

MEDIAÇÃO ALGORÍTMICA E APRENDIZADO DE MÁQUINA

uma caracterização baseada em patentes da Google sobre técnicas de modulação de atividade

RAFAEL GONÇALVES



Universidade Estadual de Campinas (Unicamp)
Instituto de Filosofia e Ciências Humanas (IFCH)
Programa de Pós-Graduação em Sociologia (PPGS)

Dissertação de mestrado em sociologia

Mediação algorítmica e aprendizado de máquina: uma caracterização baseada em patentes da Google sobre técnicas de modulação de atividade

Rafael Gonçalves (PPGS/IFCH/Unicamp)

Pedro P. Ferreira (DS/IFCH/Unicamp, orientador)

Campinas, SP
23 de dezembro de 2024

*Às conversações,
aos encontros,
e às danças*

que atravessaram a feitura desta dissertação.

Mediação algorítmica e aprendizado de máquina: uma caracterização baseada em patentes da Google sobre técnicas de modulação de atividade

Resumo

Esta pesquisa de mestrado é uma investigação sobre as características da tradução de ações operada pela mediação de algoritmos, em particular, de técnicas de aprendizado de máquina (AdM) utilizadas para modulação de atividade patenteadas pela Google. A passagem do século XX para o XXI provocou transformações a um só tempo tecnocientíficas e político-econômicas que foram denominadas, por diferentes autoras, como "mundo de ciborgues" (Haraway), "sociedades de controle" (Deleuze) e "virada cibernética" (Garcia dos Santos). Em especial, o desenvolvimento tecnocientífico protagonizado pelo complexo industrial-militar (e, progressivamente, corporativo), sobretudo do norte global, amparou-se na noção de informação para inventar novas formas de apropriação econômica e controle político, inclusive sobre processos virtuais, antes de sua efetivação: é o caso da informação genética e digital, e do próprio sistema de propriedade intelectual. Nesse sentido, o campo da inteligência artificial e, particularmente, o AdM – mecanismo utilizado para o processamento automático de dados – ocupam um papel cada vez mais central, sendo hoje utilizados amplamente para classificação, predição e geração de dados em diversas áreas. Aliado ao fato de que a maioria das pessoas carrega consigo um smartphone a maior parte do dia, abre-se a oportunidade de utilização dessas técnicas para prever e afetar a atividade de usuárias nessas situações algorítmicamente mediadas, problema conhecido como "modulação de comportamento" ou de atividade.

Diante disso, este trabalho analisa sociotecnicamente nove técnicas patenteadas pela Google – entendida como primeira e principal empresa a aliar métodos de inteligência artificial para quantificar e abstrair aspectos da vida social com a acumulação capitalista – a fim de caracterizar a modulação de atividade operada pelo AdM. Nosso trabalho é um estudo qualitativo de patentes amparado em um instrumental teórico-metodológico da "teoria ator-rede" (Latour) e em um programa de pesquisa interessado nos "processos tecnicamente mediados de associação" (Ferreira). Nossa caracterização propõem entender o AdM como a operação simultânea de três processos: captura, imaginação e modulação. A captura descreve a criação e extração de dados que é feito de modo ativo e unilateral, abstraindo e quantificando informações digitais heterogêneas em dados, bases de dados, e parâmetros numéricos que codificam o funcionamento do AdM. A imaginação denota como o AdM funciona criando uma imagem de mundo a partir dos dados capturados que é representação, tanto quanto potencial de ação retroativa sobre o sistema de computação. O terceiro e último processo é a modulação propriamente dita, em que a imagem criada se atualiza na forma de predições comportamentais, classificações estereotipadas e controle de fluxos informacionais que moldam continuamente a atividade da usuária.

Como conclusão mais geral, propomos também o termo "mediação algorítmica" como conceito explicativo da agência de algoritmos. Em contraste com outras propostas que, por exemplo, atribuem a algoritmos algum nível de agência por meio da constatação de que eles possuem vieses; propomos considerar que todos os algoritmos agem socialmente, na medida em que, dentro de um sistema de computação, eles são os mediadores da informação digital. Desse modo, concluímos apresentando mediações algorítmicas particulares ao AdM – estereotipia, reprodução do passado e racismo e sexismo algorítmico – e apontando para a necessidade de refletirmos, para além da ideia de desviesamento, sobre a possibilidade de criação de mediações algorítmicas alternativas, que contestem a tecnologia dominante, em favor de um mundo mais diverso e politicamente responsável.

Palavras-chave: algoritmos; inteligência artificial; modulação; Google; patentes; sociologia da tecnologia.

Algorithmic mediation and machine learning: a characterization based on activity modulation techniques patented by Google

Abstract

This master's research is an investigation into the characteristics of the translation of actions operated through the mediation of algorithms, in particular machine learning (ML) techniques of activity modulation patented by Google. The transition from the 20th to the 21st century caused transformations that were both techno-scientific and political-economic. Different authors have called these transformations "world of cyborgs" (Haraway), "control societies" (Deleuze) and the "cybernetic turn" (Garcia dos Santos). In particular, the techno-scientific development spearheaded by the military-industrial complex (and, progressively, by corporations), especially in the global north, has relied on the notion of information to invent new forms of economic appropriation and political control. These include those that apply over virtual processes, that is, before they take place: examples are genetic and digital information, and the intellectual property system itself. In this sense, the field of artificial intelligence and, in particular, ML – a mechanism used for automatic data processing – occupy an increasingly central role, and is now widely used for tasks such as: classification, prediction and data generation in various areas. Combined with the fact that most people carry a smartphone with them most of the day, this opens up the opportunity to use these techniques to predict and affect user activity in these algorithmically mediated situations, a problem known as "behaviour modulation" or "activity modulation".

In view of this, this dissertation analyzes nine techniques patented by Google, that is here understood as the first and main company to combine artificial intelligence methods to quantify and abstract aspects of social life with capitalist accumulation. The analysis is made in order to characterize the modulation of activity operated by ML in these patented techniques. Our research is a qualitative investigation of patents that is based on "actor-network theory" (Latour) as our main theoretical-methodological perspective and is aligned with a research program interested in "technically mediated processes of association" (Ferreira). The characterization proposed here understands ML as the simultaneous operation of three processes: capture, imagination and modulation. Capture describes the creation and extraction of data, what is done in an active and unilateral way, abstracting and quantifying heterogeneous digital information into data, databases, and numerical parameters that encode the operation of ML. Imagination denotes how ML works by creating an image of the world from the captured data that is both representation and the potential for retroactive action on the computer system. The third and final process is modulation itself, in which the image created is actualized in the form of behavioral predictions, stereotyped classifications and control of informational flows that continually shape the user's activity.

As a broader conclusion, we also propose the term "algorithmic mediation" as an explanatory concept for the agency of algorithms. In contrast to other proposals which, for example, attribute some level of agency to algorithms by noting that they have biases, we propose that all algorithms act socially, to the extent that, within a computer system, they are the mediators of digital information. In this way, we conclude by presenting algorithmic mediations particular to ML – stereotyping, reproduction of the past and algorithmic racism and sexism – and pointing to the need to reflect, beyond the idea of de-biasing, on the possibility of creating alternative algorithmic mediations that challenge the dominant technology, in favor of a more diverse and politically responsible world.

Keywords algorithms; artificial intelligence; modulation; Google; patents; sociology of technology.

Mediación algorítmica e aprendizaje automático: una caracterización basada en patentes de Google sobre técnicas de modulación de la actividad

Resumen

Esta disertación de máster es una investigación sobre las características de la traducción de acciones operadas a través de la mediación de algoritmos, en particular las técnicas de aprendizaje automático (AA) utilizadas para modular la actividad patentadas por Google. La transición del siglo XX al XXI trajo consigo transformaciones tanto tecnocientíficas como político-económicas, que diferentes autores han denominado «mundo de los cyborgs» (Haraway), «sociedades de control» (Deleuze) y «giro cibernético» (García dos Santos). En particular, el desarrollo tecnocientífico encabezado por el complejo militar-industrial (y, progresivamente, corporativo), especialmente en el norte global, se ha apoyado en la noción de información para inventar nuevas formas de apropiación económica y control político, incluso sobre los procesos virtuales, antes de que se hagan efectivas: es el caso de la información genética y digital, y del propio sistema de propiedad intelectual. En este sentido, el campo de la inteligencia artificial y, en particular, la AA – un mecanismo utilizado para el procesamiento automático de datos – ocupan un papel cada vez más central, y en la actualidad se utilizan ampliamente para la clasificación, la predicción y la generación de datos en diversos ámbitos. Combinado con el hecho de que la mayoría de las personas llevan consigo un teléfono inteligente la mayor parte del día, esto abre la oportunidad de utilizar estas técnicas para predecir y afectar a la actividad de los usuarios en estas situaciones mediadas por algoritmos, un problema conocido como «modulación de comportamiento» o de actividad.

En vista de ello, este trabajo analiza nueve técnicas patentadas por Google – entendida como la primera y principal empresa en combinar métodos de inteligencia artificial para cuantificar y abstraer aspectos de la vida social con la acumulación capitalista – para caracterizar la modulación de la actividad operada por AA. Nuestro trabajo es un estudio cualitativo de patentes basado en un instrumento teórico-metodológico de la «teoría actor-red» (Latour) y en un programa de investigación interesado en los «procesos de asociación mediados técnicamente» (Ferreira). Nuestra caracterización propone entender AA como la operación simultánea de tres procesos: captura, imaginación y modulación. La captura describe la creación y extracción de datos, que se realiza de forma activa y unilateral, abstrayendo y cuantificando información digital heterogénea en datos, bases de datos y parámetros numéricos que codifican el funcionamiento de la AA. La imaginación denota cómo funciona la AA creando una imagen del mundo a partir de los datos capturados que es a la vez representación y potencial de acción retroactiva sobre el sistema informático. El tercer y último proceso es la modulación propiamente dicha, en la que la imagen creada se actualiza en forma de predicciones de comportamiento, clasificaciones estereotipadas y control de los flujos informativos que moldean continuamente la actividad del usuario.

Como conclusión más general, también proponemos el término «mediación algorítmica» como concepto explicativo de la agencia de los algoritmos. A diferencia de otras propuestas que, por ejemplo, atribuyen cierto nivel de agencia a los algoritmos a través de la constatación de que tienen sesgos, proponemos considerar que todos los algoritmos actúan socialmente, en la medida en que, dentro de un sistema informático, son los mediadores de la información digital. De este modo, concluimos presentando mediaciones algorítmicas propias de la AA – estereotipos, reproducción del pasado y racismo y sexismo algorítmicos – y asimismo señalamos la necesidad de reflexionar, más allá de la idea de eliminación de sesgos, sobre la posibilidad de crear mediaciones algorítmicas alternativas que desafíen la tecnología dominante en favor de un mundo más diverso y políticamente responsable. **Palabras clave:** algoritmos; inteligencia artificial; modulación; Google; patentes; sociología de la tecnología.

Lista de Figuras

1.1	Ciclo de acumulação de capital da Google: O dinheiro (M) investido em tecnologia (c) e força de trabalho pago (v_1), isto é, no salário de trabalhadoras da Google, produz serviços (P_1) que serão utilizados gratuitamente por usuárias em processos de trabalho não-pago (v_2), gerando dados e conteúdos (P_2). Esta nova mercadoria de prosumidor (C') é então vendida para anunciantes, gerando lucro (M' maior que M). Fonte: Fuchs (2012).	15
1.2	Diagrama simplificado do PageRank. A relevância de uma página (tamanho do nó no diagrama) depende da quantidade e da qualidade (função dos links que apontam para cada página) de links que apontam para ela (setas). Fonte: baseado em Pasquinelli (2009).	19
2.1	Esquema de um neurônio tal qual proposto por McCulloch e Pitts (1943) e função matemática que ele computa. A saída y é o resultado da função $g(\cdot)$ para a soma de entradas ($x_1 \cdots x_N$). Ou seja, se a soma das entradas é maior que θ , a saída é igual a um, senão, ela é igual a zero. Fonte: adaptado de Von Zuben e Attux (2007).	31
2.2	À direita, esquema de um neurônio de tipo perceptron, à esquerda, esquema de uma rede neural MLP. Cada círculo cinza na rede representa um neurônio, de modo que, conectados, a rede é capaz de computar uma função complexa. Enquanto que X_1, \dots, X_N representam as entradas da rede e Y sua saída; $x_1 \cdots x_N$ representam as entradas de cada neurônio e y suas saídas. Os pesos sinápticos ($w_1 \cdots w_N$) que multiplicam as entradas de cada neurônio são parâmetros aprendidos no processamento de dados que determinam a função a ser computada. Fonte: adaptado de Gonçalves (2022).	32
2.3	Exemplo de documento de patente (Ni 2022:1,7,11).	37
2.4	Captura de tela da ferramenta PPubS (sem data) obtida no dia 17/11/2024. Visualização de informações textuais da patente de Krishna et al. (2021).	38
2.5	Captura de tela da ferramenta PPubS (sem data) obtida no dia 17/11/2024. Visualização das figuras da patente de Krishna et al. (2021).	38
2.6	Captura de tela da ferramenta Google (sem data) obtida no dia 23/11/2024. Visualização da patente de Carbune e Feuz (2022).	39
3.1	Situação exemplo de evento em desenvolvimento. Aqui o símbolo de uma chama no canto superior direito indica um incêndio na catedral, enquanto que o bloco “FD” representa o departamento de bombeiros e os balões, a localização e o conteúdo de mensagens postadas por usuárias cujas mensagens se relacionam com o evento de incêndio. Respectivamente, da esquerda para a direita e de cima para baixo: “Há muitos caminhões de bombeiro a caminho do centro”, “Código 12345 @ state and main”, “Eu espero que os paroquianos escaparam do incêndio...”, “Meu deus! Há um incêndio na catedral!”, “Deve haver dez caminhões de bombeiro lutando contra o incêndio da igreja”, “Alguém sabe por que há fumaça do outro lado do rio?”. Fonte: Carbune e Feuz (2022).	63

3.2	Exemplos do funcionamento dos métodos de resolução de conjunto de atividades. Em ambos os casos, há uma conversa em grupo do ponto de vista de Alice, que conversa com Bob e Chris (apenas na imagem da esquerda). À esquerda, podemos ver que a conversa é interrompida para sugerir atividades que devem ser aceitas ou recusadas por Alice (310 e 315). À direita, a conversa é interrompida para que Alice também inclua explicitamente uma data e horário para o agendamento de uma atividade (415). Fonte: Lance e Koonce (2021).	65
3.3	Esquema do processo de captura de informação digital. Fonte: elaboração própria.	70
4.1	Exemplo esquemático de uma rede neural utilizada para classificar imagens visuais (em um computador digital representadas como matrizes de pixels RGB – vermelho, verde e azul) de cachorros e de gatos. São representados os dois momentos do algoritmo: treinamento e inferência. Na figura, W se refere aos parâmetros numéricos que caracterizam a saída da rede neural, enquanto que $J(.)$ é a função objetivo, métrica utilizada para comparar o resultado real e esperado na etapa de treinamento. O rótulo é o resultado esperado para a saída do modelo em cada iteração do treinamento, isto é, zero ou um. Fonte: Gonçalves (2022).	93
4.2	Domínio (X) em vermelho, contradomínio (Y) em azul e imagem ($Im(f)$) em amarelo. Fonte: elaboração própria, baseado em https://commons.wikimedia.org/wiki/File:Codomain2.SVG (acesso: 02/07/2024).	93
4.3	Representação bidimensional (via t-SNE, algoritmo de redução de dimensionalidade) de vetores de palavras multidimensionais de uma rede neural de modelagem de linguagem. Os eixos X e Y são duas dimensões numéricas abstratas. Para nós, o que importa desta representação é que uma proximidade entre palavras no gráfico indica uma proximidade entre os valores numéricos dos vetores que representam essas palavras no espaço latente do modelo. Desse modo, é possível perceber que existe uma relação entre uma proximidade numérica e uma proximidade semântica. Fonte: Le Cun, Bengio e Hinton (2015).	94
4.4	Visualização de duas camadas internas de uma rede neural convolucional e da projeção do espaço latente na forma de pixels via desconvolução. A ideia implícita neste trabalho é mostrar que o espaço latente nas redes neurais convolucionais registraria informações visuais que vão de estruturas mais simples (linhas e degradês na primeira camada) até estruturas mais complexas (quinas e curvas na segunda camada). Fonte: Zeiler e Fergus (2014).	94
4.5	Manifold para uma rede neural de tipo autocodificador treinada na base de dados MNIST. Cada imagem representa a saída produzida pelo modelo após um deslocamento de posição no espaço latente de um autocodificador treinado para gerar dígitos escritos a mão. A posição de um resultado na grade do manifold se relaciona com sua posição relativa no espaço latente. Percebe-se como há uma continuidade do espaço que, por sua vez, captura informação: características estruturais dos dígitos escritos a mão. Assim, os dígitos gerados na primeira fileira vão de representações claras do número 6 até o 5 passando por formas intermediárias que não são claramente um 6 ou um 5. Fonte: gerado pelo autor.	95
4.6	Redução de dimensionalidade via t-SNE dos dígitos da base de dados MNIST. Respectivamente: zeros (vermelho), seis (azul), dois (amarelo), cinco (azul claro), três (verde claro), uns (laranja), setes (roxo), oitos (lilás), nove (magenta) e quatro (verde). Fonte: Olah (2014b).	96
4.7	Exemplo de dois conjuntos não separáveis linearmente (esquerda), de uma transformação não-linear no conjunto que o faz separável linearmente (centro) e do resultado dessa transformação como uma classificação não-linear no conjunto de entrada (direita). Desse modo, a operação de um algoritmo de aprendizado de máquina não-linear pode ser visto tanto como uma classificação não-linear quanto uma deformação não-linear de um espaço que passa a ser linearmente separável. No AdM, o treinamento visa calcular números que vão parametrizar essa transformação. Fonte: Olah (2014a).	96
4.8	Exemplos animados (gif1,gif2) de um caso de sucesso e de falha na separação dos dados por transformação não-linear. Fonte: Olah (2014a).	97

4.9	Exemplo animado (gif) de um caso de falha de separação dos dados em um plano bidimensional (esquerda), exemplo estático de sua separação em três dimensões (centro) e exemplo animado (gif) do sucesso de separação observando um corte bidimensional do caso tridimensional. Fonte: Olah (2014a).	97
4.10	Exemplos do funcionamento dos métodos de sugestão de mensagens. À esquerda, um fluxo de mensagens entre Tony e Greg é apresentado como série histórica a ser captada e utilizada pelo método da patente. À direita, as mensagens já processadas proveem sugestões à Greg para a mensagem de Tony. Fonte: Braun e Patel (2022).	101
5.1	Respectivamente: gif animado comparando a modulação por amplitude e por frequência de um sinal de entrada; modulação por amplitude e modulação por frequência. Fontes: https://commons.wikimedia.org/wiki/File:Amfm3-en-de.gif , https://commons.wikimedia.org/wiki/File:Amplitude-modulation.png e https://commons.wikimedia.org/wiki/File:Frequency-modulation.png (adaptado) (acesso: 10/06/2024).	112
5.2	Exemplos do funcionamento dos métodos de sugestão de ação. Fonte: Krishna et al. (2021).	118
5.3	Esquema do processo de modulação de atividade	121
5.4	Esquema geral da captura, imaginação e modulação	121
5.5	Mapa da aprendizagem de máquina (visualizar online). Fonte: elaboração própria.	123
7.1	Três momentos de um algoritmo: (a) como regras matemáticas abstratas; (b) como código-fonte ou programa escrito; (c) como operações materiais em um aplicativo ou software. Entre eles há ainda traduções ativas entre um modo e outro, embora geralmente o termo algoritmo se refira aos três momentos de forma indistinta.	139
8.1	Exemplo fictício de gráfico de treinamento (acima), a ordenada mede o erro, enquanto que a abcissa representa o tempo de treinamento (número de épocas). São mostradas as zonas de sobajuste (A), ponto ótimo (B) e zona de sobreajuste (C). Quanto maior a distância entre o erro de treinamento e o de validação, isto é, de uma inferência realizada para testar o modelo, menor a capacidade de generalização do modelo. Abaixo, um exemplo fictício de resultado de uma função para os três casos (A, B e C): o sobajuste resulta em um resultado mais estereotipado, mas um sobreajuste projeta interpolações e extrapolações muito pouco estáveis, o que faz com que o modelo não acerte exemplos que não foram utilizados no treinamento. Fonte: elaboração própria, baseada em Goodfellow et al. (2016:111, 113, 118, figuras 5.2, 5.3 e 5.5).	157
8.2	Execução do modelo gerativo Craiyon v4 para o prompt “doctor”.	163
8.3	Execução do modelo gerativo Craiyon v4 para o prompt “nurse”.	164
8.4	Execução do modelo gerativo Craiyon v4 para o prompt “couple”.	165

Lista de Tabelas

2.1	Lista de patentes analisadas	40
3.1	Elementos de captura encontrados nas patentes.	53
3.2	Comparação dos diferentes aparelhos de captura, ênfase na nossa proposta de aparelho de captura de informação digital.	61
3.3	Exemplo fictício de uma tabela de dados com a estrutura: ID (identificação ou número do exemplo ou dado), localização, acelerômetro, giroscópio, última mensagem de texto e se aceitou ou não determinada sugestão.	68
5.1	Elementos da tipologia de modulação proposto em Machado (2019): M1 – “Recomendação que leva a ação”; M2 – “Decidindo pelo usuário”; M3 – Predição, antecipação e inferência; M4 – “Atribuindo valor à interação”; M5 – “Análise de emoções” e M6 – “Possibilidades para coleta de dados”.	115

Lista de Abreviaturas e Siglas

ACM *Association for Computing Machinery* [Associação para a Maquinaria Computacional].

AdM aprendizado de máquina.

AE *autoencoder* [autocodificador].

AGR *automatic gender recognition* [reconhecimento automático de gênero].

AM *amplitude modulation* [modulação por amplitude].

BBX Baobáxia.

BERT *Bidirecional Encoder Representations from Transformers* [Codificador Bidirecional de Representações a partir de Transformers].

CIA *Central Intelligence Agency* [Agência Central de Inteligência].

DARPA *Defense Advanced Research Projects Agency* [Agência de Projetos de Pesquisa Avançada de Defesa].

DRM *Digital Rights Management* [Gerenciamento de Direitos Digitais].

EHT *Event Horizon Telescope* [Telescópio do Horizonte de Eventos].

EUA Estados Unidos da América.

FM *frequency modulation* [modulação por frequência].

GAFAM Google/Alphabet, Amazon, Facebook/Meta, Apple e Microsoft.

GAN *generative adversarial network* [rede adversarial gerativa].

GPT *Generative Pre-trained Transformer* [Transformer Gerativo Pré-treinado].

GrEGS Grupo de Estudos Gilbert Simondon.

IA inteligência artificial.

IAg inteligência artificial gerativa.

IoT *internet of things* [internet das coisas].

ISP *internet service provider* [provedor de serviço internet].

LaSPA Laboratório de Sociologia dos Processos de Associação.

LSTM *long short-term memory* [memória de curto longo prazo].

m2m *machine-to-machine* [máquina-para-máquina].

MdL modelo de linguagem.

MLP *multilayer perceptron* [perceptron de múltiplas camadas].

MNIST *Modified National Institute of Standards and Technology* [Instituto Nacional de Padrões e Tecnologia Modificado].

NASA *National Aeronautics and Space Administration* [Administração Nacional da Aeronáutica e Espaço].

NGA *National Geospatial-Intelligence Agency* [Agência Nacional de Informação Geoespacial].

NLP *natural language processing* [processamento de linguagem natural].

NSA *National Security Agency* [Agência de Segurança Nacional].

PatFT *Patent Full-text Databases* [Bases de dados de Texto Completo de Patente].

PPB *Pilot Parliaments Benchmark* [Referência Piloto de Parlamentares].

PPubS *Patent Public Search* [Busca Pública de Patente].

RLHF *reinforcement learning from human feedback* [aprendizado por reforço com reatualimentação humana].

RNA rede neural artificial.

SPI sistema de propriedade intelectual.

STS *science and technology studies* [estudos de ciência e tecnologia].

SVM *support vector machines* [máquinas de vetores de suporte].

SWPA *Sony World Photography Award* [Prêmio Mundial de Fotografia da Sony].

UI *user interface* [interface de usuário].

UNESCO *United Nations Educational, Scientific and Cultural Organization* [Organização das Nações Unidas para a Educação, a Ciência e a Cultura].

USPTO *United States Patents and Trademark Office* [Escritório de Patentes e Marcas Registradas dos Estados Unidos].

Sumário

	Resumos	iv
	Lista de Figuras	vii
	Lista de Tabelas	x
	Lista de Abreviaturas e Siglas	xi
	Apresentação	1
Parte I	Introdução: tecnociência e capitalismo contemporâneo	9
Capítulo 1	A Google como empresa e como tecnologia	10
Capítulo 2	Propriedade intelectual, sociologia de patentes e os algoritmos de modulação de atividade	24
Parte II	Uma cartografia do aprendizado de máquina	44
Capítulo 3	Captura	45
Capítulo 4	Imaginação	71
Capítulo 5	Modulação	106
Parte III	Mediação algorítmica	127
Capítulo 6	Do aprendizado de máquina sobre a ação social à ação social do aprendizado de máquina	128
Capítulo 7	O que é um algoritmo? por uma definição de mediação algorítmica	137
Capítulo 8	Mediações algorítmicas na aprendizagem de máquina	153
	Considerações finais	166
	Referências	173
	Apêndices	192
Apêndice A	Lista de título e resumo das patentes	192
Apêndice B	Breve cronologia da IA	196
	Índice completo de conteúdo	197

Apresentação

Esta dissertação de mestrado versa sobre o problema específico da modulação de atividade de usuárias por técnicas de aprendizado de máquina (AdM)¹ em sistemas digitais e sobre o tema mais geral da agência algorítmica e da possibilidade de uma descrição sociológica desta. Ela é o resultado do projeto de pesquisa de mestrado inicialmente intitulado *Mediação algorítmica nas sociedades de controle: caracterização da modulação de comportamento de usuárias por meio de técnicas de aprendizado de máquina em patentes da Google* apresentada ao Programa de Pós-Graduação de Sociologia (PPGS) do Instituto de Filosofia e Ciências Humanas (IFCH) da Universidade Estadual de Campinas (Unicamp) e financiada pelo processo no. 2023/01858-0, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), que teve início em março de 2023.

