMA, Pranshu; TAN, Shelly. A bottle of water per email: the hidden environmental costs of using AI chatbots. *The Washington Post*, 18 de setembro de 2024. Disponível em: <https://www.washingtonpost.com/technology/2024/09/18/energy-ai-use-electricity-water-data-centers/>
- VIEIRA, Carla. Inteligência Artificial e Vieses Algorítmicos - Loggi. *Medium*, set. 2020. Disponível em: <https://partiu.loggi.com/intelig%C3%A2ncia-artificial-e-vieses-algor%C3%ADtmicos-992ee4786ca>. Acesso em: 4 fev. 2022.
- VINGE, Vernor. *Technological Singularity*, 1993. Disponível em: <https://frc.ri.cmu.edu/~hpm/book98/com.ch1/vinge.singularity.html>. Acesso em: 4 fev. 2022.
- VIVEIROS DE CASTRO, Eduardo. *Máquinas sobrenaturais e outros habitantes da triplíce fronteira antropológica*. [Transcrição de palestra] XVI Seminário Internacional de Estudos de Literatura, Programa de Pós-Graduação em Literatura, Cultura e Contemporaneidade da PUC-Rio, nov. 2023. Disponível em: https://www.academia.edu/112157885/M%C3%A1quinas_sobrenaturais_e_outros_habitantes_da_tr%C3%ADplice_frenteira_antropol%C3%B3gica. Acesso em: 30 abril 2024.
- VOCK, Ido. ChatGPT proves that AI still has a racism problem. *The New Statesman*, 9 de dezembro de 2022. Disponível em: <https://www.newstatesman.com/thestaggers/2022/12/chatgpt-shows-ai-racism-problem>. Acesso em: 27 ago. 2024.
- WALL STREET JOURNAL. Inside TikTok’s Algorithm: A WSJ Video Investigation. *The Wall Street Journal*, 21 de julho de 2021. Disponível em: <https://www.wsj.com/articles/tiktok-algorithm-video-investigation-11626877477>. Acesso em: 20 de out. 2023.

- WANG, Jackie. “This is a Story About Nerds and Cops”: Predpol and Algorithmic Policing. *E-flux journal*, n. 87, 2017. Disponível em: <https://www.e-flux.com/journal/87/169043/this-is-a-story-about-nerds-and-cops-predpol-and-algorithmic-policing/>. Acesso em: 4 fev. 2022.
- WARK, Mackenzie. *O capital está morto*. São Paulo: Editora Funilaria e sobinfluência edições, 2022.
- WEST, Sarah M. Data Capitalism: Redefining the Logics of Surveillance and Privacy. *Business & Society*, vol 58, n. 1, pp. 20-41, 2017.
- WEI, Eugene. Seeing Like an Algorithm. *Remains of the Day* (blog), 18 de setembro de 2020. Disponível em: <https://www.eugenewei.com/blog/2020/9/18/seeing-like-an-algorithm>. Acesso em: 15 out. 2023.
- WIENER, Norbert. *Cibernética: ou controle e comunicação no animal e na máquina*. São Paulo: Perspectiva, [1948] 2017.
- WILLIAMS, Daniel. Predictive coding and thought. *Synthese*, n. 197, pp. 1749–1775, 2020.
- WISNIK, Guilherme. Dentro do nevoeiro: o futuro em suspensão. In: NOVAES, Adauto (org). *Mutações: o futuro não é mais o que era*. São Paulo: Edições SESC São Paulo, 2015.
- WOLFRAM, Stephen. What Is ChatGPT Doing...and Why Does It Work? Disponível em: <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/>. Acesso em: 29 ago. 2024.
- ZUBOFF, Shoshana. *A Era do Capitalismo de Vigilância: a luta por um futuro humano na nova fronteira do poder*. São Paulo: Intrínseca, 2021.