Dividimos o presente trabalho em três partes, para além desta apresentação inicial. Aqui, incluímos: esta apresentação geral da dissertação; um breve relato de nossa trajetória; o problema sobre o qual nos propomos dissertar; e uma exposição inicial do nosso argumento principal. A primeira parte, [Introdução: tecnociência e capitalismo contemporâneo](#), serve de introdução ao trabalho, que opera na interface entre ciências sociais e engenharia. O primeiro capítulo, [A Google como empresa e como tecnologia](#), apresenta a Google e utiliza a sua origem com a invenção do PageRank como caso emblemático da intersecção entre tecnociência e capitalismo contemporâneo, de onde nascem os predicados de nosso objeto, em especial aqueles relativos ao sistema de propriedade intelectual (SPI) e aos algoritmos. O segundo capítulo, [Propriedade intelectual, sociologia de patentes e os algoritmos de modulação de atividade](#), apresenta brevemente algumas definições preliminares, a metodologia que empregamos, isto é, a análise qualitativa de patentes, e o material, patentes da Google sobre algoritmos de modulação de atividade. A segunda parte, [Uma cartografia do aprendizado de máquina](#), é o núcleo empírico de nossa pesquisa, onde descrevemos propriamente os mecanismos sociotécnicos operantes nos algoritmos de AdM patenteados, focando em três momentos: captura (capítulo 3), imaginação (capítulo 4) e modulação (capítulo 5). A conclusão dessa parte desemboca em um experimento cartográfico de caracterização do AdM. A terceira e última parte, [Mediação algorítmica](#), visa sintetizar o debate teórico feito ao longo do trabalho, propondo uma exposição inicial sobre a questão da agência de algoritmos baseada no debate ao redor do “viés” (capítulo 6), uma definição alternativa de algoritmo e de sua agência (capítulo 7) e uma apresentação de alguns casos específicos de mediação algorítmica operada pelo AdM (capítulo 8). Finalmente, concluímos o trabalho com algumas [Considerações finais](#) que retomam o argumento apre-

¹O termo em inglês, *machine learning*, é diversificadamente traduzido por “aprendizado de máquina” e “aprendizagem de máquina” em português. Decidimos manter o primeiro no título, pois seu uso é mais corrente (uma busca por ambos os termos no Google Search realizada no dia 25 de novembro de 2024 apontou “aproximadamente 3.340.000 resultados” para “aprendizado de máquina” e “aproximadamente 213.000 resultados” para “aprendizagem de máquina”). Entretanto, no corpo do texto, utilizaremos ambas as grafias, dado que “aprendizagem de máquina” explicita a natureza processual do algoritmo que será por nós explorada.

sentado na dissertação e apontam para possibilidades alternativas ao modelo tecnocientífico dominante em relação aos algoritmos e ao AdM.

Escolhemos como capa do trabalho uma representação gráfica de um espaço latente² de um algoritmo de AdM treinado sobre a base de dados de dígitos escritos a mão Instituto Nacional de Padrões e Tecnologia Modificado (MNIST, *Modified National Institute of Standards and Technology*). Essa base de dados é composta por pequenas imagens em preto e branco de dígitos escritos por estudantes do ensino médio e funcionários do Departamento do Censo dos Estados Unidos da América (EUA)³. Ela é frequentemente utilizada para a avaliação de algoritmos de classificação de imagens, podendo ser considerada “canônica” e de reconhecimento imediato por alguém da área. Essa imagem foi criada com um modelo autocodificador (AE, *autoencoder*), que é utilizado para reduzir a dimensionalidade de dados, criando um espaço latente de codificação com um aspecto de continuidade que é chamado, com referência às estruturas topológicas da matemática (como a esfera, o toro e a garrafa de Klein), de “manifold”, de difícil tradução para o português, sendo por vezes traduzido como “variedade”, mas podendo ser traduzido também por “multiplicidade”, entre outros. Se, por um lado, esta imagem ilustra as notáveis capacidades que observamos no AdM em sua extrapolação que gera dígitos perceptivelmente intermediários em relação aos algarismos que conhecemos, por outro, nos instiga a pensar como o campo da inteligência artificial (IA) progrediu vertiginosamente nos últimos anos, passando da classificação e geração de dígitos, para a geração de imagens e textos verossímeis, a classificação precisa de imagens que possibilita a condução de um carro autônomo e, finalmente, a modulação de atividade de usuárias online.

Neste trabalho, optamos por utilizar a flexão para o plural masculino ao se referir ao autor e a flexão para o feminino como gênero neutro. O primeiro se justifica, por um lado, por identificar o gênero gramatical com o gênero do autor, e, por outro lado, por entender que toda produção é coletiva, fruto de uma multiplicidade de enunciações, de muitos encontros. As poucas exceções são para registros que consideramos muito majoritariamente biográficos, como na seção logo abaixo, em que preferimos manter o singular. Já o segundo, visa chamar atenção para o aspecto inerentemente político da linguagem. Embora saibamos que utilizar o feminino não supre todos os problemas ligados ao binarismo de gênero, estereotipia e sexismo incorporados no português brasileiro, acreditamos que essa escolha é um passo nessa direção por direcionar a atenção, tanto da leitora quanto do autor, às escolhas relativas a gênero naturalizadas na nossa língua. É, portanto, uma escolha política tanto quanto metodológica.

Este trabalho foi escrito, sempre que possível, utilizando software livre. Sua escrita foi feita nos editores de texto Vim e Neovim⁴, utilizando as linguagens de marcação de texto abertas Markdown e \LaTeX ⁵ e compilado para pdf utilizando o programa Pandoc⁶. A escrita foi feita em dois ambientes GNU/Linux: um computador de mesa antigo com o sistema operacional Arch Linux e um notebook ThinkPad T14 com Ubuntu⁷ instalado via Windows Subsystem for Linux (WSL). As imagens foram geradas, sempre que possível, utilizando o webapp de software livre Penpot⁸. Referências foram geridas

²Essa discussão será o centro do capítulo 4.

³Disponível em: <https://yann.lecun.com/exdb/mnist/> (acesso: 19/11/2024).

⁴Respectivamente <https://www.vim.org/> e <https://neovim.io/> (acesso: 14/11/2024).

⁵Respectivamente <https://www.markdownguide.org/> e <https://www.latex-project.org/> (acesso: 14/11/2024).

⁶<https://pandoc.org/> (acesso: 14/11/2024).

⁷Respectivamente <https://archlinux.org/> e <https://ubuntu.com/> (acesso: 14/11/2024).

⁸<https://penpot.app/> (acesso: 14/11/2024).

com o software livre Zotero⁹. Durante a pesquisa, o código-fonte e os arquivos brutos desta dissertação foram armazenados no repositório de código da plataforma software livre Gitlab¹⁰. Ao longo da dissertação, incluímos animações (arquivos .gif) integradas ao pdf: apenas alguns softwares possibilitam a visualização delas no pdf, como o caso do leitor de pdf software livre Okular¹¹, entretanto, providenciamos links para que os mesmos sejam facilmente acessados por um navegador de internet no caso de utilização de outro leitor de pdf. Finalmente, a versão mais atualizada desta dissertação pode ser encontrada em <https://rafaelg.net.br/dissertacao.pdf>, disponibilizada como arquivo em minha página pessoal que está disponível pela funcionalidade de hospedagem gratuita de sites estáticos oferecido pela plataforma Gitlab¹².

0.1 Sobre o autor: explicitação de uma perspectiva parcial

Várias autoras têm enfatizado recentemente como a biografia de um sujeito, sua história e a materialidade de seu corpo estão diretamente implicados com o conteúdo do conhecimento produzido. Donna Haraway (1988) é uma dessas autoras que apresenta uma interessante reflexão sobre as noções de “objetividade” e de “responsabilidade” no contexto do debate sobre a possibilidade de uma ciência feminista. Seu desafio era construir uma noção de objetividade que seria capaz de levar em conta simultaneamente uma visão da radical contingência histórica dos conhecimentos, uma prática crítica para a construção de significados e uma noção significativa de “real” (Haraway 2023a). Fugindo tanto do relativismo absoluto do construtivismo radical que leva a um entendimento solipsista do mundo, quanto aos essencialismos que afirmam uma perspectiva universal – o “truque divino” que postula a possibilidade de observar a realidade de uma perspectiva privilegiada tanto quanto desmaterializada e, portanto, tão absoluta, quanto vazia – Haraway insiste em uma objetividade encarnada: “objetividade feminista significa simplesmente *saberes situados*” (Haraway 2023a:329). Essa perspectiva leva a uma aproximação necessária da situacionalidade do conhecimento com suas implicações políticas, o que a autora denomina “responsabilidade”: “não é possível se realocar a partir de um ponto privilegiado possível [*possible vantage point*] sem se responsabilizar por esse movimento”, pois “[a] visão é *sempre* uma questão de poder enxergar – e talvez da violência implícita em nossas práticas de visualização” (Haraway 2023a:336). Seu argumento desemboca na centralidade do “se posicionar”, que implica assumir responsabilidade por nossas práticas de produção de conhecimento. Haraway defende uma visão “a partir de um corpo” contra a “visão a partir de cima, de lugar nenhum” (Haraway 2023a:341). Uma “perspectiva parcial”, um “conhecimento situado”. Busquei levar seus apontamentos comigo durante toda a escrita da dissertação, de modo que a pesquisa levada a cabo por mim não tem pretensão de neutralidade e de universalidade, mas, antes, de parcialidade e, portanto, de objetividade, tanto quanto de responsabilidade (nos termos de Haraway). Além disso, achei importante apresentar aqui uma breve trajetória que busca me ajudar a explicitar para a leitora a situação do conhecimento por mim aqui apresentado.

Nasci e cresci em Limeira, uma cidade do interior do estado de São Paulo, tendo desde cedo me interessado simultaneamente por diversas áreas do conhecimento (como a matemática, a física, as hu-

⁹<https://www.zotero.org/> (acesso: 14/11/2024).

¹⁰<https://gitlab.com/> (acesso: 14/11/2024).

¹¹<https://okular.kde.org/> (acesso: 16/12/2024).

¹²Aproveito para mencionar que mantenho parte dos textos e fichamentos que são bastidores da pesquisa de mestrado que condicionou esta dissertação publicada na página <https://maquinacoes.rafaelg.net.br>.

manidades e as artes). Em 2016 ingressei no curso de graduação em engenharia elétrica na Faculdade de Engenharia Elétrica e de Computação (FEEC) da Unicamp. Desde então, moro em Campinas. Concluí o curso de graduação em 2022, após o período conturbado de pandemia de COVID-19. Durante o curso, a FEEC se mostrou um lugar especialmente adequado para os interesses que o curso me suscitariam, pois, além do curso de engenharia elétrica, a FEEC congrega o curso de engenharia da computação na modalidade “sistemas e processos”, incluindo portanto, pesquisadoras interessadas na computação. Durante o curso, aproximei-me progressivamente dessa área e, depois, da engenharia de sistemas que denominamos amplamente de “inteligência artificial”. Nesse espírito, cursei, além das disciplinas de graduação obrigatórias para o curso, algumas disciplinas de pós-graduação¹³, realizando também uma pesquisa de iniciação científica sobre os algoritmos de AdM denominados “redes neurais artificiais (RNAs)” sob a orientação de Romis R. F. Attux¹⁴, a quem serei sempre grato por ter me iniciado na pesquisa. Antes de finalizar o curso, também fiz um estágio de dois anos em que trabalhei como desenvolvedor web e designer de interface (UI) na equipe de tecnologias de informação da empresa de tecnologias educacionais digitais Clickidea. Finalmente, ainda no que diz respeito à computação, aproximei-me do movimento software livre e da cultura hacker, onde conheci o Laboratório Hacker de Campinas (LHC)¹⁵ e o grupo de estudos de projetos de software livre LKCamp¹⁶, tendo feito pequenas contribuições de programação e participando da construção, na FEEC, do Laboratório de Experimentação, Desenvolvimento e Inovação Tecnológica (LED-IT).

Entretanto, ainda na graduação em engenharia, interessei-me também pela aproximação possível entre meu curso e as humanidades¹⁷. Nesse espírito, desde meu ano de ingresso, procurei me aproximar de debates sobre “engenharia e sociedade”¹⁸ e cursei diversas disciplinas de sociologia no IFCH¹⁹. A primeira, e também a mais marcante, foi “HZ291 - Tópicos especiais em humanidades I”, disciplina de serviço do IFCH, ministrada majoritariamente por um estagiário docente (PED), na ocasião, o professor, no melhor sentido da palavra, Guilherme Flynn Paciornik. Nela, tive contato pela primeira vez com a sociologia da tecnologia, tendo lido textos, entre outros, de Laymert Garcia dos Santos, Pedro Ferreira, Gilles Deleuze e Gilbert Simondon. Acredito que esse encontro foi determinante para tudo o que fiz

¹³Os temas das disciplinas foram: “aprendizado de máquina” (2019); “aprendizado profundo aplicado ao processamento de linguagem natural” (2020); “aprendizado profundo para computação musical” (2021); “Paulo Freire e a educação em engenharia” (2021); “processos estocásticos para engenharia” (2021); e “teoria da informação” (2022).

¹⁴O trabalho síntese dessa pesquisa foi apresentado no “XXVII Congresso de Iniciação Científica Unicamp” (Gonçalves e Attux 2019).

¹⁵<https://lhc.net.br/> (acesso: 14/11/2024).

¹⁶<https://lkcamp.dev/> (acesso: 14/11/2024).

¹⁷Acredito que soma-se a isso a minha participação, também desde o primeiro ano da graduação, no movimento estudantil da Unicamp, tendo participado do Centro Acadêmico Bernardo Sayão (CABS), entidade representativa dos estudantes de graduação em engenharia elétrica, interagido constantemente com o Diretório Central de Estudantes (DCE) e assumido cargos de representação discente junto à instituição. Essas experiências permitiram um deslocamento da tendência tecnicista e academicista que uma ou um estudante de engenharia pode sentir ao frequentar o curso somente nos espaços de ensino em que, quase sempre, a técnica é apresentada abstraída da sociedade, e a vida acadêmica, da política.

¹⁸Principalmente, participei da organização, na FEEC, de uma disciplina sobre o tema “engenharia e sociedade”, oferecida por Romis R. F. Attux, experiência que me aproximou de temas como “tecnologia social”, “economia solidária” e “engenharia popular”, e de coletivos como a Incubadora Tecnológica de Cooperativas Populares (ITCP) da Unicamp e a Rede de Engenharia Popular Oswaldo Sevá (REPOS).

¹⁹Os temas das principais disciplinas foram: “humanidades para engenharia” (2016); “sociologia de Marx” (2018); “sociologia de Durkheim” (2019); “sociologia de Weber” (2020); “sociologia da tecnologia” (2022); “pensamento social do Brasil” (2022) e “sociologia contemporânea” (2022).

depois. Comecei meu diálogo com Pedro, que desembocaria na orientação deste mestrado, em 2019 quando, em uma disciplina sobre Durkheim, ante ao meu interesse no que as humanidades tinham a dizer da tecnologia, recomendou-me que lesse Gilbert Simondon e Bruno Latour. Depois, fiz duas iniciações científicas sob sua orientação, respectivamente sobre viés de gênero no modelo de linguagem (MdL) BERT/Google (Gonçalves e Ferreira 2022a) e sobre discurso neoliberal em artigos científicos da Google sobre AdM (Gonçalves e Ferreira 2022b). Além disso, entrei, em 2020, no Grupo de Estudos Gilbert Simondon (GrEGS) e, em 2021, no Laboratório de Sociologia dos Processos de Associação (LaSPA), ambos coordenados por Pedro P. Ferreira e dos quais ainda faço parte.

Meu trabalho de conclusão de curso em engenharia elétrica, *Redes neurais artificiais: uma introdução indisciplinar* (não publicado), orientado por Romis R. F. Attux e coorientado por Pedro P. Ferreira, visou ser uma dupla introdução: à socióloga, do funcionamento técnico das RNAs; e à engenheira, do funcionamento social do mesmo algoritmo. Ele pode ser entendido, então, como uma tentativa de síntese e convergência de ambos os meus interesses durante o curso.

Ao ingressar no mestrado, em março de 2023, percorri uma trajetória heterogênea, seguindo o que já fizera anteriormente: perseguir mais um objeto (as tecnologias digitais) que uma área disciplinar. Assim, cursei disciplinas da sociologia no IFCH – “Seminário de dissertação” (ministrada por Bárbara Castro, em 2023), “Teoria sociológica I” (Fábio Querido, 2023), “Sociologia do trabalho II” (Ricardo Antunes, 2023) e “Sociologia da ciência” (Pedro Ferreira, 2024) –; do mestrado em divulgação científica e cultural no Laboratório de Estudos Avançados em Jornalismo (Labjor) – “Estudos culturais da ciência” (Daniela Manica, 2023) e “Sócio-antropologia da ciência e da tecnologia” (Rafael Evangelista, 2023) –; e das artes visuais no Instituto de Artes – “Imagens, meios e conhecimento” (Cesar Baio, 2023). Além disso, participei, na condição de estudante especial, da disciplina “Tecnoceno, virada cibernética e conflitos cosmotécnicos”, oferecida por Henrique Z. Parra na Universidade Federal do Estado de São Paulo (Unifesp) em 2023. A estas docentes, e em especial a Bárbara Castro e Marta Kanashiro, que fizeram contribuições importantíssimas na ocasião de exame de qualificação do meu mestrado, registro meus agradecimentos.

Entendo que essa trajetória híbrida, na intersecção entre ciências sociais e engenharia, informou fortemente a escolha do projeto de pesquisa e os desenvolvimentos do mestrado, de modo que eu acredito que tanto as falhas quanto as contribuições que este trabalho porventura possua podem ser melhor entendidas à luz do caminho aqui brevemente esboçado.

0.2 Posição inicial do problema

O desenvolvimento das tecnologias digitais em convergência com novos modos de acumulação capitalista, a proliferação de dispositivos computacionais em nossa vida cotidiana e o aumento significativo de situações mediadas por algoritmos fez com que descrever as formas sociotécnicas de funcionamento dessas novas tecnologias se tornasse uma tarefa cada vez mais importante para o entendimento de nossa vida social contemporânea. Em especial, a profusão de sistemas e do discurso ao redor da chamada “inteligência artificial” – denominação ampla de um campo tecnocientífico, econômico e cultural – e particularmente do aprendizado de máquina – inicialmente entendido como conjunto de técnicas automatizadas de modelagem e inferência estatística a partir de dados – fazem destes, elementos centrais da sociedade contemporânea (cf. capítulo 2, particularmente seção 2.2 para uma exposição de tal tecnologia). Nesse sentido, o presente trabalho tem como principal objetivo entender como algorit-

mos de AdM mediam interações sociais participando da atividade resultante de usuárias que com eles interagem. É o problema que parte da literatura da sociologia da tecnologia denomina “modulação de comportamento”²⁰. No discurso coletivo, é também o que por vezes é intuitivamente considerado como a ação d’o algoritmo”. Nesse sentido, vale ressaltar que nosso esforço vai na direção de lançar luz sobre uma problemática sociológica muito mais ampla, a saber, a da agência ou ação sociotécnica²¹ dos algoritmos. Para isso, realizamos nossa análise sobre um *corpus* composto por nove patentes da empresa Google, considerada aqui como pioneira e mais importante agente a unir AdM e novas estratégias de acumulação capitalista (ver capítulo 1). Em outras palavras, buscamos especificamente *caracterizar como técnicas de aprendizado de máquina patenteadas pela Google são mobilizadas na modulação de atividade de usuárias*. Para que isso fosse possível, foi necessário: 1) identificar o uso de tais técnicas nas patentes da empresa, 2) entender e descrever como o AdM é utilizado para a modulação de atividade nessas patentes e 3) relacionar o funcionamento observado com uma extensa bibliografia sobre a posição que as novas tecnologias ocupam no capitalismo contemporâneo, a fim de explicitar o que há de específico no aprendizado de máquina.

Tendo isso em vista, consideramos importante destacar que ao menos dois acontecimentos foram de suma importância para a construção do problema sobre o qual este visa dissertar:

(1)

A publicação de um post, em seu blog, por Alexander Galloway (2019) que faz uma provocação sobre a realidade inerentemente política de algoritmos, construções matemáticas e linguagens de programação, lançando a provocativa hipótese da existência de “vieses inerentes” a tais construções:

Um tópico que tem chamado minha atenção já faz vários anos – embora eu só tenha escrito sobre isso episódica e tangencialmente – é o modo como software (incluindo sua matemática e sua lógica) pode em si mesmo ser sexista, racista ou classista. E eu não quero dizer sobre os *usos* do software. Uso é muito óbvio; nós já sabemos essa resposta. Eu quero dizer números como 5 ou 7. Ou a variável x . Ou uma estrutura de controle *se/então* [if/then]. Ou uma linguagem de computação inteira como C++ ou Python. Essas coisas contêm vieses inerentes? Elas poderiam ser sexistas ou racistas?

A inquietação explicitada nesse parágrafo ressoara com aquela implícita nos incômodos que carregávamos do nosso próprio processo de aprendizado sobre essas técnicas e que, primeiramente, levou-nos a buscar na sociologia um complemento necessário para a engenharia. Por outro lado, qualquer resolução dessa questão que não apelasse a uma análise puramente discursiva, ou à redução do objeto técnico à matéria totalmente inerte resultada da ação humana, implicaria passar obrigatoriamente pela

²⁰Por exemplo, Machado (2019), importante referência para nosso trabalho. Neste trabalho, entretanto, optamos por priorizar o uso do termo “atividade” sobre “comportamento”, pois o segundo nos remete a uma literatura mais ligada à psicologia e às ciências cognitivas, enquanto que buscamos analisar justamente o que há de social, isto é, do comportamento coletivo que resulta de associações, e não o caráter psicológico do fenômeno.

²¹Como esperamos que fique mais claro abaixo, entendemos o funcionamento e a agência do AdM como a um só tempo sociais e técnicos. Nesse sentido, utilizaremos alternadamente ação sociotécnica, ação técnica e ação social para nos referirmos ao mesmo fenômeno, sendo o adjetivo apenas índice do nosso foco em determinada ocasião. Desse modo, “ação social” não se refere ao conceito weberiano, mas sim a um entendimento de que quanto mais técnico é algo, mais social ele o é, e vice e versa (Latour 2011).

espinhosa questão da agência de algoritmos e objetos técnicos e, em última instância, às possibilidades de uma antropologia e sociologia dos “não humanos”, para utilizar o termo de Latour (2019).

Em relação ao primeiro primeiro ponto, acreditamos que uma pesquisa é fruto não só da compatibilidade teórica em relação a um campo minimamente estabelecido e da relevância social de sua execução, mas também do fato de uma pesquisadora ter sido efetivamente afetada por uma questão, pela existência de um demônio que a inquieta, de desejo que a move. No nosso caso, o fato de termos cursado uma graduação em engenharia elétrica (entre 2016 e 2022), realizado uma pesquisa de graduação sobre um algoritmo de AdM (entre 2018 e 2019) e estagiado na equipe de tecnologias da informação de uma empresa privada (entre 2020 e 2022), desembocou na seguinte inquietação: *como seria possível que, a despeito das óbvias alterações que a invenção e desenvolvimento de tecnologia promoveriam nas relações sociais, se estude engenharia, pesquise algoritmos e construa tecnologia sem que se leve em conta as causas e efeitos sociais desses processos?* Em outras palavras, incomodava-nos como algumas professoras, estudantes e desenvolvedoras não pareciam se sentir responsáveis, agindo – conscientemente ou não – como se a tecnologia fosse *neutra* ou desprovida de conteúdo social. Desejávamos fazer outra coisa.

O caminho que encontramos para elaborar esse problema dentro da sociologia foi, logo de início, no tensionamento de algumas certezas que, ao longo da pesquisa pudemos constatar, são constituintes de uma razão colonial, moderna, ocidental, eurocentrada e capitalista. Bruno Latour (2019, originalmente publicado em 1991) – amparado em uma tradição filosófica “menor” (Simondon, Deleuze e Guattari) – e Donna Haraway (2023a, “Um manifesto ciborgue” [1985]) – no centro das discussões sobre ciência e feminismo – são duas figuras importantes que propuseram o afrouxamento de divisões como natureza/cultura, técnico/social, natural/artificial, sujeito/objeto. Assim, liberadas de suas posições estanque, humanidade e não humanidade passam a ser resultados e não premissas de uma pesquisa, pois “toda interação humana é sociotécnica” (Latour 2017:253) e a humanidade deve ser encontrada, não nos polos de divisões binárias, mas justamente “no próprio ato de delegação, no passe, no envio, na troca contínua das formas” (Latour 2019:174). Partindo da suspensão desses divisores, torna-se possível a proposição de uma análise empírica da agência de um algoritmo sem que esta possibilidade seja preliminarmente sufocada, antes mesmo de uma pesquisa, pela constatação de que um objeto não poderia agir. Esperamos que esta perspectiva – em diálogo com diversos outros trabalhos referenciados ao longo da dissertação – seja tão convincente para a leitora quanto julgamos eficaz para nossa análise.

(2)

Nosso encontro com o trabalho *Modulações algorítmicas* de Débora Machado (2019), até onde sabemos pioneiro na análise da ação social de algoritmos a partir do estudo de patentes. Neste, Machado analisa a “modulação de comportamento” de usuárias realizada por técnicas patenteadas pela Facebook/Meta, propondo categorias tipológicas de modulação em algoritmos. A coincidência no foco na participação de algoritmos na ação, especificamente utilizando o conceito deleuziano de modulação para caracterizá-la, nos é extremamente valioso; porém é o fato de ser um trabalho de sociologia da tecnologia cujo material empírico foi a descrição de técnicas em documentos de patentes que se mostrou para nós o aspecto mais importante do trabalho e que motivou o desenvolvimento do nosso. Ao longo da dissertação, apresentamos outras autoras que utilizaram patentes em suas análises sociológicas, mas nenhum outro trabalho é baseado em um levantamento sistemático de um conjunto de patentes da mesma empresa para descrever o que há de comum na ação dessas técnicas. Desse modo, esse foi

um trabalho importante a partir do qual pudemos nos inspirar e nos amparar metodologicamente na construção desta pesquisa.

Diante disso, conectado a questões amplas como a da agência não humana e do papel das tecnologias digitais no capitalismo contemporâneo, nosso trabalho visa contribuir para uma acepção socio-técnica do AdM, caracterizando seu funcionamento ou sua agência, bem como contribuindo no debate sobre os processos sociais implicadas com a construção e uso dessa técnica. Isso é feito por meio de uma pesquisa empírica sobre patentes da Google em conexão com uma literatura heterogênea que busca relacionar tecnologia e sociedade. A apresentação mais detida do material e da metodologia empregada se encontra no capítulo 2.

0.3 Argumento da dissertação

Antes de iniciar o nosso trabalho propriamente dito, gostaríamos de explicitar rapidamente o argumento que esta dissertação desenvolve. Iniciamos com uma hipótese – entendida aqui menos como uma suposição inicial anterior a qualquer pesquisa e que deve ser testada, e mais como um apontamento teórico-metodológico que é já cientificamente informado e que nos serve de “prumo” – de que o AdM, com seus processos envolvendo a maximização de acertos, a classificação, predição e geração de padrões, etc., ocuparia uma centralidade na vida social contemporânea: tais técnicas influenciariam, modificariam ou substituiriam a atividade de usuárias que interagem com sistemas de computação digital. É o processo que chamamos, influenciados principalmente por Deleuze (2013), de modulação.

A nossa análise sobre o material empírico da Google, empresa sobre a qual conscientemente escolhemos nos debruçar, mostrou com complexidade sinergias e desvios em relação ao nosso entendimento intuitivo inicial. No fim e ao cabo, foram duas as principais conclusões que chegamos:

1. A agência algorítmica é o modo particular como a ação e significação surgem em sistemas de computação digital. Em um computador digital, os fluxos elétricos e eletromagnéticos e os padrões eletrônicos que servem de substrato para a informação digital são sempre controlados por funcionamentos algorítmicos. *Os algoritmos são os mediadores da informação digital* e sua ação, a mediação algorítmica, determina tudo o que ocorre em um sistema de computação digital, entre a informação de entrada (advinda de sensores e outros aparatos conectados a computadores) e a informação de saída (de uma tela ou outros atuadores igualmente a eles conectados);
2. A modulação de atividade é apenas um dos aspectos da aprendizagem de máquina. A descrição de sua operação exige a consideração de outros dois processos. Em primeiro lugar, a captura da informação por meio de uma infraestrutura ativa que visa captar sinais de informação analógicos e digitais, da usuária e do meio, transformando-as em dados e, posteriormente, em parâmetros numéricos. Em segundo lugar, o processamento desses dados e o estabelecimento de tais parâmetros criam uma representação (uma imagem) que confere ao AdM performatividade, conectando a captura e a modulação. Desse modo *o aprendizado de máquina pode ser caracterizado por três processos inter-relacionados e interdependentes que são: a captura, a imaginação e a modulação.*

O desenvolvimento de ambas consta nas páginas seguintes. Desejamos, sem mais delongas, uma boa leitura!

Parte I

Introdução: tecnociência e capitalismo contemporâneo

Capítulo 1

A Google como empresa e como tecnologia

Don't be evil.

Do the right thing.

Google

E aí eu perguntei por que os caras tinham colocado esse nome Gooooogle. O que quer dizer Google? Naquele momento eu fiz essa pergunta e eu sei que muita gente nunca se fez essa pergunta até hoje, to certo? Então, Google vem do googol que é o símbolo do infinito que é aquele do oito deitado, a representação numérica do googol é um seguido de cem zeros. Um mais um zero é dez, mais dois zeros é cem, com três zeros é mil, dez mil, cem mil, um milhão... O cem zeros como chama? Isso é a Google. E aí eu falei, gente, se o cara carrega isso no nome, o que esses caras querem? Que tipo de coisa que a gente pode enrolar com esses caras? *É melhor a gente ficar esperto, esses caras não parecem coisa boa não. E aí você vai ver em relação a apropriação de dados o que a Google dá conta de fazer hoje. Vocês imaginam o tamanho do datacenter da Google? Então, essa trajetória...*

TC Silva

O slogan “*Don't be evil* [Não seja malvado]” foi utilizado como lema empresarial da Google Inc. até o ano de 2015, marco da reestruturação que fez da Alphabet Inc. o conglomerado “pai” [*holding*] e a Google LLC sua principal subsidiária. A partir desse ano, a Alphabet passa a utilizar o lema “*Do the right thing* [Faça a coisa certa/direita]”. Entretanto, o slogan inicial continua no código de conduta da empresa até hoje¹. Em uma versão anterior do código, lemos:

Nosso lema corporativo informal é “Não seja malvado”. Nós Googlers geralmente relacionamos essas palavras com a maneira que nós atendemos nossas usuárias, e também como deveríamos atender. Mas ser “um tipo diferente de companhia” significa mais do que os

¹O histórico do lema é traçado na página da Wikipedia, disponível em https://en.wikipedia.org/wiki/Don%27t_be_evil (acesso: 24/06/2024).

produtos que fazemos e o negócio que estamos construindo; significa ter certeza que nossos valores centrais informam nossa conduta em todos os aspectos de nossas vidas como empregadas da Google.²

Entretanto, como aponta Antônio Carlos Santos Silva, mais conhecido como TC, importante nome da cultura digital brasileira, mestre e fundador da Casa de Cultura Tainã e membro da Rede Mocambos, em fala feita no evento “Teia Cultura Digital”, no dia 29 de agosto de 2020, há uma divergência patente entre os lemas utilizados pela empresa e sua atuação concreta, por exemplo no que diz respeito à “apropriação de dados”³. Com uma pequena ressalva, concordamos com Sérgio Amadeu da Silveira ao dizer que “o grande valor do Google não está [somente] em seus milhares de servidores, prédios, veículos ou terrenos, mas na sua capacidade de obter dados de milhões de pessoas no planeta e cruzá-los a fim de formar perfis de consumidores potenciais, organizar as características finas dos comportamentos e agrupar os diversos tipos de preferências culturais, econômicas e até mesmo ideológicas”, e ainda que, a partir da captura de informação constante e o máximo possível em tempo real, “o grande negócio do capitalismo cognitivo [como o capitalismo contemporâneo é por alguns denominado] é a modulação de comportamento” (Assange 2015b:15).

1.1 Alguns aspectos sociais e econômicos

A Google/Alphabet (doravante simplesmente Google, por entender que tal distinção é um artefato jurídico-econômico que não se reflete em uma cisão na prática do objeto de nossa pesquisa) é a empresa detentora do Google Search – site mais acessado do mundo –, do Youtube – maior plataforma de compartilhamento de vídeos, e o segundo site mais acessado – e principal empresa por trás do sistema operacional Android – o mais usado por smartphones e por dispositivos em geral⁴. Desse modo, os fluxos informacionais controlados por ela são não só gigantes, como também, mundiais. Além disso, a Google é uma importante agente não só na aquisição de startups de tecnologia (e.g. Rodolfo Avelino em Cassino, Souza, e Silveira 2021:70–71), mas também na produção de artigos científicos, notadamente da área de IA⁵.

Nesse sentido, fazemos coro com Zuboff (2018) que constata que a Google, detentora “[d]o serviço de busca mais popular do mundo[, ... e] considerada por muitos como a pioneira do *big data*” é empresa central da convergência contemporânea entre tecnociência e capitalismo, argumentando que ela “foi a pioneira na lógica de acumulação mais ampla que [a autora denomina] de capitalismo de vigilância, da qual o *big data* é tanto uma condição quanto uma expressão” (Zuboff 2018:24–25).

²No original: “Our informal corporate motto is ‘Don’t be evil.’ We Googlers generally relate those words to the way we serve our users – as well we should. But being ‘a different kind of company’ means more than the products we make and the business we’re building; it means making sure that our core values inform our conduct in all aspects of our lives as Google employees” (Google 2005).

³A transcrição completa de sua fala pode ser encontrada em Paciornik (2021).

⁴Baseados respectivamente nas listas compiladas em https://en.wikipedia.org/wiki/List_of_most-visited_websites e https://en.wikipedia.org/wiki/Usage_share_of_operating_systems (acesso: 24/06/2024).

⁵Acreditamos que a pesquisa informal realizada por Gleb Chuvpilo em 2019 <https://chuvpilo.medium.com/ai-research-rankings-2019-insights-from-neurips-and-icml-leading-ai-conferences-ee6953152c1a> e em 2020 <https://chuvpilo.medium.com/whos-ahead-in-ai-research-in-2020-2009da5cd799> (acesso: 24/06/2024), embora não sejam suficientes definir a Google como principal instituição de pesquisa em IA hoje, ao menos dão uma dimensão da sua posição.

1.1.1 O império das mentes na nova era digital: Assange encontra Jared Cohen e Eric Schmidt

O livro *Quando o Google encontrou o WikiLeaks* é a resposta de Julian Assange (2015b) em relação à entrevista concedida à Eric Schmidt, presidente executivo da Google, e Jared Cohen, diretor do Google Ideas, em 2011, quando Assange se encontrava em prisão domiciliar em Norfolk, Reino Unido⁶. Além de ambos, que realizavam a entrevista no intuito de escrever um livro sobre o futuro das tecnologias digitais, também se encontravam ali o editor do livro, Scott Malcomson que já trabalhara como consultor nas Nações Unidas e que depois entraria para o Departamento de Estado, e Lisa Shields, namorada de Schmidt da época e vice-presidente do Conselho de Relações Exteriores⁷.

O livro em questão, inicialmente concebido com o título *The Empire of the Mind [O império da mente]*, mas publicado com outro: *The New Digital Age: Reshaping the Future of People, Nations and Business*⁸ (2013), segundo a avaliação de Assange (2015b:42), “não cumpriu a promessa [de ser uma antecipação visionária das mudanças tecnológicas globais] e deixou até de imaginar um futuro, bom ou ruim, consideravelmente diferente do presente”. O argumento do livro, ainda segundo Assange (2015b:42) é “uma junção simplista da ideologia do ‘fim da história’ de Fukuyama [...] e smartphones mais velozes”. Para Assange (2015b:42), o livro não deixa de ser uma “serenata de amor do Google ao governo norte-americano”, em que “[o] Google, um super-estado digital, se oferecia como o visionário geopolítico de Washington”. “*A nova era digital* é, antes de mais nada, uma tentativa do Google de se posicionar como um visionário geopolítico dos Estados Unidos, a única empresa capaz de responder à pergunta ‘Para onde a América deveria ir?’” em que “os senhores Schmidt e Cohen assumem alegremente o fardo do ‘nerd branco’”. Por outro lado, o livro “é uma obra funestamente seminal, e nenhum dos autores parece ter a capacidade de enxergar, e muito menos expressar, a titânica perversidade centralizadora que estão construindo”. Para Assange, “[o] avanço da tecnologia da informação simbolizada pelo Google prenuncia a morte da privacidade para a maioria das pessoas e aproxima o mundo do autoritarismo” (Assange 2015b:55, 56, 58, 57), de modo que: “Sem nem ao menos entender como, eles atualizaram e discretamente implantaram a profecia de George Orwell. Se você quiser uma visão do futuro, basta imaginar o Google Glasses – que é financiado por Washington – em rostos humanos inexpressivos... para sempre” (Assange 2015b:58).

Entretanto, o debate que a entrevista e os livros de Assange e Cohen e Schmidt levantam são bastante interessantes por lançarem luz sobre aspectos menos evidentes da Google, explicitando a trama

⁶Julian Assange é o jornalista, programador de computador, ativista cypherpunk (cf. Assange 2015a; *apud* Assange 2015b:136, n. 62) e, mais notadamente, fundador e editor chefe do WikiLeaks, “uma organização de mídia multinacional e biblioteca associada” (WikiLeaks 2015). Descrever a trajetória de Assange ou do WikiLeaks fugiria ao escopo desse trabalho, basta apenas lembrar a leitora que o site ganhou atenção após a publicação de documentos sigilosos do governo dos Estados Unidos, incluindo registros de guerras, que foram vazados por Chelsea Manning em 2010. Desde então, as autoridades estadunidenses criminalizam o WikiLeaks e seu fundador, em um conflito transnacional que incluiu anos de exílio em embaixadas e duas prisões. Após um acordo com a justiça americana firmado em 24 de junho deste ano em que ele se declarou culpado “de jornalismo”, em suas palavras, Assange foi solto da Prisão de Belmarsh, no Reino Unido, onde estava detido desde 2019.

⁷Desse modo, Assange (2015b:29) constata que “[a] essa altura, a delegação era um quarto o Google e três quartos o *establishment* da política externa norte-americana”, embora depois Assange perceba que mesmo Schmidt, uma espécie de “‘ministro das relações exteriores’ do Google” tinha fortes ligações com o governo estadunidense, de modo que, naquela conversa “o Departamento de Estado norte-americano conseguiu se infiltrar no centro de comando do WikiLeaks” (Assange 2015b:37, 32).

⁸O qual, na tradução brasileira, se torna ainda mais brando: *A nova era digital: como será o futuro das pessoas, das nações e dos negócios* (2013).

política em que a empresa se imbrica com diversos departamentos e instituições políticas dos EUA com interesses geopolíticos globais. Assange (2015b) apresenta como, desde antes de sua formalização como empresa, as pesquisas que foram base para a criação do buscador por Lawrance “Larry” Page e Sergei Brin, àquela altura estudantes em Stanford, foram parcialmente financiadas pela Agência de Projetos de Pesquisa Avançada de Defesa (DARPA, *Defense Advanced Research Projects Agency*) e pela Administração Nacional da Aeronáutica e Espaço (NASA, *National Aeronautics and Space Administration*) (Assange 2015b:38, 52, n. 69). Depois, ela também recebeu financiamento da Agência de Segurança Nacional (NSA, *National Security Agency*) e participou do PRISM, programa de vigilância da agência, tal qual denunciado por Edward Snowden, ex-funcionário da NSA. Ela também comprou, em 2004, a Keyhole, start-up de mapeamento cofinanciada pela Agência Nacional de Informação Geoespacial (NGA, *National Geospatial-Intelligence Agency*) e pela Agência Central de Inteligência (CIA, *Central Intelligence Agency*), a partir de quando desenvolve o Google Maps e uma versão empresarial fornecida ao Pentágono e outros órgãos estadunidense (Assange 2015b:38 *et seq.*, 50-51, n. 68). Ainda outros casos de aproximação da Google com os EUA são narrados por Assange, mas acreditamos que essa pequena lista é suficiente para o que nos interessa.

Retornando ao lema da empresa utilizado por nós como primeira epígrafe, Assange (2015b:37) constata que “ao que tudo indica”, seus líderes realmente acreditam no “poder civilizador das multinacionais esclarecidas” que formata o mundo segundo o julgamento de uma “superpotência benevolente”: “Eles dizem que uma mente aberta é uma virtude, mas os pontos de vista que ameaçam a orientação excepcionalista que reside no centro da política exterior norte-americana continuarão invisíveis para eles. Essa é a banalidade impenetrável do lema: ‘Não seja mal’. Eles realmente acreditam que estão fazendo o bem. E isso é um problema” (Assange 2015b:37).

Em todo caso, não saberíamos dizer se a substituição do “não seja mal” pelo “faça a coisa certa/direita” da Alphabet significaria uma mudança conscienciosa que explicita um posicionamento político específico. Mas, independente disso, acreditamos que sua posição privilegiada de influência sobre as pessoas “dentro” e “fora” da internet é, cada vez mais, inquestionável:

Não importa se o Google é apenas uma empresa ou “mais do que apenas uma empresa”, suas aspirações geopolíticas estão firmemente enredadas na agenda de política externa da maior superpotência do mundo. À medida que cresce o monopólio do Google na área de busca e serviços de internet, ele estende a vigilância industrial para a maior parte da população do planeta, dominando rapidamente o mercado de telefonia móvel e apressando-se para ampliar o acesso à internet no hemisfério sul, ele se *torna* praticamente a própria internet para muitas pessoas. **A influência do Google sobre as escolhas e o comportamento de todos os seres humanos se traduz em um poder concreto de influenciar o rumo da história** (Assange 2015b:41, grifo nosso).

Acreditamos que o relato de Assange nos é importante para eliminar sumariamente uma perspectiva acrítica que entenda a Google como “mera produtora de tecnologia”. Pois, deslocando a política para o centro do debate, fica evidente que não há como sua tecnologia ser neutra, estando diretamente a serviço e constituindo a própria política levada a cabo pela empresa.

1.1.2 Gerando excedente econômico e político: Zuboff (e Fuchs) encontra Hal Varian

Já Zuboff (2018), analisando dois documentos escritos por Hal Varian, economista-chefe da Google, discorre sobre a opinião da empresa sobre alguns dos desdobramentos esperados para a disseminação de transações econômicas mediadas por computador, tendência que nos se apresenta. Os “novos usos” constatados pelo economista são: 1) a extração e análise de dados dessas transações, 2) novas formas contratuais baseadas em um monitoramento melhorado, 3) personalização e customização para as pessoas envolvidas, e 4) experimentos contínuos. Para Zuboff (2018:26), isso seria índice de “uma lógica emergente de acumulação”, dando também pistas sobre “a divisão de aprendizagem que ela forma e o caráter da civilização da informação para a qual ela conduz”. Para que as quatro possibilidades levantadas por Varian se concretizem, é necessário que os fluxos informacionais passem a ser não só mediados por computador, mas também capturados pelas empresas. Desse modo, a ambição corporativa passa a ser a de controlar um fluxo informacional cada vez maior: “dados de bilhões de sensores incorporados em uma ampla gama de objetos, corpos e lugares” (Zuboff 2018:27). Em relação à empresa que nos interessa:

Os novos investimentos da Google em machine learning [aprendizado de máquina], drones, dispositivos vestíveis, carros automatizados, nanopartículas que patrulham o corpo procurando por sinais de doenças e dispositivos inteligentes para o monitoramento do lar são componentes essenciais dessa cada vez maior rede de sensores inteligentes e dispositivos conectados à internet destinados a formar uma nova infraestrutura inteligente para corpos e objetos. (Zuboff 2018:27, grifo nosso)

Dessa perspectiva, a Google – e outras grandes empresas de tecnologia⁹ – faz tudo o que está sob seu alcance para aumentar seu controle sobre a informação. Nesse sentido, o exemplo da extração ilegal de dados pela invasão de redes wi-fi por carros do Google Street View mostra essa lógica que pode ser chamada, seguindo Siva Vaidhyanathan, citado por Zuboff, de um “imperialismo de infraestrutura”:

Mas isso só é possível porque essas empresas se constituíram, hoje, como pontos de passagem obrigatória, no sentido que dá a eles Latour (2011:187–88, *passim*). Não apenas a empresa é dona dos dois sites mais acessados do mundo, mas também seus produtos, cada vez mais integrados em uma lógica de plataforma, são utilizados como se não houvesse concorrência. Foge do escopo deste trabalho entender por que, apesar da existência de outros buscadores, serviços de nuvem, serviços de e-mail, serviços de compartilhamento de vídeos, serviços de videoconferência e serviços de mapas¹⁰, por exemplo, a Google ainda se estabelece como se não houvesse alternativas. Entretanto, talvez a resposta esteja justamente

⁹Remetemos a leitora principalmente à conhecida sigla Google/Alphabet, Amazon, Facebook/Meta, Apple e Microsoft (GAFAM), também chamadas de “big five” estadunidense e que mais ou menos corresponde ao também conhecido FAANG: Facebook/Meta, Amazon, Apple, Netflix e Google/Alphabet. Couldry e Mejias (2020) ainda adicionam Baidu, Alibaba, e Tencent como correlatas chinesas. O acrônimo NATU também é utilizado mais recentemente para: Netflix, Airbnb, Tesla e Uber. Independente das siglas e denominações, todas essas companhias (e ainda outras como Spotify, IBM, Nvidia e Xiaomi) podem ser consideradas “big techs”, ou seja, grandes empresas de tecnologia: “A empresa não pergunta se pode fotografar casas para seus bancos de dados, ela simplesmente pega o que quer. A Google, então, esgota seus adversários no tribunal ou eventualmente concorda em pagar multas que representam um investimento negligenciável para um retorno significativo” (Zuboff 2018:30).

¹⁰O buscador duckduckgo (<https://duckduckgo.com/>), nuvens nextcloud (<https://nextcloud.com/>), e-mails riseup (<https://riseup.net>) ou proton (<https://proton.me/>), peertube (<https://joinpeertube.org/>), jitsi (<https://meet.jit.si/>) e open street map

na sua incrível capacidade de mediar a informação digital e capturar dados, gerando um potencial de predição e de ação que é, este sim, inigualável.

Se a Google é hoje a principal detentora de dados sobre a “cotidianidade”, isto é, “small data” (Zuboff 2018), gerado continuamente por meio da mediação algorítmica de atividades não mercantis, é porque foram pioneiras em um modelo de negócio baseado justamente na produção e extração de dados: “Eles então optaram por um modelo de propaganda. A nova abordagem dependia da aquisição de dados de usuários como matéria-prima para análise e produção de algoritmos que poderiam vender e segmentar a publicidade por meio de um modelo de leilão exclusivo, com precisão e sucesso cada vez maiores” (Zuboff 2018:32).

Christian Fuchs, analisando o “capitalismo da Google”, ampara-se na noção de prosumidor (*prosumer*, ou produtor-consumidor) dos estudos de jogos (*game studies*)¹¹, para lançar luz sobre como a Google capitaliza a partir de serviços gratuitamente disponibilizados (figura 1.1). Para o autor, são duas as formas como a empresa acresce valor nesses serviços:

Por um lado, ela indexa conteúdo gerado por usuárias que é enviado para a rede e, portanto, age como uma meta-exploradora de toda produtora de conteúdo gerado por usuário. [...] Portanto a Google explora todas as usuárias que criam conteúdos na Rede Mundial de Computadores [World Wide Web] (WWW). Por outro lado, usuárias utilizam serviços da Google e, com isso, conduzem um trabalho não-pago de geração de mais-valor (Fuchs 2012:43).

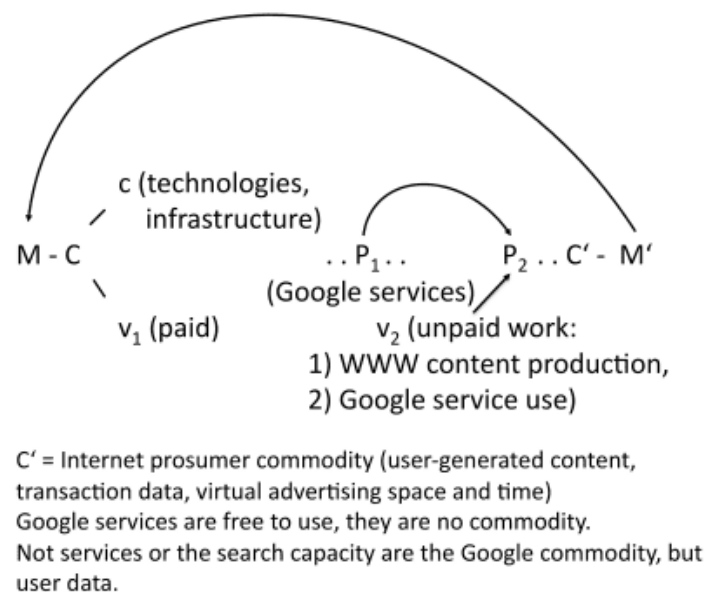


Figura 1.1: Ciclo de acumulação de capital da Google: O dinheiro (M) investido em tecnologia (c) e força de trabalho pago (v1), isto é, no salário de trabalhadoras da Google, produz serviços (P1) que serão utilizados gratuitamente por usuárias em processos de trabalho não-pago (v2), gerando dados e conteúdos (P2). Esta nova mercadoria de prosumidor (C') é então vendida para anunciantes, gerando lucro (M' maior que M). Fonte: Fuchs (2012).

(<https://www.openstreetmap.org/>) são, respectivamente, alternativas baseadas em software livre para os serviços da Google listados.

¹¹Para uma aproximação interessante dos *game studies* com a sociologia da tecnologia, elaborando sobre as noções de prosumidor e jogador (jogador-trabalhador), conferir Garcia dos Santos e Ferreira (2008).

Desse ponto de vista, serviços como o Google Search e Youtube não chegam a ser mercadorias. Os dados de usuárias, por outro lado, vendidos para anunciantes, sim. O autor argumenta ainda que esta situação pode ser interpretada como uma “terceirização [*outsourcing*] do trabalho produtivo para usuárias, que trabalham completamente de graça e ajudam a maximizar as taxas de exploração” (Fuchs 2012:45). Em consonância com Zuboff, Fuchs (2012) constata que essa situação só é possível por meio da utilização de práticas intensas de vigilância. Além disso, argumenta que “[o] problema não são as tecnologias providas pela Google, mas as relações capitalistas de produção na qual estas são organizadas. O problema é que a Google, para prover seus serviços, necessariamente precisa explorar usuárias e engajar na vigilância e mercantilização [*commodification*] de dados de usuárias [*user-oriented data*]” (Fuchs 2012:46).

O paradigma analítico de Fuchs é esclarecedor e fornece elementos importantes para entender a prática, hoje generalizada, de empresas que vivem do fornecimento de serviços gratuitamente disponibilizados. Entretanto, ao separar a vigilância e a exploração dos serviços materialmente providos pela Google, acreditamos que Fuchs recai em duas limitações que Zuboff nos ajuda a resolver. Em primeiro lugar, a inseparabilidade da agência técnica e da racionalidade embutida em sua materialização: serviços como a ordenação por relevância de sites, a sugestão de palavras na escrita de e-mails e documentos e a otimização de rotas não são possíveis sem uma infraestrutura ativa de produção e extração de dados. Em segundo lugar, que, para além da venda de dados para anunciantes, o controle do meio sobre o qual as trocas informacionais acontecem permite um modo de dominação sobre a atividade de usuárias que ultrapassa a noção de exploração econômica.

Para Zuboff (2018), além da exploração tão bem analisada por Fuchs, o que uma infraestrutura ativa de captura informacional (o que a autora denomina “Big Other”) permite é a antecipação e a influência da atividade das usuárias. Para a autora, em uma situação limite:

A conformidade agora desaparece na ordem mecânica de coisas e de corpos, não como ação, mas como resultado, não como causa, mas como efeito. Cada um de nós pode seguir um caminho distinto, mas esse caminho já é moldado pelos interesses financeiros e/ou ideológicos, que imbuem o Big Other e invadem todos os aspectos da “vida privada” de cada um (Zuboff 2018:45).

Essa capacidade gerada pelo processamento de dados, que a autora denomina “excedente comportamental [*behavioral surplus*]”, mas que, com Parra (2022), preferimos chamar de excedente informacional, é o que orienta as práticas de vigilância e captura e o que possibilita um deslocamento nas finalidades da empresa “em direção à observação, à comunicação, à análise, à previsão e à modificação em tempo real do comportamento atual e futuro” (Zuboff 2018:55). Em outras palavras, em direção à modulação, cujo objetivo é o “lucro” e o “controle”.

Sua influência geopolítica, o escopo e a escala de seus serviços, a ubiquidade de seus produtos, sua invencibilidade legal e superioridade mercadológica nos permite ver a Google como um “império” transnacional (Assange 2015b), uma “exploradora global” (Fuchs 2012), bem como a principal agente no “organismo mundial” (Zuboff 2018) que controla os fluxos informacionais planetários. Desse modo, acreditamos ainda que não seria exagero dizer que ela seja hoje a instituição que mais se aproxima de um *governo global*¹².

¹²Agradecemos a Pedro Ferreira por nos oferecer essa intuição, corroborada pelos argumentos expostos em Assange

1.2 Do PageRank™ ao Google Search: tecnologia e modulação de atividade

O PageRank, um dos principais algoritmos que compõem o buscador Google Search, é um ótimo ponto de partida para introduzir as relações entre tecnociência e capitalismo que a Google congrega. Assim, tomando o buscador como caso concreto de partida, buscamos mostrar brevemente como essas dimensões exploradas acima se relacionam com algoritmos, patentes e com a própria modulação. Esperamos que a análise mais panorâmica delineada aqui sobre um serviço bastante conhecido ajude a justificar a relevância da análise mais detida e específica que fazemos sobre as patentes de AdM na parte II.

Inicialmente desenvolvido por Larry Page e Sergei Brin na Universidade de Stanford em 1998, o PageRank foi primeiro descrito em um artigo científico (Brin e Page 1998), depois patentado (Page 2001) e sua marca foi registrada¹³ (2004). A intuição por trás desse mecanismo é considerar a relevância de uma certa página da internet baseado na quantidade e na qualidade de outras páginas interligadas a ela. Relevância neste caso, é uma função matemática relacionada com a quantidade e com a origem de links que apontam para determinada página. No mesmo ano de sua concepção, a Google Inc. foi fundada por eles. Interessante notar portanto como, já em sua origem, a Google se constituía de um nó emaranhado de empresa, pessoas, algoritmos e patentes – o que sustenta a nossa escolha de levar em conta, a um só tempo, essas diversas dimensões em sua análise.

Cecilia Diaz-Isenrath (2005) faz um estudo deste tipo, analisando o artigo e a patente que deram origem à empresa. Sua análise mostra como, desde sua concepção, a Google é um entrelaçamento de servidores, empresas doadoras, instituições governamentais, usuárias, sistemas operacionais, patentes, programas, etc., “em 1998 [a Google] era um híbrido: uma mistura de algoritmos, linguagens formalizadas e projeto de pesquisa, uma trama de mecanismos de cálculo, codificação e gestão” (Diaz-Isenrath 2005:103)“. Além da NASA e da DARPA, já citadas, os agradecimentos do artigo citam a IBM, a Intel e a Sun, importantes empresas de tecnologia por doarem equipamentos à Stanford, e o Projeto de Biblioteca Digital dessa universidade. Mesmo assim, em um primeiro momento, ela não é muito mais que “um centro de pesquisa na Universidade de Stanford”¹⁴, enquanto que “em 2004, [ela é] uma companhia com quase 2000 funcionários”, “tem escritórios de suporte técnico, vendas e centros de P&D em diferentes locais dos EUA e em outras cidades”, “[p]ossui 95 nomes de domínios de internet registrados e interfaces com o usuário [sic] disponíveis em 97 idiomas”¹⁵ (Diaz-Isenrath 2005:108).

A diferença da Google em relação aos mecanismos de busca correntes daquela época é que seu produto “busca não só rastrear e indexar uma vasta quantidade de documentos”, tarefa já bem executada por outros serviços, “mas também retornar resultados ‘relevantes’” (Diaz-Isenrath 2005:101). Enquanto que a primeira tarefa é realizada pelo “Googlebot”, o algoritmo “PageRank” se encarrega da segunda. Diaz-Isenrath (2005) mostra também como a intenção inicial era que tal funcionalidade beneficiaria as consumidoras, diferentemente de outros produtos preocupados unicamente com o anúncio de propagandas, de modo que, seguindo o “não seja malvado”, a Google apresentaria apenas “anúncios discretos

(2015b), Fuchs (2012), e Zuboff (2018).

¹³Disponível em https://tsdr.uspto.gov/#caseNumber=78242279&caseSearchType=US_APPLICATION&caseType=DEFAULT&searchType=statusSearch (acesso: 25/06/2024).

¹⁴“Fora desse centro, ‘Google’ era uma maneira de pronunviar a palavra ‘googol’, usada por matemáticos para designar 10 elevado à centésima potência” (Diaz-Isenrath 2005:102).

¹⁵A autora ainda acrescenta que “todas as palavras deste mundo parecem não lhe bastar: o buscador apresenta [em 2004] interfaces em duas línguas ‘extraterrestres’ (Klingon e Leet), tomadas de uma série de ficção científica” (Diaz-Isenrath 2005:108).

e bem diferenciados da pesquisa” (Diaz-Isenrath 2005:110). Se bem que a própria imagem que a Google apresenta muda desde sua concepção: “No paper de 1998, Page e Brin justificavam o uso da palavra ‘google’ (10 elevado à centésima potência) porque combinava com o objetivo do projeto de construir uma máquina de busca de larga escala. Em 2004, essa palavra ‘reflete a missão da companhia de organizar o imenso, aparentemente ilimitado volume de informação disponível na web’”¹⁶ (Diaz-Isenrath 2005:109).

Tarleton Gillespie (2018), analisando os algoritmos de relevância pública, dos quais o PageRank seria um exemplo, insiste na sua natureza eminentemente política: tais algoritmos executam critérios de ordenamento específicos e institucionalmente informados que influenciam a usuária a eles conectada, ao mesmo tempo que se apresentam como neutros e evitam qualquer responsabilização por “erros” e “vieses”¹⁷. Desse modo, “devemos ver os algoritmos não apenas como códigos com consequências, mas sim como o mais recente mecanismo construído socialmente e institucionalmente gerenciado para convencer o julgamento público: uma nova lógica de conhecimento” (Gillespie 2018:117). O autor apresenta como há uma agência assimetricamente distribuída entre quem desenha o algoritmo e suas usuárias, embora estas ainda participem em alguma medida nos sentidos gerados por ele. Entretanto, o programa de pesquisa elaborado por Gillespie, embora importante e lastreado em uma extensa literatura de estudos sobre tecnologias e algoritmos e em muitos exemplos concretos, ainda pretende “resistir firmemente à tentação de tratar a tecnologia como condutora nas explicações” e “desvendar as escolhas humanas e institucionais que estão por trás desses mecanismos frios” (Gillespie 2018:98).

Por outro lado, Zeynep Tufekci (2015), de um modo que entendemos que pode ser complementar a Gillespie, considera os algoritmos como agentes computacionais e também alerta para o caráter invisível dos processos computacionais, sobretudo em sua intersecção com os “big data”, mostrando como eles atuam como “porteiros [gatekeepers]”, mediando a informação digital, e levantando novos problemas relativos à transparência, assimetria e privacidade, mas sem, com isso, reduzi-los completamente à intencionalidade de humanos “por detrás”. Isso é especialmente verdade no caso de algoritmos de ordenamento de buscas em um site indexador como o Google Search, pois é a partir deste que, muitas vezes, a usuária se direciona para procurar ou encontrar um conteúdo na internet, sem que seja possível que cada busca tenha sido previamente antecipada e seu sentido minuciosamente direcionado por programadoras da Google.

Matteo Pasquinelli (2009) analisa especificamente o algoritmo PageRank, argumentando que ele pode ser visto como um “diagrama” do próprio “capitalismo cognitivo”:

O diagrama do PageRank ilustrado acima [no texto original, abaixo aqui, na figura 1.2,]

¹⁶Que continua, evidenciando a trama que une tecnociência e capitalismo, ainda sobre o nome: “Trata-se de uma marca da empresa, registrada nos EUA e em outros países. Outras marcas e serviços da Google são PageRank, Froogle, AdSense, AdWords, Blogger e Gmail. Também são marcas registradas ‘I’m Feeling Lucky’ (o texto sobre um dos botões de comando na interface do motor de busca) e ‘It’s all about results’ (UNITED STATES OF AMERICA, 2004). A ampla difusão de ‘Google’ como marca tem trazido alguns efeitos não previstos pela empresa: muitos internautas, sobretudo falantes de inglês, usariam ‘to google’ como um verbo, sinônimo de ‘procurar alguma coisa no Google’ ou, em geral, de ‘fazer uma busca na web’. Os representantes do Google, lemos no documento submetido à SEC, preocupam-se com que o nome se torne uma marca genérica e com o conseqüente risco de que a forma verbal possa diluir a imagem da empresa” (Diaz-Isenrath 2005:109). Hoje percebemos, entretanto, que o uso do nome da empresa como verbo não diminuiu em nada seu renome, talvez seja mesmo o contrário.

¹⁷Para o autor, há seis dimensões do valor político dos algoritmos, os quais ele denomina “padrões de inclusão”, “ciclos de antecipação”, “avaliação de relevância”, “promessa de objetividade algorítmica”, “entrelaçamento com a prática” e “produção de públicos calculados” (Gillespie 2018:98).

não tem qualquer semelhança com a estrutura centralizada do Panóptico descrito por Foucault em *Vigiar e Punir*. A natureza líquida e hipertextual da rede (e da noosfera em geral) exige ilustração adicional. Um diagrama do capitalismo cognitivo pode ser intuitivamente traçado se – na estrutura de um hipertexto – cada relação [*link*] simétrica é substituída por um vetor assimétrico de energia, dados, atenção ou valor. O que o PageRank revela e mede é precisamente essa constituição assimétrica de qualquer hipertexto e rede (Pasquinelli 2009:5)

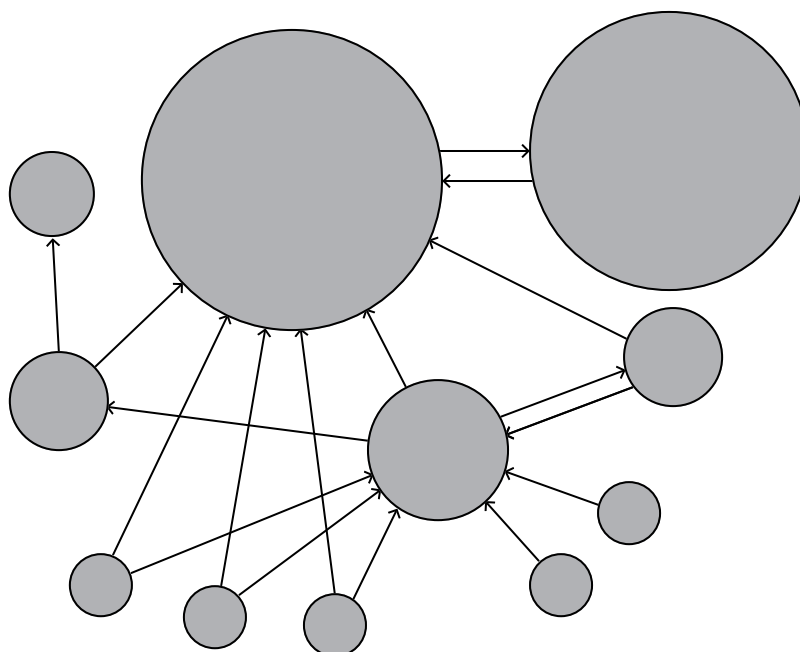


Figura 1.2: Diagrama simplificado do PageRank. A relevância de uma página (tamanho do nó no diagrama) depende da quantidade e da qualidade (função dos links que apontam para cada página) de links que apontam para ela (setas). Fonte: baseado em Pasquinelli (2009).

Sua análise possui alguma similaridade com Fuchs (2012), pois se propõe a explicar a amplificação de valor e geração de lucro a partir de um serviço da Google, mas partindo de premissas bastante distintas sobre a possibilidade de apropriação capitalista de uma capacidade coletiva de conhecimento, o que Marx (2015) chamou de “intelecto geral [*general intellect*]”. Especificamente, sua perspectiva dialoga com autores que pensaram o “capitalismo cognitivo”, entre outros, pela criativa e controversa conceituação de “trabalho imaterial”¹⁸. Desta perspectiva, o PageRank da Google seria um “aparato de captura parasitário do valor produzido pela inteligência comum” (Pasquinelli 2009:6), já que o PageRank é capaz de “dar um valor de ordenamento a cada nó da rede” que é depois “não-oficialmente reconhecido como moeda da economia global da atenção e influencia crucialmente a visibilidade online de indivíduos e companhias” (Pasquinelli 2009:7). A isso o autor dá o nome de “valor” e “mais-valor” de rede.

Em outro artigo, Pasquinelli (2013) avança em seu argumento: partindo de noções de “máquina” e “informação” em Simondon, “informação valorizante” em Alquati e “máquina abstrata” e “mais-valor máquina” em Deleuze e Guattari, dá precisão à ideia de “mais-valor de rede” em um diálogo que visa aproximar Marx da realidade contemporânea, chamada, pelo autor, de “sociedade do metadado” (Pas-

¹⁸Desenvolvido por exemplo em Lazzarato e Negri (2001). Para uma leitura interessante do conceito, conferir Alves da Silveira e Ferreira (2009).

quinelli 2013). O argumento principal é que “operando como interface numérica entre o domínio do conhecimento e o do capital, *o código digital transforma a informação em valor*” (Pasquinelli 2013:17). Como toda máquina, os dispositivos digitais participam de processos de amplificação de valor, com a diferença de que eles não estão mais confinados à fábrica como outrora estiveram seus correlatos termodinâmicos: “[a]s máquinas cibernéticas [...] escapam da fábrica e cada vez mais convertem a cooperação social e a comunicação em forças produtivas” (Pasquinelli 2013:28). É essa ideia, central para a escola do capitalismo cognitivo, de captura e automatização da cooperação social, que vai, depois, fundamentar a proposição mais ousada de uma “teoria da inteligência-de-máquina-trabalho [*labour theory of machine intelligence*]”¹⁹ (Pasquinelli 2023). Em sua “história social da inteligência artificial”, Pasquinelli (2023:120) explica que: “[s]eguindo Braverman e Schaffer, poder-se-ia adicionar que Babbage proveu não só uma teoria da máquina-trabalho [*labour theory of the machine*], mas uma teoria da inteligência-de-máquina-trabalho”, pois suas “máquinas de calcular”, “‘máquinas inteligentes’ de seu tempo”, “foram uma implementação do olho analítico do mestre da fábrica” (Pasquinelli 2023:120). É esta ideia que o autor vai aplicar para mostrar como a IA e as máquinas informacionais contemporâneas, enquanto “infraestrutura computacional”, são “concretização do trabalho em comum” (Pasquinelli 2023:120), ou ainda, “a automação da inteligência comum” (Pasquinelli 2023, “Conclusion: The automation of General Intelligence”).

Mas de volta ao PageRank, é interessante notar como, em confluência com Zuboff (2018) e Fuchs (2012), Pasquinelli (2009) localiza no negócio de propagandas a geração de lucro da empresa:

O AdSense da Google provê uma infraestrutura leve para propagandas que infiltra em cada interstício da rede como um parasita sutil e mono-dimensional, extraindo lucro sem produzir nenhum conteúdo. Dinheiro entra no ciclo através do Adwords e é então distribuído através do AdSense para blogueiros individuais ou companhias de rede. Na economia da internet, ambos o tráfego de um site e a redistribuição de valor são hoje extensivamente governadas pelo PageRank (Pasquinelli 2009:7).

Dessa perspectiva, a Google pode ser vista como uma empresa que extrai parasitariamente uma “renda cognitiva” praticamente “sem custo”:

A Google pode ser descrita como uma rentista global que está explorando novas terras da internet sem precisar de cercamentos estritos ou produzir conteúdo. Nesse cenário, o Google aparece como um puro aluguel sobre a meta dimensão da informação que é acumulada através das redes digitais. A Google não possui a informação da internet, mas sim o diagrama mais rápido para acessar e medir a inteligência coletiva que a produziu (Pasquinelli 2009:10).

Hoje, porém, o PageRank não é o único algoritmo por trás do Google Search e da Google em geral. As recentes técnicas de aprendizagem de máquina, objeto deste trabalho, têm uma posição central não só para o buscador, mas também para outras tarefas envolvendo processamento de linguagem natural, como o serviço de tradução automática do Google Translate²⁰. Um caso particularmente relevante é o

¹⁹A referência óbvia aqui é em relação à teoria do valor-trabalho (*labour theory of value*) em Marx.

²⁰Em relação ao PageRank, a patente “*Node Embedding via Hash-Based Projection of Transformed Personalized PageRank* [Representações de nós via projeção baseada em cifra de pagerank personalizado transformado]” (US20230214425A1) é um

do MdL Codificador Bidirecional de Representações a partir de Transformers (BERT, *Bidirecional Encoder Representations from Transformers*) (Devlin et al. 2019), similar aos modelos Transformer Gerativo Pré-treinado (GPT, *Generative Pre-trained Transformer*) utilizados no ChatGPT/OpenAi. Em linhas gerais, o BERT é um algoritmo capaz de gerar representações numéricas de palavras ou sequências de caracteres, que são posteriormente utilizadas em diversas tarefas de processamento de linguagem natural²¹. Recentemente, a Google anunciou que o BERT seria usado em quase todas as buscas em inglês que fossem inseridas no Google Search (Google 2020). Segundo a empresa, o uso do algoritmo “seria particularmente útil para entender a intenção por trás da pergunta” e a incorporação do BERT teria “ajudado a melhorar as buscas em uma escala massiva, impactando 1 em cada 10 buscas em inglês nos EUA” (Google 2020).

Tendo tudo isso em vista, podemos agora analisar como a Google e o PageRank, escolhido aqui como caso exemplar, se relacionam com o efeito que mais precisamente nos interessa: a modulação. Ochigame e Ye (2021) propõem uma abordagem muito interessante para observar as diferenças de resultados do buscador com base na localização que a busca foi feita. O “atlas da busca [*Search Atlas*]” visa combater a ideia de que a Google seria “um deus onisciente benevolente”, enfatizando a situacionalidade de seus resultados e explicitando a presença de “regiões” e “barreiras” informacionais. Desse modo, seu trabalho pretende contribuir para a explicação de “como esses resultados [do Google Search] moldaram gradualmente suas [da usuária] opiniões, comportamentos, identidades, anseios e esperanças; sua ideologia, seu ciclo de amizade e sua visão de mundo” (Ochigame e Ye 2021:1971), em outras palavras, como o buscador modula as usuárias.

Como mostrado ao longo do artigo, os resultados de uma busca é significativamente diferente a depender da região que se faz a pesquisa, tanto no caso de páginas (Ochigame e Ye 2021, e.g. figuras 3 e 4), quanto no caso de imagens (Ochigame e Ye 2021, e.g. figuras 1,5-7). Em um primeiro caso, as representações textuais e imagéticas para a palavra “deus” parecem coincidir com aspectos culturais de determinado país, ou melhor, com uma ideia de “deus” ligada à religião dominante deste (Ochigame e Ye 2021, *passim*, figuras 1, 3 e 5). Em outro caso, resultados da busca “anexação da Crimeia [*crimean annexation*]” mudam significativamente se a busca partiu da Rússia, da Crimeia ou da Holanda, focando, respectivamente: no pertencimento ou não da península à Federação Russa; na “ocupação” russa na Crimeia; e nas sanções da União Europeia sobre a Rússia (Ochigame e Ye 2021:1975, figura 4). Um outro exemplo interessante da ferramenta criada pelas pesquisadoras é o agrupamento de resultados de pesquisa mostrado nas figuras 8 e 9 do artigo. A nona figura mostra um possível agrupamento de países em “regiões informacionais” para a busca “como combater mudança climática”, evidenciando uma tendência em países da Europa continental em medidas preemptivas, enquanto que, em alguns países insulares, o foco maior são nas ameaças imediatas, por exemplo. Se em países como a Alemanha, o

caso claro de uso de aprendizado de máquina para melhoria do algoritmo de ordenamento. Embora seja importante enfatizar que nem sempre é fácil relacionar uma patente com um produto e nem mesmo ter certeza de que a invenção registrada foi de fato materializada, há ainda outros exemplos recentes de patentes da Google sobre aprendizado de máquina que provavelmente se relacionam com o PageRank e o Google Search. São elas: “*Dynamic injection of related content in search results* [Injeção dinâmica de conteúdo relacionado em resultados de pesquisa]” (US20210382905A1), “*Training and/or utilizing a model for predicting measures reflecting both quality and popularity of content* [Treinando e/ou utilizando um modelo para prever medidas que refletem qualidade e popularidade de conteúdos]” (US20220004918A1) e “*Systems and Methods of Anomaly Detection* [Sistemas e métodos para detecção de anomalia]” (US20230138371A1).

²¹Para uma brevíssima apresentação do algoritmo com foco na sua propensão a reproduzir estereótipos de gênero, conferir nosso trabalho (Gonçalves e Ferreira 2022a).

foco recai em políticas governamentais e institucionais, na Holanda, Aruba e Suriname, as principais palavras (“uso”, “consumo”) remetem a escolhas individuais (Ochigame e Ye 2021:1980). Além disso, a maioria dos países insulares como Tonga e Suriname mostram palavras ligadas a práticas de gestão domiciliar (“eletrônicos”, “luz”, “desligar [*unplug*]”), enquanto que em alguns países de ilhas mais afetadas por mudanças climáticas o foco recai sobre escolhas alimentares (“diariamente”, “carne”, “evitar”, “vacas”, “fazenda”, “comer”). Este resultado, segundo Ochigame e Ye (2021), é coincidente com estudos etnográficos das ilhas Marshall, que mostram, a despeito de sua “pegada ecológica” pequena, um foco discursivo na culpabilização ao invés de afronta e protesto²².

Embora seja impossível afirmar exatamente os critérios através dos quais essas diferenças nos resultados são criadas – pois além de não termos acesso ao código-fonte do buscador²³, “o Google Search se apoia fortemente em novos sistemas de aprendizado profundo cujas decisões são notoriamente difíceis de se interpretar, até mesmo por pesquisadores da própria Google” (Ochigame e Ye 2021:1981) –, é possível supor que elas são consequência de estratégias de incentivar o uso do buscador, o que gera dados para a empresa e, conseqüentemente, um excedente político e econômico. As autoras insistem que “buscadores não são condutores totalmente neutros que respondem aos interesses de usuárias com resultados objetivamente ‘relevantes’”, mas também não são “reduzíveis a editoras puramente subjetivas com poder ilimitado para selecionar e escolher” (Ochigame e Ye 2021:1981). Para o funcionamento do buscador são utilizados tanto métodos algorítmicos quanto pessoas, como é o caso de trabalhadoras terceirizadas contratadas para avaliar “a qualidade aparente dos resultados” ou moderar “conteúdo inapropriado ou ilícito”, sendo, portanto, “produto tanto de padrões culturais quanto de escolhas de design dos algoritmos” (Ochigame e Ye 2021:1981). Além disso:

Mecanismos de busca respondem tanto aos interesses imediatos de usuárias quanto a imperativos financeiros de corporações. O design do mecanismo de busca da Google é inseparável das prioridades de seu negócio de propagandas. Buscadores também respondem a pressões políticas e legais. O mecanismo chinês [*China-based*] Baidu favorece resultados que se alinham com as visões de autoridades do governo Chinês, e a Google remove resultados para cumprir com as leis europeias de proteção de dados (Ochigame e Ye 2021:1981).

Desse modo, a mediação de uma busca por algoritmos internos ao Google Search compõe uma série de agências que condicionam os resultados da busca. Os interesses são múltiplos: de órgãos reguladores, de algoritmos de ranqueamento, de trabalhadoras investidas no produto, da usuária que busca e, também, da empresa que visa o lucro. Assim, o Google Search modula a atividade da usuária, na medida em que seus resultados a informam e condicionam ações a serem executadas por ela baseadas em tais resultados. Nosso intuito aqui foi mostrar a modulação de atividade no Google Search, em que tal operação é mais evidente – por se tratar de um produto que utilizamos cotidianamente –, para dar mais realismo ao fenômeno. Ao longo deste trabalho, exploraremos alguns outros modos que a mediação algorítmica participa da determinação de ações da usuária, especialmente na parte II.

²²Artigo citado: Peter Rudiak-Gould. 2014. Climate Change and Accusation: Global Warming and Local Blame in a Small Island State. *Current Anthropology* 55, 4 (August 2014), 365–386. DOI: <https://doi.org/10.1086/676969>.

²³Como exemplo dessa limitação, as autoras comentam que “a Google intencionalmente não provê uma API para resultados de buscas na internet e implementa várias táticas para bloquear raspadores [de dados, *scrapers*]. Ironicamente, a Google é uma raspadora ela mesma e lucra massivamente da raspagem [*scraping*] de sites” (Ochigame e Ye 2021:1981).

Esta seção visou explicitar o caráter heterogêneo da Google e de seus produtos, ao mesmo tempo desenvolvimentos tecnocientíficos e meios de obtenção de poder e lucro, envolvendo questões corporativas, especificidades técnicas e leis de propriedade intelectual. A Google se constitui, nesse sentido, como um agente central do capitalismo contemporâneo, sendo simultaneamente um nó de relações heterogêneas, instituição de extensão e influência global e, por isso, um centro de cálculo, no sentido de Latour (2011), em diversas atividades tecnocientíficas. Agora, podemos avançar para outras questões, tais quais: o que mais precisamente é um algoritmo? Qual o papel da inteligência artificial e do aprendizado de máquina nesses produtos? O que significa fazer uma sociologia a partir de patentes de métodos computacionais? O capítulo seguinte apresenta algumas definições iniciais, bem como os métodos e materiais que utilizamos na pesquisa.

Capítulo 2

Propriedade intelectual, sociologia de patentes e os algoritmos de modulação de atividade

Como entender esse processo de colonização do virtual e de capitalização dessa mesma dimensão? ... E assim como o valor de um homem foi reduzido pelo capitalismo ao valor do trabalho abstrato transferido para a mercadoria, agora o valor da informação passa pela mesma redução, através dos diferentes sistemas de propriedade intelectual. Abrem-se assim um horizonte e um campo de atuação insuspeitados para a apropriação capitalista: o plano molecular do finito ilimitado no qual, lembrando Deleuze, um número finito de componentes produz uma diversidade praticamente ilimitada de combinações. ... No plano molecular, a relação entre a tecnociência e o capital se institucionaliza através dos laços jurídicos que unem a invenção e a propriedade intelectual. Mas para que isso ocorresse, foi preciso transferir para este terreno [o da informação] o regime de patentes, que vigorava na esfera industrial e selava as relações entre o direito e a ciência, protegendo a propriedade de artefatos e máquinas, isto é, das coisas que não existiam na natureza inanimada.

Politizar as novas tecnologias

Laymert Garcia dos Santos

A U.S. patent gives you, the inventor, the right to “exclude others from making, using, offering for sale, or selling” an invention or “importing” it into the U.S. A plant patent gives you additional rights on the “parts” of plants (e.g., a plant patent on an apple variety would include rights on the apples from the plant variety). What is granted is not the right to make, use, offer for sale, sell or import the invention, but the right to stop others from doing so. If someone infringes on your patent, you may initiate legal action. U.S. patents are effective only within the U.S. and its territories and possessions.

What is a patent?

USPTO

2.1 Patentes, informação, software e algoritmos

Uma patente é um documento legal que originalmente visa simultaneamente apresentar (tornar patente) uma invenção tecnológica e garantir a exclusividade da possibilidade de construção da mesma pela inventora (França 1997). Nesse sentido, como descreve a epígrafe¹ do Escritório de Patentes e Marcas Registradas dos Estados Unidos (USPTO, *United States Patents and Trademark Office*), órgão oficial do Departamento de Comércio dos EUA, cuja responsabilidade é “(...) conceder patentes para a proteção de invenções e registrar marcas [*trademarks*]” (USPTO sem data-a), ela se institui principalmente como um direito negativo. Em uma versão anterior do mesmo site, lia-se que “o direito que a patente provê é, na linguagem do estatuto e da garantia mesma, ‘o direito de excluir outras de fazer, usar, oferecer para venda ou vender’ a invenção nos Estados Unidos ou ‘importar’ a invenção para os Estados Unidos” (USPTO sem data-a). Ou seja, uma patente é sobretudo o direito de impedir que outros construam a invenção projetada. Um instrumento legal que, em comunhão com a assimetria econômica e política de grandes empresas em relação a indivíduos, é utilizada na competição neoliberal em defesa da concentração econômica. Nesse sentido, as patentes, parte do SPI², são um dos pilares da acumulação capitalista contemporânea, sendo reconhecido por diversos autores como um importante instrumento legal de poder ou controle tecnocientífico (Dardot e Laval 2017; Garcia dos Santos 2003b, 2003a; Mattelart 2005; Villares 2007).

Pierre Dardot e Christian Laval (2017, no capítulo 3 “A grande apropriação e o retorno dos ‘comuns’”), por exemplo, apresentam brevemente uma genealogia do SPI, enquadrando-o como parte importante do que ficou conhecido como “novos cercamentos” dos comuns. “[O]s camponeses autóctones expropriados do controle das sementes pelas grandes multinacionais da indústria agroalimentar, como a Monsanto; o desenvolvimento de patentes sobre seres vivos por pressão das empresas de biotecnologia; ou ainda a monopolização das patentes dos *softwares* mais usados no mundo pelas gigantes da informática, como a Microsoft” (Dardot e Laval 2017:106–7) seriam exemplos particularmente significativos da expansão do capitalismo por meio de novos cercamentos. Essa chave de leitura também relaciona a propriedade intelectual, e particularmente as patentes, com papel central que a expropriação ou a acumulação por despossessão teria mesmo no capitalismo contemporâneo da perspectiva de alguns de seus críticos (Dardot e Laval 2017).

Interessante perceber que, após uma mudança recente, o site do USPTO explicita os “direitos adicionais sobre as ‘partes’” no caso de uma “patente de planta”, justamente a mudança que Garcia dos Santos (2003b) aponta na epígrafe proveniente de seu texto “A virtualização da biodiversidade” e em outros ensaios deste livro que tem como subtítulo “o impacto sociotécnico da informação digital e genética”. Para o autor, o SPI ocupa posição central no que ele denomina “virada cibernética”, movimento recente “que selou a aliança entre o capital e a ciência e a tecnologia, e conferiu à tecnociência a função

¹Cuja tradução é: “Uma patente estadunidense dá a você, a inventora, o direito de ‘excluir outras de fazer, usar, oferecer ou vender’ uma invenção ou ‘importá-la’ para os EUA. Uma patente de planta lhe dá ainda direitos adicionais sobre as ‘partes’ de plantas (e.g., uma patente de planta sobre uma variedade de maçã incluiria direitos nas maçãs desta variedade de planta). O que é cedido não é o direito de fazer, usar, colocar a venda, vender ou importar a invenção, mas o direito de impedir outras de fazê-lo. Se alguém infringe a sua patente, você pode iniciar uma ação legal. Patentes estadunidense só tem efeito nos Estados Unidos e seus territórios e propriedades” (USPTO sem data-b).

²Para uma breve apresentação sobre o sistema internacional de patentes, o modo de implementação no Brasil e uma defesa de seu uso como fonte de descrição tecnológica, ver França (1997). Para uma apresentação endógena do sistema de patentes estadunidense, conferir USPTO (sem data-a).

de motor de uma acumulação que vai tomar todo o mundo existente como matéria-prima à disposição do trabalho tecnocientífico” (Garcia dos Santos 2003a:10–11). Essa mudança é perceptível ao constatar que “na ótica do biotecnólogo uma planta, um animal e até mesmo o ser humano reduz-se a um pacote de informações – porque o que interessa é o agenciamento das suas informações genéticas”, é o caso da apropriação privada de uma sequência genética de uma planta, vida em potencial. Mas também “a noção de trabalho e até mesmo a de produção de conhecimento também são profundamente alteradas, agora não pela informação genética, mas pela digital”, pois o “trabalho vai inscrever-se no software dos computadores, cuja capacidade de processar a informação na própria cadeia da produção lhe confere uma flexibilidade até então desconhecida”, de modo que “o próprio saber é transformado: cada vez mais é reelaboração, reprocessamento, reprogramação e recombinação de conhecimentos já existentes” (Garcia dos Santos 2003a:17). Em tempos de inteligência artificial gerativa, isso é mais evidente do que nunca³.

Segundo Dardot e Laval, podemos até mesmo falar de um “devir-mundo do capital” (Dardot e Laval 2017:144), na medida em que o capitalismo não cessa de expandir os horizontes de sua capacidade de geração de excedentes. Para os autores, “[e]xiste hoje um mercado mundial de direitos de propriedade sobre o conhecimento”, que são mercados criados através de estratégias jurídicas de criação de conhecimentos codificáveis e mercantilizáveis: “O argumento que leva à ampliação desse mercado e, com ela, à transformação do conhecimento é o da concorrência: o país, ou a empresa, que não cria cercamentos perderá posições no mercado mundial” (Dardot e Laval 2017:123).

Em relação às tecnologias digitais, Dardot e Laval (2017), amparados em Carlotta Hess e Elinor Strom, relacionam a apropriação capitalista do conhecimento com uma “captura do digital”, significando que as tecnologias digitais tornam possível a apropriação do conhecimento na forma de uma propriedade exclusiva⁴. Isso é evidente no caso de softwares proprietários (i.e., não livres), em que estratégias tecno-jurídicas limitam as possibilidades de reprodução de determinado programa de computador. O argumento de Kittler (1995) é similar ao constatar que “o software, se existisse [como algo separado do hardware], seria apenas um negócio de bilhões de dólares baseado nos elementos mais baratos do planeta [notadamente o silício, segundo elemento químico mais abundante na crosta terrestre]”. Segundo Kittler (1995), ao se separar as regras lógicas abstraídas do hardware por linguagens de programação de alto-nível, o software – que escondem o funcionamento material da usuária –, torna-se possível abstraí-las em propriedade intelectual pela criação de licenças de software e outros artefatos que determinam tais lógicas abstratas como produtos comerciais.

Segundo Kittler, evitaremos uma perspectiva sobre algoritmos como entidades “imateriais”, dizendo provisoriamente que um algoritmo é uma unidade elementar da computação digital: um programa ou um subprograma que executa uma operação em um computador digital, uma operação que converte um conjunto de entradas em um conjunto de saídas conforme uma determinada finalidade⁵.

³Conferir *infra* capítulo 3.

⁴Segundo Charlotte Hess e Elinor Strom, o problema fundamental do conhecimento se resume hoje a um problema de captura do digital. Segundo elas, embora o conhecimento seja um bem público puro, ao mesmo tempo não exclusivo e não rival, existem meios técnicos de transformá-lo em um bem que seja propriedade exclusiva de um detentor que só aceite cedê-lo mediante pagamento. Por ser capturável pelas novas tecnologias, o conhecimento é comparável a um recurso comum, muito semelhante aos comuns naturais. Por isso ‘precisa ser gerido, controlado, protegido para que lhe sejam garantidas sustentabilidade e preservação’. É nesse sentido exato que se deve falar de ‘comuns do conhecimento’ ou ‘comuns da informação’”. (Dardot e Laval 2017:173).

⁵Hoje, objeto de admiração, medo e debate, o termo algoritmo é, por vezes, utilizado em sentido amplo como sinônimo de

Um programa de computador, portanto, é geralmente composto por diversos algoritmos, embora também seja possível entender o próprio programa como um algoritmo que delega parte de suas funções a outros algoritmos. No caso apresentado no capítulo anterior, o PageRank é um dos principais algoritmos mobilizados pelo buscador Google Search.

2.2 Inteligência artificial, aprendizado de máquina e a propriedade intelectual de máquinas de apropriação intelectual

Gostaríamos de mostrar brevemente como, embora em nossa pesquisa documentos constitutivos do SPI sejam utilizados como fonte de descrição tecnológica, a própria aprendizagem de máquina executa uma operação análoga à patente, na medida em que suas capacidades derivam da “mineração de dados”, isto é, do ajuste de seus parâmetros numéricos a partir de relações encontradas em conjuntos de dados. Há então uma analogia funcional bastante interessante entre a patente e o AdM que, embora não aprofundaremos, sugerimos brevemente aqui⁶.

2.2.1 Inteligência artificial

Até então tratamos desse tipo de algoritmo sempre em paralelo com o termo “inteligência artificial”, mas sem precisar nenhum dos dois termos. Apresentaremos então breves linhas sobre seus significados e sua relação. Formalmente, o termo IA surge em 31 de agosto de 1955 quando John McCarthy, Marvin Lee Minsky, Nathaniel Rochester e Claude Elwood Shannon publicam “Uma proposta para o projeto de pesquisa de verão de Dartmouth sobre inteligência artificial [*A proposal for the Dartmouth summer research project on artificial intelligence*]”:

Propomos que um estudo de 2 meses e 10 homens [*a 2 month, 10 man study*] sobre inteligência artificial seja realizado durante o verão de 1956 no Dartmouth College em Hanover, New Hampshire. *O estudo deve prosseguir com base na conjectura de que todo aspecto do aprendizado ou qualquer outra característica da inteligência pode, em princípio, ser descrito com tanta precisão que uma máquina pode ser feita para simulá-lo.* Será feita uma tentativa de descobrir como fazer as máquinas usarem a linguagem, formarem abstrações e conceitos, resolverem tipos de problemas agora reservados aos humanos e se aprimorarem. Acreditamos que um avanço significativo pode ser feito em um ou mais desses problemas se um grupo cuidadosamente selecionado de cientistas trabalharem juntos nele por um verão. (McCarthy et al. 1955:2, grifo nosso)

Entretanto, uma rápida passada de olho no documento mostra como, na verdade, propunha-se ali um desenvolvimento de temas já existentes em diversas áreas do conhecimento. Segundo Elish e plataforma, inteligência artificial ou até mesmo internet. Uma forma muito presente no discurso contemporâneo é se referir à ordenação automática de conteúdos online como expressão d’o algoritmo” (no *feed* de um aplicativo, por exemplo). Essa formulação passa a impressão de que o termo se referiria a uma entidade autônoma e quase mítica e que, por si só, poderia ser objeto de uma pesquisa sociológica. Em certo sentido, obter uma definição tecnicamente precisa e sociologicamente eficaz é um dos objetivos de nosso trabalho, sendo grande parte dele um esforço para construir uma tal definição e que tenha capacidade explicativa para o problema mais específico de sua agência. Para nossa formulação final, conferir especialmente o capítulo [O que é um algoritmo? por uma definição de mediação algorítmica](#).

⁶Os eventos apresentados aqui podem ser também ser consultados, de maneira sintetizada, no apêndice B.

boyd (2018:60) o projeto “reuniu os últimos avanços nas ‘ciências de sistemas’, incluindo cibernética, teoria da informação, teoria de sistemas e ciência cognitiva” ao redor do “problema da inteligência artificial”. Entre os aspectos explicitados ao redor desse problema⁷, por exemplo, jaz “redes neurais”, do qual voltaremos a falar abaixo, mostrando que o termo já era estabelecido no momento de proposição do documento⁸. Por esse motivo, outras autoras argumentam que os esforços mobilizados em relação a uma inteligência artificial antecedem a criação o termo. De fato, mesmo a famosa “Maquinaria computacional e inteligência [*Computational machinery and intelligence*]” de Alan Turing (1950) que propôs o que hoje conhecemos como “teste de Turing”⁹ precede a data do documento.

Yuk Hui (2020), é um dos autores que faz esse deslocamento, apontando que “somos frequentemente lembrados dos trabalhos de Warren McCulloch e Walter Pitts – pesquisadores que participaram das Conferências Macy¹⁰ – quando nos referimos a redes neurais artificiais” (Hui 2020:165), um dos algoritmos de IA mais conhecidos¹¹. No mesmo sentido, em um texto endógeno, Goodfellow, Bengio e Courville (2016:12 *et seq.*) separam o desenvolvimento das RNA em três momentos: a *cibernética* entre 1940s e 1960s, o *conexionismo* entre 1980s e 1990s e o *aprendizado profundo* desde 2006. Os momentos de menor debate intelectual no campo são comumente chamados de “invernos da IA” (cf. Cardon, Cointet, e Mazieres 2018). Portanto, podemos considerar que o campo tecnocientífico que ainda consideramos hoje como IA, surge nesse momento de efervescência que podemos denominar “cibernética”.

Entretanto, para dar conta do que a IA se tornou, acreditamos que defini-la como um campo principalmente tecnocientífico, mesmo que desde a origem relacionado com outros¹², seria insuficiente. Elish e boyd (2018) apresentam como seu significado mudou ao longo da história, tendo mais recentemente substituído parcialmente o termo “big data” que dizia respeito mais a um campo mercadológico ou em-

⁷Que são: 1) Computadores automáticos; 2) Como pode um computador ser programado para utilizar a linguagem; 3) Redes neurais; 4) Teoria do tamanho de um cálculo; 5) Auto-melhoramento; 6) Abstrações; 7) Aleatoriedade e criatividade (McCarthy et al. 1955:2).

⁸“Como um conjunto de neurônios (hipotéticos) pode ser arranjado de modo a formar conceitos. Considerável trabalho teórico e experimental foi feito sobre esse problema por Uttley, Rashevsky e seu grupo, Farley e Clark, Pitts e McCulloch, Minsky, Rochester e Holland, e outros. Resultados parciais foram obtidos, mas o problema precisa de mais trabalho teórico” (McCarthy et al. 1955:2).

⁹O “teste de Turing”, ou “jogo de imitação” como chamado pelo autor, foi uma tentativa de avançar os esforços de uma definição de máquina inteligente: diante da incapacidade de responder de forma não-ambígua a questão “máquinas podem pensar?”, Turing (1950) propõe responder, ao invés disso, se computadores digitais poderiam interagir com humanos em uma situação específica sem serem percebidos como máquinas. Ou seja, para uma máquina “passar no teste”, ela não necessita responder sinceramente, ou responder corretamente, mas sim ser percebida como humana. Interessante notar como os recentes MdL, como o chatGPT da OpenAI, ainda operam sob o mesmo princípio (conferir capítulo 3).

¹⁰Em uma nota da tradução, lemos que elas foram: “um total de 160 conferências interdisciplinares realizadas na cidade de Nova York entre 1946 e 1953 com o objetivo declarado de fomentar a comunicação entre diversos ramos da ciência e reestabelecer uma unidade entre eles. Parte de seus trabalhos levou ao lançamento da base fundacional da cibernética” (Hui 2020:165, n. 7).

¹¹“Esta interpretação da história é aparentemente diferente da narrativa clássica sobre a origem do termo ‘inteligência artificial’, que dizem ter sido criado durante a Conferência de Dartmouth de 1956, associado a cientistas e pesquisadores como Marvin Minsky, John McCarthy e Claude Shannon, entre outros. Essa narrativa clássica continua com o avanço no desenvolvimento de uma inteligência artificial simbólica fraca (ou o que John Haugeland chama de ‘boa e velha inteligência artificial’) para uma mais forte e, por fim, leva à fantasia dos dias de hoje de uma superinteligência” (Hui 2020:165–66).

¹²Haraway (2023a, “Um manifesto ciborgue”) mostra como podemos falar de uma origem militar das tecnociências cibernéticas. Latour (2011:267) também mostra como, ao menos nos EUA, a maioria da pesquisa e desenvolvimento em geral é realizada pela indústria, com financiamento público e destinado à “defesa”. Finalmente, mesmo na ocasião de proposição de McCarthy et al. (1955), embora McCarthy e Minsky fossem acadêmicos, Shannon estava na então Bell Telephone Laboratories e Rochester, na IBM.

presarial. Nesse sentido, soluções e produtos “inteligentes”, como o smartphone [“celular inteligente”, em oposição ao *dumbphone*, ou seja, um “telefone estúpido” comum], cidades inteligentes [*smart cities*] e automação residencial são frequentemente incluídas no termo.

Diante disso, neste trabalho tomaremos a seguinte decisão metodológica: consideramos como IA um campo, originalmente tecnocientífico, mas que se tornou, progressivamente, descritivo de produtos e métodos em setores econômicos e corporativos. Paralelamente a isso, também incluímos no termo o universo cultural ao redor de robôs, máquinas inteligentes e etc. que criam um imaginário sociotécnico associado (literatura e cinema de ficção científica, por exemplo). Ou seja, tomaremos IA como um campo amplo, heterogêneo e com limites não tão bem definidos, centrados na ideia de “máquinas inteligentes”¹³. Em contrapartida, partimos da premissa de que é possível encontrar uma definição com sentido mais estrito para o termo AdM, entendido inicialmente como um conjunto específico e bem definidos de técnicas e métodos que são utilizados de formas diversas¹⁴ para resolver tarefas por meio do processamento de dados. Assim, durante o trabalho, falaremos de “sistemas de IA” e “tecnologias de IA”, mas de “técnicas de AdM”, “modelos de AdM” e “algoritmos de AdM”.

2.2.2 Aprendizado de máquina e redes neurais artificiais

O artigo “Alguns Estudos em Aprendizado de Máquina Utilizando o Jogo de Damas [*Some Studies in Machine Learning Using the Game of Checkers*]” (Samuel 1959) é frequentemente considerado como marco consolidador do termo AdM, sendo a ideia presente nele sintetizada como “aprendizado de máquina é o campo de estudos que dá a computadores a habilidade de aprender sem serem explicitamente programados” (e.g. Brown 2021). Essa definição tem o problema de conter o termo “aprendizado” na própria explicação do AdM. Uma definição de aprendizado amplamente utilizada e apresentada por Goodfellow, Bengio e Courville (2016:97 *et seq.*) é a de Mitchell (1997): “um programa de computador é tido como tendo aprendido da experiência E com relação a uma classe de tarefas T e métrica de desempenho [*performance metric*] P, se seus desempenhos nas tarefas T, medidas por P, melhoram com a experiência E”. Alternativamente, a IBM (Anón 2024) define AdM em seu site como “um ramo da inteligência artificial (IA) e ciência da computação que foca em utilizar dados e algoritmos para capacitar [*enable*] IA a imitar o modo como humanos aprendem, aumentando gradualmente sua acurácia [taxa de acertos].” O AdM diz respeito então a uma classe mais restrita de algoritmos, que funcionam por meio do processamento de dados.

Como já mencionado, um dos primeiros algoritmos de AdM e um dos mais importantes hoje são as RNAs. De acordo com a seção “o que é uma rede neural?” de Haykin (2009), importante referência da área, elas são: “um processador distribuído massivamente paralelo feito de unidades de processamento que tem uma propensão natural para armazenar conhecimento exponencial e torná-lo disponível para

¹³Segundo Goodfellow, Bengio e Courville (2016:1), “inventoras sonham há muito tempo em criar máquinas que pensam. Este desejo remonta, pelo menos, ao tempo da Grécia antiga” citando Pigmalião, Dédalo e Hefesto como inventores e Galateia, Talos e Pandora como vidas artificiais. E ainda, remetendo ao texto de Ada Lovelace, que: “[q]uando os computadores programáveis foram concebidos pela primeira vez, as pessoas interrogavam-se sobre a possibilidade de essas máquinas se tornarem inteligentes, mais de cem anos antes de uma ter sido construída”. Desse modo, é possível entender a própria história da computação de maneira imbricada com a história da inteligência artificial.

¹⁴Sendo a principal técnica observada na pesquisa os algoritmos de RNAs e suas variações como aprendizado profundo e modelagem de linguagem. Entretanto, outros algoritmos são explicitamente citados nas patentes como, por exemplo, máquinas de vetores de suporte (SVM, *support vector machines*) e árvores de decisão (*decision trees*).

uso” (Haykin 2009:2). Segundo o autor, elas foram originalmente baseadas no sistema nervoso de animais e lembram o cérebro por ao menos dois motivos: “o conhecimento é obtido pela rede de seu ambiente através de um processo de aprendizagem” e “as forças de conexão interneural, conhecidas como pesos sinápticos, são usadas para armazenar o conhecimento adquirido” (Haykin 2009:2). Nesta definição, fica evidente a intenção de aproximação das redes neurais “naturais” e “artificiais”, de modo a entender as primeiras como computadores humanos e as segundas como modelos do cérebro. Embora nem todas as pesquisadoras da área compartilhem dessa premissa, de fato as primeiras redes neurais foram desenvolvidas baseadas em estudos das ciências biológicas, enquanto que, hoje, é muito comum falarmos de nossa cognição em termos computacionais¹⁵. Goodfellow, Bengio e Hinton expõem esta questão da seguinte forma:

Enquanto que a neurociência é uma fonte importante de inspiração, é preciso não tomá-la como um guia rígido. Nós sabemos que os neurônios reais [biológicos] computam funções muito diferentes que as modernas unidades lineares retificadas [unidades ReLU, um dos principais tipos de neurônio utilizados em aprendizado profundo], mas maior realismo neural ainda não levou a uma melhoria no desempenho do aprendizado de máquina. Ademais, enquanto a neurociência inspirou com sucesso várias *arquiteturas* de redes neurais, nós ainda não sabemos o suficiente sobre o aprendizado biológico para que a neurociência seja muito orientadora para os *algoritmos de aprendizado* que usamos para treinar essas *arquiteturas* (Goodfellow et al. 2016:15).

No mesmo livro, *Deep learning* [Aprendizado profundo], que versa sobre a vertente contemporânea de aprendizagem de máquina utilizando redes neurais profundas, isto é, de múltiplas camadas, podemos ler que:

O precursor mais antigo do aprendizado profundo moderno eram modelos lineares simples motivados pela perspectiva neurocientífica. Esses modelos foram designados para receber um conjunto de valores de entrada x_1, \dots, x_n [, ou seja, o vetor \mathbf{x}] e associar a eles uma saída y . Esses modelos aprenderiam uma série de pesos [sinápticos] w_1, \dots, w_n [, ou seja, \mathbf{w} ,] e calculariam sua saída $f(\mathbf{x}, \mathbf{w}) = x_1 w_1 + \dots + x_n w_n$ [isto é, a saída y de uma rede neural linear, ou seja, de um único neurônio artificial, poderia ser calculada como uma função da soma dos valores de entrada x_i ponderada pelos pesos sinápticos w_i associados a cada uma delas para todo i] (Goodfellow et al. 2016:14).

Mais especificamente, o primeiro modelo de IA neuroinspirado foi o proposto por McCulloch e Pitts em seu artigo “Um cálculo lógico das ideias imanentes na atividade nervosa [*A logical calculus*

¹⁵Como sugerimos em Gonçalves e Padovan (2024a:12), acreditamos que “há [hoje] uma inversão: os algoritmos que pretendiam, originalmente, modelar o sistema nervoso animal passam a ser utilizados para explicá-lo. O cérebro se torna um ‘computador de processamento distribuído’ e a visão humana, ‘uma tarefa de reconhecimento de padrões’. Desse modo, o aprendizado humano é referido como a criação de regras comportamentais por meio da plasticidade cerebral e da adaptação ao ambiente, da mesma forma como um algoritmo de rede neural artificial calcularia os pesos sinápticos ao ser exposto a uma tarefa computacional específica: ‘uma *rede neural* é uma máquina que foi desenhada para modelar a forma como o cérebro realiza uma tarefa particular ou uma função de interesse’ (Haykin, 2009, p. 2), embora hoje também expliquemos o cérebro humano com uma surpreendente proximidade em relação a uma rede neural artificial”. Segundo Goodfellow, Bengio e Courville (2016:15), “hoje, a neurociência é tida como uma importante fonte de inspiração para pesquisadoras do aprendizado profundo, mas não é mais um guia proeminente para o campo”.

of the ideas immanent in nervous activity]” (McCulloch e Pitts 1943) (representado na figura 2.1). O artigo propõe um modelo matemático do neurônio que seria capaz de ser utilizado para representar proposições lógicas. Apesar de apresentar algumas diferenças em relação ao modelo neuronal usado com mais frequência atualmente, esse modelo já possui os seus principais elementos (entradas, função de ativação e saída). Já as principais diferenças são que no modelo de McCulloch e Pitts as entradas e a saída do neurônio são necessariamente binárias (0 ou 1) e suas entradas não são mediadas por pesos. O modo de funcionamento do AdM e das RNAs em particular serão apresentados em detalhe na parte II.

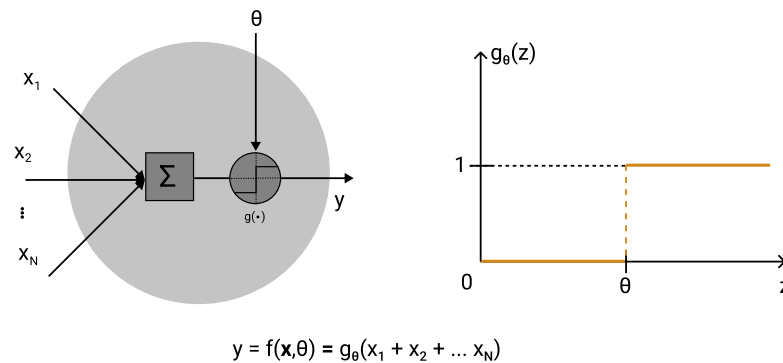


Figura 2.1: Esquema de um neurônio tal qual proposto por McCulloch e Pitts (1943) e função matemática que ele computa. A saída y é o resultado da função $g(\cdot)$ para a soma de entradas $(x_1 \dots x_N)$. Ou seja, se a soma das entradas é maior que θ , a saída é igual a um, senão, ela é igual a zero. Fonte: adaptado de Von Zuben e Attux (2007).

Outros trabalhos importantes deste momento, a cibernética, na separação de Goodfellow et al. (2016), foram a publicação do livro *The organization of behavior* [A organização do comportamento] escrito pelo psicólogo Donald Hebb (1949) que propõe um mecanismo de plasticidade sináptica frequentemente sintetizado como “células que disparam juntas, permanecem juntas”; a proposição da arquitetura neuronal “perceptron” por Rosenblatt (1958) que, diferente do neurônio de McCulloch-Pitts, aceita entradas não binárias e já possui pesos; e a proposição de um método automatizado de aprendizado, o ADALINE (*adaptive linear element*, ou elemento linear adaptativo) proposto por Widrow e Hoff (1960).

Já na transição do paradigma cibernético para o conexionista¹⁶, percebemos que a teoria hebbiana foi complementada pela teoria conexionista da mente proposta pelo economista neoliberal Friedrich Hayek em seu *The sensory order* [A ordem sensorial] (1952), de modo que poder-se-ia dizer que Hayek “roubou [a ideia de] reconhecimento de padrão [da neurologia e da cibernética] e a transformou em um princípio neoliberal de regulação do mercado” (Pasquinelli 2021:160). Pasquinelli comenta que “[e]m 1958, Frank Rosenblatt definiu o Perceptron (a primeira rede neural artificial operacional para reconhecimento de padrões) como ‘conexionista’ [em oposição à vertente dita ‘simbólica’ da IA, mais reconhecida

¹⁶Do ponto de vista da eficácia, os modelos neuronais – por serem lineares, isto é, unicamente constituídos por uma soma de valores ponderados – não apresentavam uma capacidade notável: por exemplo, não é possível representar linearmente a operação lógica de disjunção exclusiva (XOR) (cf. Goodfellow et al. 2016:14). A necessidade de um aprendizado e a não-garantia de otimalidade faz com que eles não sejam vantajosos em relação à chamada “solução fechada”, o cálculo exato da solução por vias matemáticas.

naquele momento^{17]} e reconheceu que o trabalho de ‘Hebb e Hayek’ foram ‘o mais sugestivo’ para o seu próprio” (Pasquinelli 2021:160).

O perceptron, entendido como “primeira verdadeira máquina conexionista” (Cardon et al. 2018:185), foi muito influente no campo, de tal maneira que ainda hoje uma das principais classes desses algoritmos é chamada de perceptron de múltiplas camadas (MLP, *multilayer perceptron*) (figura 2.2). Já sobre a proximidade do economista com o campo, podemos dizer que a influência foi bidirecional. Se, por um lado, a cibernética e a modelagem computacional do cérebro influenciaram a obra de Hayek; por outro lado, é bastante evidente que os desenvolvimentos do campo do aprendizado profundo, ainda hoje, são fortemente pautados por ideias de maximização de acertos e minimização de erros, que ressoam com as tendências neoliberais de maximização do desempenho e do prazer (Dardot e Laval 2016).

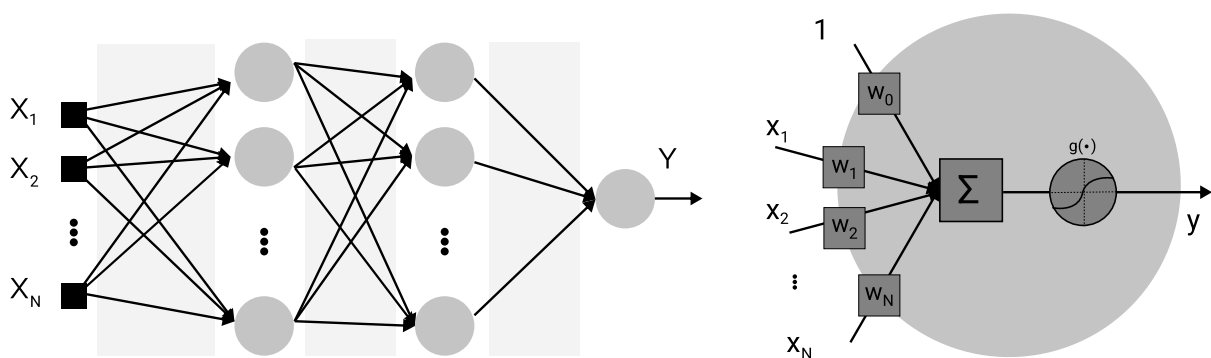


Figura 2.2: À direita, esquema de um neurônio de tipo perceptron, à esquerda, esquema de uma rede neural MLP. Cada círculo cinza na rede representa um neurônio, de modo que, conectados, a rede é capaz de computar uma função complexa. Enquanto que X_1, \dots, X_N representam as entradas da rede e Y sua saída; $x_1 \dots x_N$ representam as entradas de cada neurônio e y suas saídas. Os pesos sinápticos ($w_1 \dots w_N$) que multiplicam as entradas de cada neurônio são parâmetros aprendidos no processamento de dados que determinam a função a ser computada. Fonte: adaptado de Gonçalves (2022).

Nesse contexto, o principal avanço do conexionismo foi a proposta de um algoritmo de cálculo de derivadas utilizando a retropropagação de erro [*error backpropagation*] em uma rede (Rumelhart, Hinton, e Williams 1986), uma das principais ideias utilizada ainda hoje. Esse método (em geral, associado com o algoritmo de otimização chamado “gradiente descendente” utilizado na etapa de aprendizado) se baseia no cálculo de taxas de variação do erro em relação aos pesos que são utilizadas no ajuste dos próprios pesos¹⁸. Outro algoritmo influente proposto nesse momento foi o neurônio memória de curto longo prazo (LSTM, *long short-term memory*) (Hochreiter e Schmidhuber 1997), um neurônio que foi muito utilizado no contexto de redes neurais recorrentes (redes com realimentações internas na etapa de inferência). Segundo Goodfellow, Bengio e Courville (2016:16), “[a] ideia central no conexionismo é que um grande número de unidades computacionais simples podem alcançar comportamento inteligente quando são enredados conjuntamente [*networked together*]. Esse insight se aplica igualmente

¹⁷Para o debate IA conexionista versus simbólica, conferir Cardon et al. (2018). Um dos principais algoritmos da IA simbólica, baseada em regras e não em dados, foram os sistemas especialistas (*expert systems*) (Cardon et al. 2018:191 et seq.).

¹⁸Uma apresentação interessante e relativamente didática do funcionamento desses algoritmos pode ser vista nos vídeos sobre o tema publicados pelo canal do YouTube 3Blue1Brown (2017a, 2017b, 2017b).

para neurônios no sistema nervoso biológico e para as unidades escondidas [*hidden units*] em modelos computacionais” (Goodfellow et al. 2016:16).

Finalmente a concepção das redes neurais profundas (*deep neural networks*) e do aprendizado profundo, marcaram a passagem para a última etapa de desenvolvimento das redes neurais, quando elas passam a ser entendidas como um aprendizado de representação (LeCun, Bengio, e Hinton 2015; cf. *infra* **Imaginação**). Para que isso seja possível, criam-se estruturas de redes com várias camadas sucessivas que transformam progressivamente a entrada, daí o adjetivo “profundo”. De acordo com Goodfellow, Bengio e Courville (2016:26):

Eu resumo, aprendizado profundo é uma abordagem para aprendizado de máquina que parte fortemente do nosso conhecimento sobre o cérebro humano, estatística e matemática aplicada, como se desenvolveu nas últimas várias décadas. Nos anos recentes, aprendizado profundo teve um grande crescimento em sua popularidade e utilidade, largamente como resultado de computadores mais potentes, bases de dados mais amplas e técnicas para treinar redes mais profundas. Os anos seguintes estão cheios de desafios e oportunidades para melhorar ainda mais o aprendizado profundo e para levá-lo a novas fronteiras.

Como apontado por Pasquinelli (2023) e Cardon et al. (2018), o evento que marca mais significativamente o sucesso do aprendizado profundo é a vitória da competição de visão computacional ImageNet pela rede convolucional AlexNet proposta por Krizhevsky, Sutskever, e Hinton (2017). Para Pasquinelli (2023:14), “[o] aprendizado profundo evoluiu da extensão de técnicas de reconhecimento de padrões visuais dos anos 1950s para dados não visuais, que agora incluem texto, áudio, vídeo e dados comportamentais das origens mais diversas”. Já para Cardon e colaboradores (2018:176), o aprendizado profundo é “[u]m método de aprendizagem [que] oferece o processamento o mais ‘bruto’ possível das entradas, eliminando qualquer modelagem explícita das características dos dados e otimizando a previsão a partir de enormes amostras de exemplos”, “a transição de uma máquina hipotético-dedutiva para uma máquina indutiva”.

Outros marcos importantes do aprendizado profundo são: a proposta de um modelo de linguagem probabilístico por Bengio et al. (2003) e a abordagem de codificação vetorial de palavras (word2vec) desenvolvida em Mikolov et al. (2013); a proposta de um mecanismo de atenção (Luong, Pham, e Manning 2015) e da arquitetura transformer (Vaswani et al. 2017), que foram determinantes para o desenvolvimento dos atuais MdLs como os utilizados no ChatGPT da OpenAI. Já em relação aos modelos geradores de imagem, como o Dall-E da OpenAI, um marco foi a proposta de uma rede adversarial gerativa (GAN, *generative adversarial network*) (Goodfellow, Shlens, e Szegedy 2015) e, mais recentemente, o mecanismo de difusão (Rombach et al. 2022), que ganhou notoriedade por sua capacidade de geração de imagens a partir de texto. Além disso, outros esforços notáveis são os avanços na área de aprendizado por reforço (Mnih et al. 2015) cujos desenvolvimentos são atualmente utilizados no ChatGPT e similares na forma de aprendizado por reforço com reatualização humana (RLHF, *reinforcement learning from human feedback*) (Christiano et al. 2023; Ouyang et al. 2022).

2.2.3 Aprendizado de máquina como apropriação intelectual

Diante disso, percebemos que a própria aprendizagem de máquina é uma forma de apropriação intelectual, uma “captura do digital” (Dardot e Laval 2017), ou uma “máquina indutiva” (Cardon et al. 2018),

que faz emergir um potencial de conhecimento (previsão, classificação, geração, entre outros) a partir do processamento de dados que cria automaticamente relações entre eles. Como no caso da patente e das licenças de software, não é que esse conhecimento seja essencialmente apropriável de maneira privada, mas sim que, por existir uma abertura para isso, a tecnociência e o capitalismo criam ativamente aparatos para uma apropriação monopolista desse potencial. Desse modo, o objeto desta pesquisa está envolvido em múltiplos níveis de apropriação intelectual, pois não só as patentes e os softwares são mecanismos de acumulação capitalista pela privatização de um comum, mas também o próprio AdM deriva suas capacidades de uma captura sobre processos sociais que são, cada vez mais, digitalizados, dadoficados e impelidos à circular em dentro de sistemas de computação com algoritmos de AdM¹⁹.

2.3 Uma sociologia da tecnologia das patentes da Google

Uma apresentação do campo da sociologia da tecnologia foge ao escopo deste trabalho²⁰, mas podemos dizer que são duas as principais influências para a perspectiva geral adotada por este trabalho. Em primeiro lugar, a sociologia do pensador brasileiro Laymert Garcia dos Santos (2003b) – e de outras cientistas sociais influenciadas por ele – que, sobretudo a partir dos anos 1990, se debruçou a pensar (e politizar o debate sobre) “as novas tecnologias”, isto é, aquelas implicadas com informação digital e genética. Sua sociologia, que com Simondon se pretende não autocrática (Garcia dos Santos 2022), mostra como há um funcionamento conjunto entre os desenvolvimentos tecnocientíficos e as novas formas de acumulação capitalista, o que já estava presente, por exemplo, na análise sociológica do capitalismo incipiente em Marx (2017), mas que se intensifica após a virada cibernética assentada na noção de informação. Em grande medida, nosso trabalho é uma tentativa de ser uma pequena contribuição para esta análise, expandindo-a para as tecnologias de IA. Em segundo lugar, a possibilidade inaugurada por Bruno Latour (2013 com Steve Woolgar; 2011, 2017), como representante de uma vertente dos estudos de ciência e tecnologia (STS, *science and technology studies*), e por Donna Haraway (2023a), dos estudos feministas de ciência e tecnologia, também na virada do século XX para o XXI, de pensar o ciborgue, o híbrido, os agregados multiespécie, as naturezas-culturas, para além das dicotomias modernas natureza/sociedade e natural/técnico. A principal consequência disso é a possibilidade de uma sociologia dos não humanos, preocupada com a agência própria de seres que, até então, tiveram sua participação social denegada (Latour 1992, 2019). Desse modo, nosso intuito de pensar a ação social do aprendizado de máquina é também legatário dessa maneira de fazer ciência social.

Em ambos os casos, há grande diálogo com a antropologia e a filosofia, sobretudo de uma tradição que podemos denominar “menor” (Deleuze e Guattari 2018): Bergson, Simondon, Deleuze, Guattari e Stengers. Entendemos que isso se justifica, na medida em que toda pesquisa sociológica possui uma metafísica e uma antropologia subjacentes, embora nem sempre explicitadas. Nesse sentido, valorizamos o esforço de Pedro Ferreira (2004) que apresenta mais explicitamente uma sociologia da tecnologia que já está esboçada, ou que pode ser derivada, a partir das noções de “máquina” e de “filo maquinico” em

¹⁹Quisemos apenas sugerir este movimento aqui, relacionando com o SPI e com a ideia de software, porém expandiremos este argumento, ao menos de maneira indireta, nos capítulos 3 e 4.

²⁰Para uma perspectiva interessante de como a questão da tecnologia é tratada em Marx e no marxismo, em Simondon e em textos de ficção científica, ver Garcia dos Santos (1981). Uma apresentação interessantes de algumas abordagens socio-antropológicas sobre a tecnologia é apresentada em Kasper (2006). Para exemplos históricos e uma síntese metodológica da abordagem dos estudos de ciência e tecnologia, da qual também partimos, conferir Latour (2011).

Deleuze e Guattari (2011b, 2011–2012). Essa perspectiva não-antropocêntrica, pois não parte de uma noção apriorística de “humano”, estaria preocupada em estudar uma história ao mesmo tempo natural e cultural, ou melhor natural-cultural, da transformação de potenciais maquínicos virtuais em mecanismos técnicos atualizados. Dito de outro modo, de uma perspectiva anterior ao corte sujeito/objeto, natureza/cultura, humano/máquina, essa sociologia visaria estudar a concretização histórica e contingente de associações heterogêneas ou híbridas de componentes considerados “humanos” e “não humanos”. Essa proposta de sociologia, em sinergia com aquelas de Garcia dos Santos e Latour, dá um papel central aos objetos técnicos e à invenção, se desenvolvendo depois em um programa para o estudo de “processos tecnicamente mediados de associação” (Ferreira 2010). Seguindo tais esforços, entendemos a tecnologia como objeto privilegiado de estudo da sociologia, pois ela é, ao mesmo tempo, concretização final de associações heterogêneas e contexto geral de onde partem novos processos associativos²¹. Daí nosso esforço em compreender a ação sociotécnica do aprendizado de máquina, uma das técnicas mais proeminentes da contemporaneidade.

Ao propor uma caracterização sociotécnica do AdM, nosso trabalho também tem uma pretensão de transdisciplinaridade, se conectando com campos das engenharia e ciência da computação. Especialmente, dialogamos com o campo de justiça algorítmica (*algorithmic fairness*): FAT, e depois FAcT, é o nome da conferência da Associação para a Maquinaria Computacional (ACM, *Association for Computing Machinery*) preocupada com questões relativas à “justiça”, “responsabilidade [*accountability*]” e “transparência”. A sigla foi posteriormente expandida para FATE (como usado pela Microsoft) ou FATES, incluindo, respectivamente, “ética” e “segurança e proteção [*safety, security*]”. Os trabalhos de Prates, Avelar, e Lamb (2020) e Cheong et al. (2024) ilustram as preocupações deste campo, bastante centrados em temas como o *viés*, que será explorado por nós na parte III.

Finalmente, encontramos nos estudos críticos de algoritmos, diálogo profícuo, sendo a abordagem cartográfico-anatômica dos trabalhos de Matteo Pasquinelli, Kate Crawford e Vladan Joler em relação à IA e ao AdM bastante inspiradora para o nosso trabalho (Crawford e Joler 2018; Joler 2020; Joler e Petrovski 2016; Pasquinelli e Joler 2021). Desse modo, a parte II desta dissertação culmina em um experimento cartográfico bastante incipiente, mas que busca encontrar em tais autores amparo para novas abordagens explicativas das tecnologias contemporâneas.

2.3.1 Dificuldades da pesquisa sobre algoritmos

Embora encontremos em Latour (2011) valiosos recursos metodológicos para o estudo da tecnociência, a pesquisa de fenômenos mediados por algoritmos – como o caso da modulação de atividade – traz algumas dificuldades próprias. Nesse âmbito, Rob Kitchin (2014b) apresenta três grandes questões inerentes ao estudo de algoritmos:

1. a *dificuldade de acesso*, muitas vezes referido como o “problema da caixa-preta”, que diz respeito não só a inacessibilidade de alguns algoritmos, propriedade privada de empresas que os guardam sob regime de segredo industrial, mas também ao uso de porções proprietárias mesmo por softwares de código-aberto. No caso dos algoritmos de aprendizado de máquina, embora a arquitetura

²¹Dessa perspectiva a tecnologia é entendida como mobilização, estabilização e instituição de outros tempos, espaços e agências (Latour 2017), aquilo que “faz a sociedade durável” (Latour 1990), ou o “suporte e símbolo do transindividual” (Simondon 2020c).

- dos modelos seja frequentemente disponibilizada na forma de software livre ou de código-aberto, os conjuntos de dados utilizados em produtos reais dificilmente o são;
2. para além da natureza hermética e especializada de cada linguagem de programação, o carácter heterogêneo e embarcado de *sistemas algorítmicos*, que envolvem a composição de muitos algoritmos que se inter-relacionam, muitas vezes produzidos por pessoas diferentes, detidos por instituições distintas, congregando diferentes temporalidades e espacialidades. Um algoritmo quase sempre é parte de um complexo e extenso sistema algorítmico. Isso é especialmente verdadeiro no caso do aprendizado de máquina, que costuma ser apenas uma parte de um software: como veremos abaixo, aquela responsável pela “representação” (cf. [Imaginação](#));
 3. o aspecto reativo e performativo que faz com que possamos entendê-los como *agenciamentos dinâmicos* que envolvem contingências, constante mudança e são dependentes inclusive de sua interação com usuárias. Novamente, o caso do AdM é ainda emblemático, pois a própria construção do algoritmo é feita durante a interação com usuárias e automaticamente, de modo que sua agência é definida e redefinida na execução.

Nesse sentido, não há uma maneira simples de estudá-los e entendê-los. Kitchin (2014b) apresenta ainda algumas possibilidades de abordagem metodológica como o exame do código-fonte, a entrevista e etnografia de programadoras e a análise de seus efeitos externos (o que pode ser feito com entrevista/etnografia de usuárias, por exemplo). Mesmo assim, cada abordagem possui limitações, de forma que a composição simultânea de mais de uma abordagem é recomendada pelo autor. No nosso caso, entretanto, a fim de caracterizar a ação social dos algoritmos, isto é, analisar sua agência, focaremos nossa análise apenas em documentos que descrevem as tecnologias em questão, particularmente, patentes. Acreditamos que nosso trabalho, que dialoga sempre que possível com outras fontes, pode ser utilizado em composição com outros, que utilizem metodologias diversas, para uma análise mais ampla dos algoritmos de AdM.

2.3.2 Análise qualitativa de patentes como método para a investigação de algoritmos

A escolha pelo estudo de patentes pode ser visto como um ponto intermediário entre o do artigo científico, que apresenta um algoritmo isolado e com maior precisão, e o da implementação efetiva, em que o algoritmo está funcionando em condições sociais específicas²². Como apresentado acima (seção 2.1), a patente é um direito negativo que visa simultaneamente apresentar publicamente e restringir sua implementação ao proponente. Desse modo, como vantagem metodológica, a precisão descritiva, a justificativa da relevância social e a apresentação do modo de utilização da tecnologia, fazem da patente uma fonte privilegiada para o estudo de tecnologias emergentes. Além disso, as patentes oferecem uma

²²Se, por um lado, o estudo de artigos científicos nos ajuda em relação às duas primeiras dificuldades apontadas por Kitchin (2014b), dado que eles são produzidos visando a clareza e precisão, muitas vezes com códigos-fonte disponibilizados e descrições/implementações que isolam o algoritmo de sua rede efetiva. Nesse caso, perdemos justamente a ligação direta com a dimensão prática. O objetivo final do “produto” e os detalhes de como e onde se dará seu acoplamento social se perdem. Por outro lado, o estudo do produto em si, isto é, dos próprios algoritmos quando estes são disponibilizados com “código-aberto” ou em simuladores, inverte-se a questão. Se sabemos exatamente onde o algoritmo é utilizado, perdemos a clareza, o isolamento e a apresentação precisa que os artigos científicos nos provém. A primeira abordagem foi utilizada por nós em Gonçalves e Ferreira (2022b), enquanto que a segunda, em Gonçalves e Ferreira (2022a). Esses trabalhos foram imprescindíveis para a construção da abordagem aqui utilizada.

contextualização (*background* e tecnologias similares) e desenhos descritivos que contribuem para um entendimento sociotécnico destas (exemplo na figura 2.3). Entretanto, a principal desvantagem metodológica está ligada com a incerteza relativa à sua efetiva construção. Embora o conjunto de patentes seja um bom indicador do imaginário tecnocientífico e do sentido para o qual este campo, é impossível saber se uma patente específica se tornará ou não um produto²³.

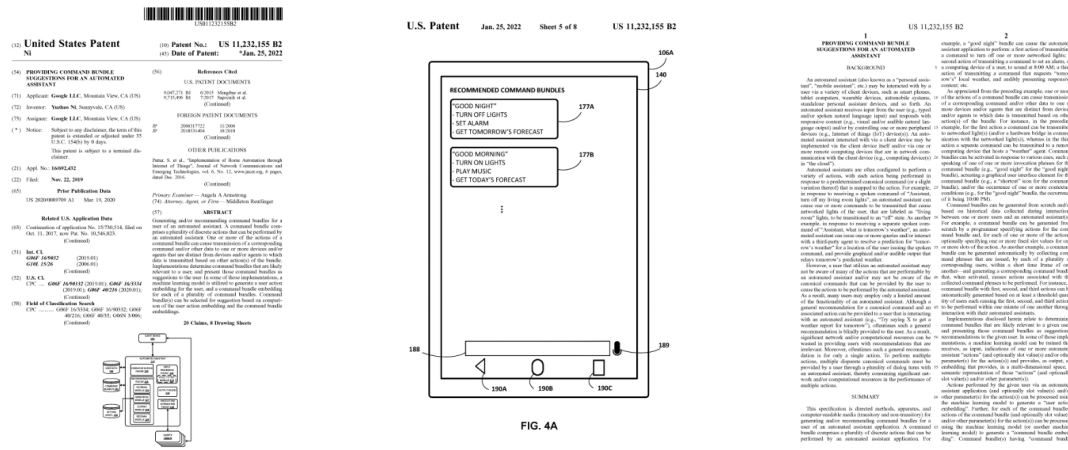


Figura 2.3: Exemplo de documento de patente (Ni 2022:1,7,11).

A construção do *corpus* de patentes sobre o qual este trabalho se apoia foi feito utilizando a ferramenta de busca Bases de dados de Texto Completo de Patente (PatFT, *Patent Full-text Databases*) do USPTO²⁴. Buscamos por patentes detidas pela Google com o termo *machine learning* (aprendizado de máquina) no título ou abstract, o que retornou 215 resultados²⁵. Destas, selecionamos as 18 patentes que se relacionavam com o fenômeno que nos propomos inicialmente a investigar, isto é, aquelas que partiriam de classificações e predições para, em alguma medida, substituir a agência da usuária, se configurando como uma modulação de atividade. No fim, analisamos mais detidamente nove patentes, apresentadas abaixo. A lista das patentes que utilizamos com título e abstract pode ser conferida no anexo [Lista de título e resumo das patentes](#).

Para a análise, utilizamos a nova ferramenta do USPTO que substituiu a ferramenta inicialmente utilizada para a busca, o Busca Pública de Patente (PPubS, *Patent Public Search*). Nela, é possível buscar patentes, ver informações detalhadas na forma de texto e, além disso, visualizar as páginas do documento (figuras 2.4 e 2.5). Além disso, por prover algumas facilidades adicionais, como por exemplo a apresentação das figuras e das reivindicações (*claims*) de modo mais sistemático, também utilizamos o serviço Patentes da Google ([Google sem data](#)) (figura 2.6). As patentes possuem modos próprios de organização, com parágrafos, figuras e detalhes das figuras enumerados, entretanto, por utilizarmos o documento como um todo em nossas análises, preferimos manter o estilo de citação de artigos, nos referindo na maioria dos casos, às páginas do arquivo oficial das patentes. Isso não quer dizer que

²³Um exemplo desse aspecto é apresentado por Latour (2011, no capítulo 3 “Máquinas”), mostrando como o motor idealizado por Diesel sofre alterações desde a invenção registrada na forma de uma patente até sua construção efetiva por engenheiros da MAN.

²⁴Esta ferramenta ([PatFT sem data](#)) foi descontinuada, tendo sido substituída pela *Patent Public Search* ([PPubS sem data](#)).

²⁵Busca realizada originalmente em 26 de junho de 2022 através da página de busca então disponível em <https://patft.uspto.gov/netahtml/PTO/search-adv.htm> com a query: ((TTL/(machine AND learning)) OR ABST/(machine AND learning)) AND AN/Google).

não utilizamos as outras referências numéricas quando conveniente, mas em geral as referências que fazemos estão no formato (Autoras Ano:Página).

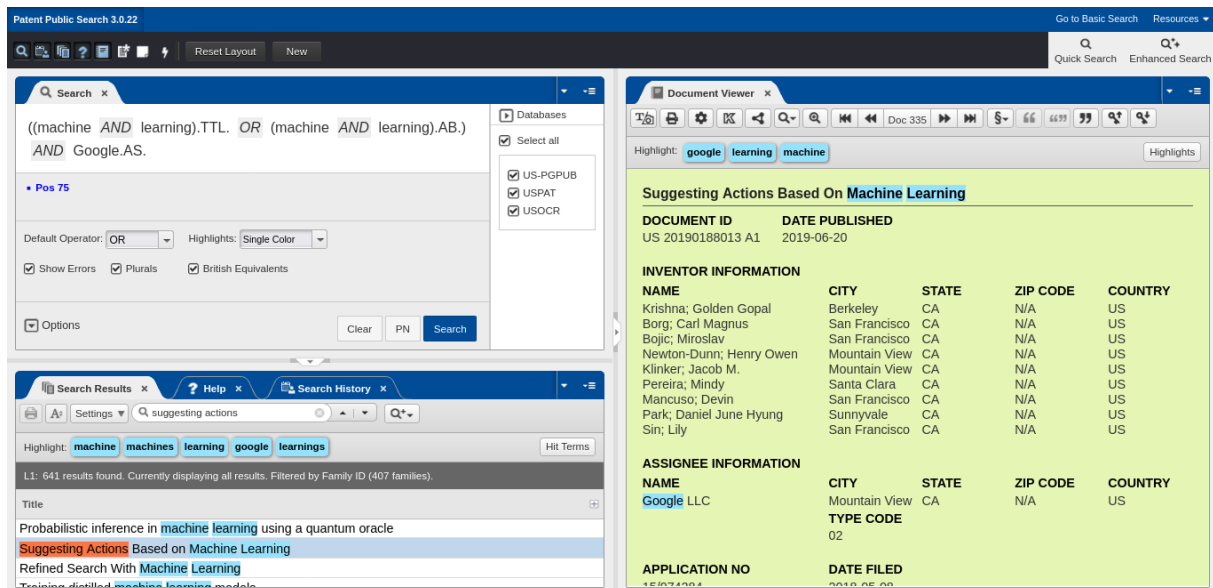


Figura 2.4: Captura de tela da ferramenta PPubS (sem data) obtida no dia 17/11/2024. Visualização de informações textuais da patente de Krishna et al. (2021).

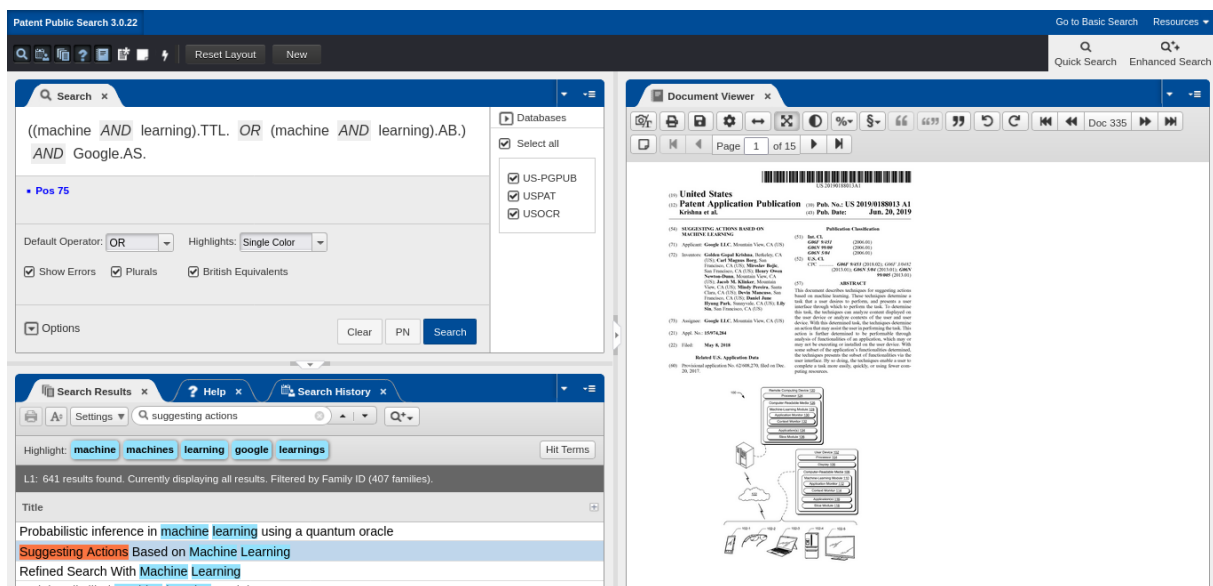


Figura 2.5: Captura de tela da ferramenta PPubS (sem data) obtida no dia 17/11/2024. Visualização das figuras da patente de Krishna et al. (2021).

Diante disso, visamos partir de uma análise qualitativa do *corpus* empírico a fim de propor uma caracterização sociotécnica do funcionamento do AdM, com foco nos processos de modulação de atividade²⁶. Isso foi feito amparado em outros trabalhos de sociologia que se debruçaram sobre patentes como material empírico tais quais Diaz-Isenrath (2005), Amadeu da Silveira (2019), Machado (2019),

²⁶Para métodos quantitativos, conferir Breitzman e Moguee (2002) e Abbas, Zhang, e Khan (2014), citados em Machado (2019).

Garcia dos Santos e Ferreira (2008) e Crawford e Joler (2018). Tendo exposto nossa metodologia, a próxima seção apresenta o material propriamente dito.

The screenshot displays the Google Patents page for patent US11366812B2. The interface includes a search bar at the top with the patent number 11366812. The main content is organized into several sections:

- Abstract:** Describes techniques and a framework for gathering information about developing events from multiple live data streams and pushing new pieces of information to interested individuals.
- Images (8):** A grid of eight thumbnail images showing various diagrams and flowcharts related to the patent.
- Classifications:** Lists the primary classification G06F16/24568 (Data stream processing; Continuous queries) and provides a link to view more classifications.
- Landscapes:** Shows two categories: "Engineering & Computer Science" and "Theoretical Computer Science", each with a search icon.
- Description:**
 - BACKGROUND:** Discusses the proliferation of internet-enabled devices and the challenge of disseminating information about newly-developing events.
 - SUMMARY:** Summarizes the techniques and framework for gathering information from live data streams and providing it to interested individuals.
- Claims (16):** Lists three claims, each describing a method implemented using one or more processors, comprising monitoring live data streams, generating a data structure, applying a machine learning model, and determining the likelihood of an event.

Figura 2.6: Captura de tela da ferramenta Google (sem data) obtida no dia 23/11/2024. Visualização da patente de Carbone e Feuz (2022).

2.3.3 Apresentação das patentes da Google sobre técnicas de modulação de atividade utilizando aprendizado de máquina

Tabela 2.1: Lista de patentes analisadas

Número	Título
11366812	Using live data streams and/or search queries to determine information about developing events
11252114	Message suggestions
11222637	Performing subtask(s) for a predicted action in response to a separate user interaction with an automated assistant prior to performance of the predicted action
11115695	Using machine learning and other models to determine a user preference to cancel a stream or download
11108889	Automatically resolving, with reduced user inputs, a set of activity instances for a group of users
10970096	Suggesting actions based on machine learning
9817869	Systems and methods for prioritizing notifications on mobile devices
9224100	Method and apparatus using accelerometer data to serve better ads
8645390	Reordering search query results in accordance with search context specific predicted performance functions

Para além das três operações descritas com mais detalhe ao longo do trabalho, constatamos algumas outras regularidades nas patentes (tabela 2.1). Em primeiro lugar, que a justificativa da propostas tecnológicas quase sempre recaem em um argumento de maximização de desempenho: seja em termos de facilidade para a usuária, economia de recursos ou maior velocidade ou precisão. Se, por um lado, de um ponto de vista estritamente técnico estes são parâmetros importantes, frequentemente utilizados para balizar a prática de engenharia; por outro lado, em conexão com o sistema de valores atuais, em especial o neoliberalismo como racionalidade pautada na generalização da concorrência e da empresa como norma institucional e subjetiva (Dardot e Laval 2016), o argumento reproduz, naturaliza e constrói uma realidade neoliberal. Isso é o que apresentamos sucintamente em outro lugar, com o nome de “viés de eficiência” (Gonçalves e Ferreira 2022b). O outro aspecto observado na maioria das patentes foi o funcionamento *online*²⁷, isto é, o treino do modelo de AdM pode acontecer concomitantemente ao seu uso para inferência/incidência. Essa constatação é interessante, pois mostra o papel ativo da usuária no treino do modelo, que ocorre também durante seu uso, e não apenas em servidores da empresa. Ao nosso ver, isto é um caso muito explícito do que Deleuze e Guattari (2011–2012) chamaram de “servidão maquínica” (ver capítulo 5). Tendo exposto esses apontamentos iniciais, passemos brevemente para a exposição de cada uma das patentes que compõe o nosso material.

²⁷“O termo online é geralmente reservado para o caso em que os exemplos [de treinamento] são extraídos de um fluxo de exemplos criados continuamente, em vez de um conjunto de treinamento de tamanho fixo sobre o qual várias passagens são feitas.” (Goodfellow et al. 2016:278). No nosso caso, também está implícito que o treinamento ocorre “em produção”, ou seja, durante o uso do algoritmo pela usuária.

US11366812 – Using live data streams and/or search queries to determine information about developing events

Na patente “Usando fluxos [*streams*] de dados e/ou consultas de pesquisa [*search queries*] para determinar informação sobre eventos em desenvolvimento” (Carbune e Feuz 2022), são apresentadas técnicas para agregar dados de múltiplos fluxos de dados (provenientes de smartphones, computadores pessoais e instituições) a fim de determinar informações sobre um evento em desenvolvimento (um exemplo dado é o caso de um incêndio em uma cidade). Além disso, a patente prevê a possibilidade de prover novas informações a indivíduos e instituições sobre novas informações “aprendidas” por meio do processamento desses dados via aprendizado de máquina. Essa informação pode ser a probabilidade da ocorrência de um evento ou uma medida da relevância deste possível evento para uma usuária específica.

US11252114 – Message suggestions

Na patente “Sugestão de mensagens” (Braun e Patel 2022), duas principais aplicações são providenciadas: em primeiro lugar, técnicas de processamento de linguagem natural, incluindo algoritmos de aprendizado de máquina, são usadas para criar uma representação das mensagens enviadas por uma usuária particular. Isso é feito por meio da extração de frases de mensagens enviadas, determinação de quais conversas incluem tal frase, geração de vetores numéricos para a representação das palavras e geração de bases de treino a partir dessas representações. Em segundo lugar, o método visa aplicar aprendizado de máquina sobre essas bases de dados para treinamento e posterior sugestão de mensagens de resposta para novas mensagens recebidas.

US11222637 – Performing subtask(s) for a predicted action in response to a separate user interaction with an automated assistant prior to performance of the predicted action

Em “Executando subtarefa(s) para uma ação prevista em resposta a uma interação separada da usuária com um assistente automático antes da execução da ação prevista”, Mirelmann et al. (2022) apresentam estratégias para armazenar dados de cache que facilitarão a execução de ações a partir de interações previstas entre usuária e assistente. Essa previsão é feita baseada no histórico de interações anteriores, bem como detalhes sobre o estado atual da interação. Desse modo, interações previstas podem ser inicializadas antes do comando da usuária. Justamente a antecipação da interação é feita utilizando aprendizado de máquina. Como consequência, este método diminui latência na resolução de tarefas que conformam com as predições do sistema.

US11115695 – Using machine learning and other models to determine a user preference to cancel a stream or download

Na patente “Usando aprendizado de máquina e outros modelos para determinar a preferência de uma usuária a cancelar uma transmissão [*stream*] ou download” (Lewis e Price 2021), é apresentado um método para treinar modelos de aprendizado de máquina utilizando informações sobre downloads e transmissões (de vídeo, por exemplo). O método visa resolver, de modo menos intrusivo à usuária, o problema de excesso de transmissões ou downloads simultâneos. Desse modo, informações contextuais e o histórico da usuária são utilizados para prever qual transmissão ou download a usuária preferiria

cancelar. O método ainda prevê seu uso simultaneamente a seu treino, isto é, a produção de mais dados sobre a preferência da usuária mesmo durante a sua execução.

US11108889 - Automatically resolving, with reduced user inputs, a set of activity instances for a group of users

Em “Resolvendo automaticamente, com entradas de usuárias reduzidas, um conjunto de instâncias de atividades para um grupo de usuárias”, Lance e Koonce (2021) apresentam técnicas para sugerir atividades, locais e horários para a realização de atividades por um grupo de usuárias a partir do processamento de mensagens de texto e de informações contextuais. Isso pode ser feito por meio da apresentação automática da sugestão na tela das usuárias, requerendo mais ou menos informações adicionais que podem ser manualmente concedidas, além do monitoramento de informações contextuais do dispositivo para, por exemplo, validar se uma usuária realizou a atividade sugerida. Desse modo, o aprendizado de máquina é treinado com os dados extraídos para prever a probabilidade de realização ou satisfação das usuárias em relação à atividade sugerida.

US10970096 - Suggesting actions based on machine learning

Na patente “Sugestão de ações baseada em aprendizado de máquina” (Krishna et al. 2021), uma técnica é apresentada para sugerir diferentes tipos de ações para uma usuária de um dispositivo (como um smartphone, por exemplo). Para isso tanto o conteúdo da tela, quanto dados do contexto da usuária podem ser analisados via aprendizado de máquina para que uma ação seja sugerida ou executada mais facilmente, rapidamente ou utilizando menos recursos computacionais. A sugestão é feita por meio de uma sobreposição na interface visual (chamada de “fatia [slice]”) que informa a usuária automática e instantaneamente.

US9817869 - Systems and methods for prioritizing notifications on mobile devices

Em “Sistemas e métodos para priorização de notificações em dispositivos móveis” (Aradhye et al. 2017), a proposta é a substituição de métodos tradicionais de ordenamento de lista de notificações (notadamente, a ordem cronológica), por uma ordem que visa representar a prioridade computada para tal ou qual notificação. Desse modo, o aprendizado de máquina sobre o histórico de ações da usuária faz com que notificações com maior probabilidade de reação pela usuária sejam graficamente enfatizadas ou apresentadas primeiro, enquanto que outras sejam rebaixadas ou mesmo seletivamente ignoradas.

US9224100 - Method and apparatus using accelerometer data to serve better ads

Em “Método e aparato utilizando dados do acelerômetro para providenciar anúncios melhores” (Chatterjee, Gawley, e Finne 2015), os dados do sensor acelerômetro de um dispositivo móvel, sensor multidimensional que quantifica aceleração, é utilizado, em conjunto com outros dados, para tipificar o movimento da usuária (por exemplo: “em pé”, “sentado” ou “deitado”) e informar a escolha de um conteúdo de anúncio a ser exibido para a usuária (por exemplo, na inicialização de um aplicativo ou entre fases de um jogo). O aprendizado de máquina é utilizado na etapa de ordenamento e seleção do conteúdo com base nos dados, entre outros, do acelerômetro, o que inclui não só o conteúdo em si,

mas o seu tipo (e.g. “grande comércio” ou “pequeno comércio”), formato (e.g. “texto” ou “áudio”) e uma previsão da ação da usuária (e.g. “visualização”, “clique”, “compra”, etc.).

US8645390 - Reordering search query results in accordance with search context specific predicted performance functions

Finalmente, a patente “Reordenando resultados de consulta de pesquisa de acordo com funções de desempenho previstas para o contexto de busca específico” (Oztekin et al. 2014) provê um método para o processamento de consultas de busca por aprendizado de máquina para identificar correlações entre o histórico de busca e ações e outras informações contextuais e resultados de mensagens. O aprendizado de máquina calcularia uma função de desempenho previsto que atuaria reordenando o resultado de uma busca visando privilegiar aqueles resultados com maior pontuação em algum critério de qualidade de predição.

Parte II

Uma cartografia do aprendizado de máquina

Capítulo 3

Captura

O roubo dos bens da Igreja, a alienação fraudulenta dos domínios estatais, o furto da propriedade comunal, a transformação usurpatória, realizada com inescrupuloso terrorismo, da propriedade feudal e clânica em propriedade privada moderna, foram outros tantos métodos idílicos da acumulação primitiva. Tais métodos conquistaram o campo para a agricultura capitalista, incorporaram o solo ao capital e criaram para a indústria urbana a oferta necessária de um proletariado inteiramente livre.

O Capital

Karl Marx

Agora se torna possível investir sobre toda criação, inclusive a criação da vida. Sabemos que por meio da privatização das telecomunicações, da colonização das redes e do próximo loteamento do campo eletromagnético, o capital global busca controlar o acesso e a exploração do ciberespaço; mas nos esquecemos de que a ambição maior da nova economia é assenhorear-se da dimensão virtual da realidade, e não apenas da dimensão da realidade virtual, do ciberespaço, como tem sido observado. Se tivermos em mente que a dimensão virtual da realidade começa a ser mais importante em termos econômicos do que a sua dimensão atual, teremos uma idéia melhor do sentido da corrida tecnológica. Aliado à tecnociência, o capitalismo tem a ambição de apropriar-se do futuro.

A informação após a virada cibernética

Laymert Garcia dos Santos

Como apontado por Marta Kanashiro (2023) o lançamento do ChatGPT no fim de 2022 pela OpenAI gerou um “evento de percepção” em relação à IA, causando muito mais uma admiração e reafirmação do desenvolvimento tecnológico do que reflexões críticas. Entretanto, a relativa e aparente qualidade das respostas, o que impressiona, também gera dúvidas quanto às origens dessa suposta capacidade discursiva. O ChatGPT é um chatbot baseado em algoritmos de modelagem de linguagem (MdL) – os GPT-3.5 e GPT-4 –, dos quais o BERT/Google, já citado, é outro exemplo¹.

Modelos de linguagem são algoritmos que visam mapear a distribuição de probabilidade de uma dada língua. Assim, a tarefa que visam resolver durante o treinamento é a de gerar complemento textual

¹Uma versão bastante inicial da ideia apresentada neste capítulo foi publicada na terceira edição da Revista Pimentalab (Gonçalves 2024a).

com uma alta probabilidade de ser reproduzido em uma língua específica. Em outras palavras, sua tarefa é gerar textos verossímeis. Com isso, embora não exista garantia de certeza sobre as informações mobilizadas em suas respostas, o chat pode ser utilizado para uma grande gama de tarefas envolvendo linguagem natural, o que é feito na forma de respostas a perguntas de uma usuária.

Mas essa capacidade de geração de texto com um grau relativo de veracidade é o resultado do processamento de uma quantidade enorme de textos captados e agrupados em bases de dados. Além disso, sua versão atual, gratuitamente disponibilizada na internet, também pode ser entendida como uma versão teste. Assim, as interações com usuárias na plataforma online visam refinar ainda mais a qualidade do potencial de inferência de texto que foi inicialmente gerado na interação com a base de dados:

É na nossa interação com o ChatGPT que fornecemos novos volumes de dados e que novos treinamentos dessa IA são feitos ou corrigidos. A interação humana é parte fundamental do desenvolvimento. [...] Enquanto na relação que o usuário estabelece parece reinar a constatação óbvia de que a máquina não é humana, no modo de funcionamento da máquina o humano já é parte do seu circuito integrado de desenvolvimento e de funcionamento (Kanashiro 2023)

Desse modo, a capacidade supostamente lógica e criativa do chatGPT é consequência da apropriação ou do controle de processos informacionais que acontecem tanto antes do funcionamento da máquina – agregados na base de dados – quanto durante a própria interação com ela, na medida em que estes também serão armazenados e utilizados no funcionamento futuro do modelo. É esse aspecto que faz com que pesquisadoras questionem a natureza gerativa de tais modelos, propondo denominá-los alternativamente como “derivativos”, por exemplo (Silva 2022). Mas o ChatGPT e os MdLs são apenas um exemplo dessa apropriação privada e unidirecional de conhecimento que embasa o potencial de tecnologias de IA e que nos remete às origens do capitalismo, tal qual apontado na epígrafe de Marx.

Em 2021, a plataforma GitHub anunciou a versão beta do GitHub Copilot, assistente de programação baseado em IA que ficaria disponível para o ambiente de desenvolvimento Visual Studio Code (VS Code). O GitHub é o principal site hospedador de software livre e de código aberto, mas, desde sua aquisição pela Microsoft em 2018, expande para a criação de outros produtos – sobretudo ligados à sua integração com o VS Code, também da Microsoft. A compra do GitHub pela Microsoft gerou tensões na comunidade de software livre, que, em mais de uma ocasião, organizaram migrações em massa para a plataforma concorrente, – o GitLab –, entretanto, isso não foi o suficiente para mudar a posição central do GitHub no desenvolvimento de softwares livre e de código aberto.

Dessa forma, o lançamento de um produto que promete ser um “par programador baseado em IA” e te permitir “focar no que importa” ao automatizar processos como criação de estruturas repetitivas ou de estruturas simples, mas recorrentes, teve um grande impacto na comunidade de desenvolvedoras. Nas palavras do site do produto: “GitHub Copilot é uma IA par programadora que ajuda você a escrever código mais rápido e com menos trabalho. Ela extrai [*draws*] contexto de comentários e códigos para sugerir linhas individuais e funções completas instantaneamente”² (GitHub 2024).

²A característica mais notável do produto é a criação automática de código funcional a partir da descrição de uma tarefa ou mesmo apenas do nome da função e das entradas que se espera receber. Um exemplo retirado do próprio site (GitHub 2024), para a linguagem de programação python, é o seguinte:

Mas de onde vêm essa capacidade de criação de código automatizada? Segundo o próprio site do produto: “GitHub Copilot é baseado [*powered*] no Codex, um modelo de IA gerativo pré-treinado criado pela OpenAI. Ele foi treinado em textos de linguagem natural e em código fonte publicamente disponíveis, incluindo código em repositórios públicos do GitHub” (GitHub 2024). Isto é, a capacidade de gerar código automaticamente vem do processamento em massa e não explicitamente consensual de códigos já criados por programadoras em diversos contextos e que estão disponibilizados publicamente na internet.

Algo similar acontece com os novos softwares geradores de imagens baseados em IA – como o Dall-E, o Craiyon (previamente chamado “Dall-E mini”) e produtos baseados na técnica stable diffusion – que, recentemente, motivaram uma onda de protestos por artistas da plataforma ArtStation (Faustino 2023). Nas palavras de Alexander Nanitchkov (2022), que precedeu o protesto coletivo: “ ‘Arte’ de IA contemporânea é criada nas costas de centenas de milhares de artistas e fotógrafos que fizeram bilhões de imagens e investiram tempo, amor e dedicação para ter seu trabalho desalmadamente [*soullessly*] roubado e usado por pessoas egoístas para lucro sem o menor conceito de ética”.

Como no caso anterior, a capacidade criativa desses algoritmos reside no processamento de grandes quantidades de imagens que são, frequentemente, extraídas de repositórios de conteúdos na internet. Desse modo, percebemos que há o avanço da tecnociência sobre novos “territórios digitais”, o que culmina em novos processos de “cercamento”, em que conteúdos muitas vezes livres são diretamente relacionados entre si e apropriados monopolisticamente por um modelo privado de IA.

```
1 import datetime
2
3 def parse_expenses(expenses_string):
4     """Parse the list of expenses and return the list of triples (date, value, currency).
5     Ignore lines starting with #.
6     Parse the date using datetime.
7     Example expenses_string:
8         2016-01-02 -34.01 USD
9         2016-01-03 2.59 DKK
10        2016-01-03 -2.72 EUR
11    """
12    expenses = []
13    for line in expenses_string.splitlines():
14        if line.startswith("#"):
15            continue
16        date, value, currency = line.split(" ")
17        expenses.append((datetime.datetime.strptime(date, "%Y-%m-%d"),
18                        float(value),
19                        currency))
20    return expenses
```

Nesse exemplo, importa-se a biblioteca `datetime` (conjunto de funções já existentes), escreve-se o cabeçalho da função (o título `parse_expenses` e a entrada `expenses_string`) e descreve-se brevemente em um comentário o que se pretende com ela (linhas 4 a 10): “Processe uma lista de gastos e retorne a lista de triplas (data, valor, moeda)./ Ignore linhas começando com #./ Processe a data utilizando `datetime`./ Exemplo de `expenses_string`:/ [...]”. O resultado são linhas de código (10-20) totalmente funcionais que realizam as operações descritas dentro dos limites estabelecidos.

3.1 Vigilância e captura

Vigilância e captura são dois termos utilizados para descrever a situação contemporânea em que grande volume de dados é gerado e extraído por corporações. Enquanto que o primeiro nos remete a uma certa passividade, a captação de rastros deixados online e uma forte aproximação com metáforas visuais como o panóptico e o Big Brother orwelliano; o segundo tende a nos remeter a um processo ativo e violento de tomada à força³.

3.1.1 Panopticismo e pós-panopticismo

Há uma extensa literatura que relaciona a vigilância com os sistemas digitais. Um dos principais nomes a conceituar a vigilância, foi Michel Foucault (2014), relacionando-a com a disciplina e culminando na imagem do panóptico como representativa das práticas de vigilância na escola, na fábrica e na prisão⁴. Entretanto, como constata Deleuze (2013) – tanto quanto Haraway (2023a) ou Garcia dos Santos (2003a) –, os fins do século XX já apontavam para uma direção distinta daquela caracterizada por Foucault em seus estudos sobre disciplina, não no sentido de sua negação, mas de complexificação e ampliação para outros modos (ver capítulo 5). Há também outras autoras que divergem da interpretação foucaultiana, preferindo não só outros nomes para denotar a era contemporânea, como também outras caracterizações. Nesse sentido, é notável a proliferação de propostas de arquiteturas “pós-panóptico”, como tentativas de expandir ou substituir o paradigma da vigilância e do Big Brother, por outras que melhor descreveriam a situação atual: “panóptico reverso”, “superPanopticon”, “synopticon”, “neo-Panopticon”, “oligopticon”, “Ubiopticon” e “Polyopticon” são algumas das propostas apresentadas por Sherman (2022). A crítica presente nessas alternativas geralmente recai no caráter centralizado da arquitetura panóptica, sendo a vigilância observada hoje em dia mais distribuída, seja por meio de uma racionalidade que internaliza a vigilância em cada sujeito (Han 2023), por uma difusão de agentes vigilantes (instituições e corporações, por exemplo) que não se reduzem ao Estado e competem com visões parciais sobre os mesmos vigiados (Latour 2012), ou pela própria natureza de técnicas como a internet das coisas (IoT, *internet of things*) e a computação ubíqua (Sherman 2022).

Um caso interessante é o de Shoshana Zuboff (2018) que, embora utilize o termo “capitalismo de vigilância”, avança no sentido de suas críticas, propondo chamar de Big Other a arquitetura-símbolo da nossa era: “Essa nova arquitetura configura-se como um ubíquo regime institucional em rede que registra, modifica e mercantiliza a experiência cotidiana, desde o uso de um eletrodoméstico até seus próprios corpos, da comunicação ao pensamento, tudo com vista a estabelecer novos caminhos para a monetização e o lucro” (Zuboff 2018:43–44).

³Gostaria apenas de mencionar que “não é só na tecnociência moderna que a captura possui uma posição de relevo”, como no caso dos sistemas digitais de vigilância apresentados aqui. Eduardo Viveiros de Castro (2018) apresenta a captura, em sua vertente canibal, como momento importante da metafísica ameríndia: “Baseado na ‘semiofagia guerreira dos Tupi’ e em estudos que aproximam a ‘caça de cabeças dos Jívaro’ a um ‘dispositivo de captura das virtualidades de pessoas’, entre outros, [Viveiros de Castro] conceitualiza uma ‘economia da alteridade predatória’ nos ameríndios: deste ponto de vista, conclui que ‘a “interioridade” do corpo social [ameríndio] é integralmente constituído pela captura de recursos simbólicos - nomes e almas, pessoas e troféus, palavras e memórias - do exterior’ (Viveiros de Castro, 2018, p. 161). Entretanto, o que diferencia o paradigma ameríndio do moderno, segundo o próprio antropólogo, é a ênfase dada para a inteligibilidade que está ou na objetividade, no caso deste, ou na subjetividade, no caso daquele” (Gonçalves e Padovan 2024a:15).

⁴David Lyon também é um nome importante da área, pois foi um dos fundadores do campo conhecido como “estudos de vigilância” e um dos que a relacionou com as tecnologias digitais contemporâneas (e.g. Lyon 2018).

Para a autora, o capitalismo de vigilância seria uma lógica econômica parasita, baseada na extração de dados visando predição e vendas. Ela se concretizaria por meio de uma arquitetura global de modificação de comportamento que sustentaria uma lógica econômica baseada na vigilância para propiciar a concentração de riqueza (Zuboff 2019). Além disso, seria a instauração de um novo modo de poder, o instrumentarianismo, que, visando uma ordem social baseada na certeza total, promoveria uma divisão social do aprendizado, o controle dos meios de modificação de comportamento e desposseção do excedente comportamental para computação, controle e predição (Zuboff 2019). Ou seja, diz respeito a uma vigilância ativa (extração de dados, predição e modificação de comportamento), que de certa forma se aproxima de uma noção de captura, isto é, da tomada à força.

Seguindo os passos sugeridos em Deleuze (2013), consideramos que não há hoje uma substituição da disciplina, da vigilância e da estrutura panóptica, mas sim uma mudança de foco: assim como a soberania continua a existir em uma sociedade baseada na disciplina, esta continua a existir em uma centrada no controle. Assim, não acreditamos que haja uma transição completa para fora do modelo panóptico, mas uma nova camada que se assenta sobre esta na medida em que as tecnologias digitais apresentam novas possibilidades de coleta distribuída de dados, comunicação instantânea e controle direto sobre processos informacionais.

3.1.2 Colonialismo digital, colonialismo de dados e novos cercamentos

É nesse sentido que algumas autoras aproximam o momento atual com as origens do capitalismo e da colonização, falando de uma “acumulação primitiva de dados”. Ao menos duas características do momento atual parecem apontar para tal interpretação, aproximando a situação contemporânea com aquela descrita por Marx (2017, especialmente no capítulo 24: ‘A assim chamada acumulação primitiva’): o caráter violento da captura de informação pelas *big techs*; e o vazio legal que faz com que muitas práticas de extração de dados estejam desregulamentadas, enquanto que outras são criadas com intervenção ativa das próprias agentes que se beneficiam da extração de dados⁵. Entretanto, o modo específico como a relação tecnologias-digitais-colonização é pensada diverge de autora para autora. Para Couldry e Mejias (2019; 2020), o colonialismo de dados pode ser visto como uma forma específica de colonialismo própria ao século XXI que une formas históricas de extrativismo colonial com novos métodos de quantificação abstrata. Desse modo, seria uma nova etapa de acumulação primitiva que “vai prover as pré-condições para um novo estágio do capitalismo que ainda não conseguimos imaginar, mas cuja apropriação da vida humana através dos dados será central” (Couldry e Mejias 2019:337). Localizando o centro de poder na díade EUA-China, seu enfoque recai sobretudo na extração de *dados humanos*, embora a isso os autores denominem hiperbolicamente de “vida humana”:

Através do que chamamos de “relações de dados” (novos tipos de relações humanas que habilitam a extração de dados para comodificação), a vida social ao redor do globo se torna um recurso “aberto” para extração que está de certo modo “aí” para o capital. Esses fluxos globais de dados são tão expansivos quanto a apropriação de terra, recursos e corpos no colonialismo histórico, embora o epicentro tenha se deslocado. O colonialismo de dados envolve não apenas um polo de poder colonial (“o Oeste”), mas ao menos dois: os Estados Unidos e a China. Isso complica nossa noção geográfica de Sul Global, um conceito que

⁵Agradecemos a Stefano Schiavetto Amancio por primeiro nos apontar essa similaridade em reunião do LaSPA em 2023.

até agora ajudou a situar a resistência e a desidentificação ao longo de divisões geográficas anteriores entre colonizadores e colonizados. Em vez disso, o novo colonialismo de dados trabalha ao mesmo tempo externamente – em uma escala global – e internamente sobre suas próprias populações. A elite do colonialismo de dados (pense na Facebook[, ou na Google]) se beneficia da colonização em ambas as dimensões e divisões Norte-Sul, Leste-Oeste não mais importam da mesma forma (Couldry e Mejias 2019:337).

Tanto a ideia de identificar o poderio da China e dos EUA, quanto a de falar em um “novo capitalismo” são tão ousadas quanto controversas. Além disso, não entendemos que Norte e Sul sejam categorias geográficas, mas antes políticas, entendendo que há um norte do sul, tanto quanto um sul do norte, por exemplo, que explicita escalas múltiplas de dinâmicas que constituem, de forma localizada espaço e temporalmente, relações centro-periferia. Entretanto, acreditamos ser valioso o esforço dos autores em retratar o momento atual como um de expropriação ou desposseção, algo que também está presente em outras autoras como Thatcher, O’Sullivan, e Mahmoudi (2016) e a própria Zuboff (2019). Paola Ricaurte (2019) também segue uma linha convergente, estendendo sua análise sobre acumulação primitiva de dados, tanto quanto da colonialidade do poder⁶, para além das relações humanas e visando incluir as tecnologias digitais: “o regime de colonialismo de dados inclui a captura de relações de dados”, tal qual apresentados por Couldry e Mejias (2019), não só nas relações entre pessoas que, através da mediação algorítmica, produzem dados capturáveis, “mas também como todo um universo de interações humano-objeto e objeto-objeto que emergiram com o desenvolvimento da ‘Internet das Coisas’ (IoT), tanto quanto biodados [*biodata*] e dados de ‘atividades derivadas de relações não relacionadas com o humano, como energia, água, estradas, redes de infraestruturas e recursos naturais’ ”⁷ (Ricaurte 2019:2).

Outras autoras preferem ainda o termo “colonialismo digital”. Para Deivison Faustino e Walter Lippold (Faustino e Lippold 2023; Lippold e Faustino 2022), o colonialismo digital diz respeito a um quadro teórico que é, em relação à proposta de Couldry e Mejias (2019), mais amplo em escopo, e mais prudente em relação às suas consequências de reformulação do capitalismo. Além disso, os autores se preocupam com uma caracterização não metafórica deste novo “colonialismo”, explicitando as continuidades da dominação colonial histórica e nas tecnologias digitais⁸. Segundo os autores, o colonialismo digital se define por duas tendências: uma delas, o “colonialismo de dados”, “é aquela que subsume cada vez mais a vida humana, o ócio, a criatividade, a cognição e os processos produtivos às lógicas extrativistas, automatizadas e panópticas do colonialismo digital”, promovendo a “manipulação intencional da cognição humana por grandes corporações empresariais a partir dessas tecnologias com vistas à ampliação da acumulação de capitais” (Lippold e Faustino 2022:58). Neste sentido, a concepção bastante explorada por Couldry, Mejias (2019; 2020) e Ricaurte (2019), diz respeito a uma escala micropolítica ou molecular, no sentido que lhe dá ao termo Deleuze e Guattari (2011–2012), apenas uma das facetas do colonialismo digital. A outra, macropolítica ou molar, é uma “nova partilha territorial do globo terres-

⁶Para uma exposição do termo “colonialidade do poder”, central no pensamento do sociólogo Aníbal Quijano do grupo Modernidade/Colonialidade, conferir Quijano (2009).

⁷A citação é de Cordova, Porciuncula, e Brebant (2018).

⁸Em relação a uma crescente literatura sobre colonialismo digital e similares que não trata da problemática racial, os autores se questionam: “O silêncio da literatura especializada em digital colonialism, i-colonialism ou data colonialism, sobre o racismo no universo digital é ensurdecador. Se o racismo foi e continua sendo a base para as velhas e novas formas de colonialismo, nos perguntamos, como nos foi possível o advento de toda uma literatura sobre COLONIALISMO (digital) que não discute o racismo?” (Lippold e Faustino 2022:69).

tre entre os grandes monopólios da indústria da informação: as chamadas Big Techs, majoritariamente concentradas no Vale do Silício, mas não apenas”, que “atualiza o imperialismo, o subimperialismo e neocolonialismo tardio ao reduzir o chamado Sul global a mero território de mineração extrativista de dados informacionais” (Lippold e Faustino 2022:58).

Do mesmo modo que a construção de linhas férreas globalmente visava uma “partilha colonial do mundo”, as fibras ópticas e cabos de rede atualizam-na para a realidade contemporânea (Lippold e Faustino 2022:64–65)⁹. Os autores nos advertem como esse “novo extrativismo”¹⁰ amplifica as assimetrias relativas à divisão internacional do trabalho. No contexto de um campo que podemos chamar, com Fuchs, de trabalho digital¹¹, o corte norte/sul se expressa entre quem detém infraestrutura técnica e financiamento em pesquisa e desenvolvimento para criação de novos algoritmos, por um lado, e o trabalho precarizado de mineração de recursos, montagem de hardware e de tarefas cognitivas repetitivas em plataformas digitais como a Amazon Mechanical Turk (MTurk), de outro. Em relação à extração de recursos naturais, podemos dizer que a infraestrutura material dos sistemas digitais é largamente baseada no trabalho precarizado de mineradoras – sobretudo no sul global. Por exemplo, entre 50% e 70% de todo o lítio – metal extensamente utilizado na produção de baterias – está localizado na Bolívia (Crawford e Joler 2018). Além disso, Crawford e Joler (2018) mostram como a IA também implica altos custos energéticos e a geração de altas quantidades de lixo eletrônico que são desigualmente distribuídos ao redor do mundo. Especificamente sobre o aprendizado de máquina, Pasquinelli e Joler (2021:1266) apontam o papel do trabalho plataformizado na criação das bases de dados como a MNIST, a ImageNet e a Labelled Faces in the Wild utilizadas por polos de conhecimento no norte global: “olhando com mais cuidado, descobrimos uma profunda divisão de trabalho que penetra [*innervates*] no sul global via plataformas de crowdsourcing que são usadas para editar e validar dados”.

Já para Kwet (2019), que descreve essa escala molar enfocando a materialidade das infraestruturas técnicas em sua conceitualização de colonialismo digital, são três os principais modos de concretização das assimetrias de poder:

Um insidioso novo fenômeno, o colonialismo digital, lança uma sombra sobre o Sul Global. Essa forma estrutural de dominação é exercida por meio da propriedade e do controle centralizados dos três pilares principais do ecossistema digital: software, hardware e conectividade de rede, o que confere aos Estados Unidos um imenso poder político, econômico e social. Dessa forma, o GAFAM (Google/Alphabet, Amazon, Facebook, Apple e Microsoft) e

⁹Não é um fato acidental que “os cabos de Internet traçam mesmas rotas dos navios negreiros”, tal qual apresentado por Solagna (2024) que comenta o artigo de Mwema e Birhane (2024) que analisa dois cabos submarinos de internet conectando África e América (Equiano, da Google; e 2Africa, da Facebook/Meta) concluindo que tais projetos “não só reproduzem a lógica colonial, como também seguem o mesmo caminho infraestrutural traçado durante a época do tráfico transatlântico de escravos”.

¹⁰Em relação à aproximação IA-extrativismo, conferir o trabalho de Crawford (2021) e a interessante cartografia audiovisual elaborada por Joler (2020).

¹¹“Como explica Fuchs, o ‘trabalho digital’ está longe de ser efêmero ou virtual, mas está profundamente incorporado em diferentes atividades. O escopo é impressionante: da servidão por contrato [*indentured labor*] em minas para extrair os minerais que formam a base física das tecnologias da informação; ao trabalho de processos de fabricação e montagem de hardware estritamente controlados e, às vezes, perigosos em fábricas chinesas; à exploração de trabalhadoras cognitivas terceirizadas em países em desenvolvimento que rotulam conjuntos de dados de treinamento de IA; às trabalhadoras físicas informais que limpam depósitos de lixo tóxico. *Esses processos criam novos acúmulos de riqueza e poder, que se concentram em uma camada social muito fina*” (Crawford e Joler 2018:7).

outros gigantes corporativos, bem como agências de inteligência estatais como a National Security Agency (NSA), são os novos imperialistas da comunidade internacional. A assimilação aos produtos, modelos e ideologias tecnológicos de potências estrangeiras - lideradas pelos Estados Unidos - constitui uma forma de colonização do século XXI (Kwet 2019:2).

Exemplos concretos apontados pelo autor são o uso de software proprietário (ao invés de software livre), o que sustenta, por exemplo, a garantia de mercantilização de conteúdos culturais utilizando Gerenciamento de Direitos Digitais (DRM, *Digital Rights Management*), – tecnologia que impede a cópia não autorizada de conteúdos digitais –; o uso de computação em nuvem, em servidores privados¹², por exemplo para acessar conteúdos digitais mercantilizados na forma de *streamings*; e, por fim, as violações na noção de “neutralidade de rede” por Provedores de Serviço Internet (ISPs, *Internet Service Providers*) ao redor do mundo, por exemplo no caso do Free Basics da Facebook/Meta, ou das políticas de *zero-rating* nas operadoras de telefonia brasileiras. Mas, além da detenção privada da principal parcela de software, hardware e protocolos de rede pelo complexo industrial-militar-corporativo estadunidense, o exemplo da África do Sul trazido por Kwet apresenta também como ainda há uma forte e contínua intrusão de países do norte sobre o sul, visando impor e intensificar ainda mais a dependência destes em relação a aqueles:

Atualmente, está ocorrendo uma nova forma de colonização corporativa. Em vez da conquista de terras, as grandes empresas de tecnologia estão colonizando a tecnologia digital. As funções a seguir são todas dominadas por um punhado de multinacionais norte-americanas: mecanismos de busca (Google [Search]); navegadores da Web (Google Chrome); sistemas operacionais de smartphones e tablets (Google Android, Apple iOS); sistemas operacionais de desktops e laptops (Microsoft Windows); software de escritório (Microsoft Office, Google G Suite); infraestrutura e serviços em nuvem (Amazon, Microsoft, Google, IBM); plataformas de redes sociais (Facebook, Twitter); transporte (Uber, Lyft); redes de negócios (Microsoft LinkedIn); streaming de vídeo (Google YouTube, Netflix, Hulu); e publicidade on-line (Google, Facebook) - entre outros. A GAFAM agora compreende as cinco corporações mais ricas do mundo, com uma capitalização de mercado combinada superior a US\$ 3 trilhões. Se os sul-africanos integrarem os produtos das Big Techs em sua sociedade, os Estados Unidos obterão enorme poder sobre sua economia e criarão dependências tecnológicas que levarão à extração perpétua de recursos (2019:5).

Dessa forma, com Faustino e Lippold, entendemos que há duas dimensões do chamado “colonialismo digital”, uma molar e outra molecular, que se articulam em um exercício de poder largamente baseado em práticas de vigilância ou de captura. Além disso, entendemos, ainda seguindo o argumento dos autores, que ambas as dimensões são marcadas por assimetrias baseadas em raça/cor, o que os autores denominam “racialização codificada”. Portanto, não se trata aqui de um uso metafórico de colonização, ou de um uso do termo desvinculado de questões de raça, mas pelo contrário, a justificação do uso do termo colonialismo está justamente na continuidade da operação de classificação racial que hierarquiza sujeitos com base em suas características físicas ainda hoje e agora por meio de tecnologias

¹²Interessante notar como o próprio termo *servidor* tem uma forte conotação colonial. Confrontar com o uso do termo mucúia (nome dado ao fruto do baobá) pela Rede Mocambos, que desenvolve o software Baobáxia: <https://www.mocambos.net/tambor/pt/baobaxia> (acesso: 01/07/2024).

digitais (Faustino e Lippold 2023:22, 96, *passim*). No nível molar, o trabalho precarizado de mineradoras, montadoras de hardware e trabalhadoras de plataforma englobam assimetrias de classe, gênero e raça de forma mais evidente. Entretanto, entendemos que ele também está presente na esfera molecular. Pasquinelli e Joler (2021) mostram como as formas históricas de opressão – como é o caso do racismo – são naturalizadas, reproduzidas nos dados e amplificadas pelos modelos algorítmicos (o que os autores chamam, respectivamente, de “viés histórico”, “viés de base de dados” e “viés algorítmico”). Um exemplo notável é o trabalho “Gender Shades [tons de gênero]” de Joy Buolamwini e Timnit Gebru (2018:10) que analisou interseccionalmente softwares comerciais de detecção automática de gênero, concluindo que “todos os algoritmos [testados] possuem um pior desempenho em indivíduos do sexo feminino e de tom de pele mais escuros [*female and darker subjects*] quando comparados com os seus homólogos do sexo masculino e mais claros [*male and lighter subjects*]”. Tarcízio Silva (2022) é uma referência importante neste aspecto, pois apresenta uma miríade de outros casos do que o autor denomina “racismo algorítmico”¹³.

3.1.3 Da vigilância à captura

No que diz respeito às patentes, fica claro que os dados são gerados não apenas por meio de uma vigilância passiva sobre rastros digitais, mas por um processo ativo de construção infraestrutural visando possibilidades de produção e captação de dados. A tabela 3.1 sintetiza os principais elementos materiais descritos nas patentes analisadas como meios para extração e geração de dados. Como é possível observar, há uma diversidade grande de sinais de informação a serem captados, sendo tempo e localização quase sempre registrados, além de ser muito comum o uso de sensores e a extração de texto. Ademais, outras possibilidades incluem sinais do microfone, câmera, histórico de ações (ou de buscas, ou navegação na web), bem como diversos sinais de informação relativos ao dispositivo.

Tabela 3.1: Elementos de captura encontrados nas patentes.

Elementos de captura	11366812	11252114	11222637	11115695	11108889	10970096	9817869	9244100	8645390
Tempo (relógios)	•	•	•	•	•	•	•		•
Localização (GPS/IP)	•	•	•	•	•	•	•	•	•
Conteúdo da tela		•							
Conteúdo de texto		•			•		•		•
Câmera	•	•				•			
Microfone	•		•			•			
Aceleômetro/giroscópio		•						•	
Orient. dispositivo							•	•	
Histórico de ações			•						•
Sensor proximidade		•					•		•
WiFi (internet)			•						
NFC				•					
RFID						•			
Bluetooth				•					
Status de mídia				•					

¹³Apontamos brevemente que tanto a captura de dados, quanto a inferência/incidência informacional estão perpassadas pela dimensão racial. Se o primeiro se evidencia na menor eficácia de tecnologias como reconhecimento facial (Silva 2023) e aferição de taxas de oxigênio no sangue – conferir <https://www.labirinto.labjor.unicamp.br/enciclopedia/oximetro/> (acesso: 01/07/2024) –; o segundo se expressa mais fortemente como racismo algorítmico nos resultados de tecnologias digitais (voltaremos a estas questões no capítulo 8).

Muitas patentes se baseiam na identificação, não somente da usuária, mas de um contexto ao seu redor. Por exemplo, Krishna et al. (2021) apresentam que o contexto de um dispositivo pode conter: *tempo* associado a uma atividade (acordar pela manhã, comutação pela manhã, etc.), *local* (casa, trabalho, aeroporto, ginásio), *situação* (no carro, com fone de ouvido conectado, com cabo de energia conectado, etc.) ou *atividade* (viajando, correndo, andando de bicicleta). Já Lewis e Price (2021), apresentam algumas possibilidades de captação de informação: da própria *usuária* (feedback de usuária, ações de navegação, ações em um reprodutor de mídia, histórico de visualização, inatividade, rótulos temporais, padrões de uso ou hábitos: como, por exemplo, uma usuária que ouve música no carro voltando do escritório, e ao chegar em casa alterna para a TV); do *contexto* (browser utilizado, aplicativo nativo, tempo do dia, local); do *conteúdo* da mídia (gênero, tamanho, título, artista, tipo – áudio ou vídeo –, e metadados); e do *dispositivo* (tipo – telefone móvel, tablet, TV, destop –, conexões de internet, etc.).

Nesse sentido, com Zuboff (2018), entendemos que há aí um mecanismo ativo que visa, por design, gerar e extrair dados das usuárias. Para a autora, assim como para Couldry e Mejias (2020), as relações sociais são material bruto apropriados pelo capitalismo (Zuboff 2019). Nossa análise é parcialmente convergente com tal resultado, pois mostra, sim, a capitalização a partir de dados extraídos de relações sociais; entretanto, nossa interpretação é mais restrita por considerar que o que é apropriado é uma dimensão particular da realidade ou das relações sociais, a saber, a informação (seção 3.2).

Para Zuboff (2018), são cinco os principais fluxos informacionais que passam a ser capturados no capitalismo de vigilância:

1. Fluxos institucionais e transinstitucionais mediados por computador (dos quais as transações econômicas são parte significativa).
2. Fluxos de sensores incorporados em objetos, corpos e lugares (IoT, IoE, etc.).
3. Fluxos de bases de dados governamentais e corporativas (incluindo bancos, serviços públicos, telecomunicações, etc.).
4. Fluxos de câmeras de vigilância públicas e privadas (desde smartphones até satélites).
5. Fluxos residuais produzidos acidentalmente por usuárias individuais (curtidas, visualizações, textos, erros ortográficos, etc.).

Nas patentes, o foco da captura recai sobre o último tipo de informação, embora os outros ainda estejam presentes em maior ou menor grau. É relevante aqui a característica constatada por Zuboff (2018:33) segundo a qual “o que importa é a quantidade e não a qualidade” dos dados. E, especificamente sobre a empresa objeto deste trabalho, que “a Google é ‘formalmente indiferente’ ao que os usuários dizem ou fazem, contanto que o digam e o façam de forma que o Google possa capturar e converter em dados”.

Zuboff chama ainda de texto eletrônico (2018) ou textos-sombra (2019) a informação digital capturada que circula nos meios digitais: “[n]esse novo regime, a arquitetura global da mediação por computador transforma o texto eletrônico, anteriormente delimitado dentro das organizações, em um organismo global inteligente que chamo de Big Other” (Zuboff 2018:58). Acreditamos que esses aspectos marcam, já na ideia de “capitalismo de vigilância”, um deslocamento da passividade representacional e da visualidade, para uma performatividade ativa e para uma textualidade.

Ainda nesse sentido, Agre (1994), nos fins do século XX, já advogava pela superioridade explicativa da captura em relação à vigilância, distinguindo-os em termos de dois paradigmas concorrentes.

Segundo o autor, “quando atividades humanas se tornam imbricadas com os mecanismos de rastreamento [*tracking*] computadorizados, a noção de interação humana com um ‘computador’ [...] começa a perder sua força. Em seu lugar encontramos sistemas de atividades que são fortemente integrados com processos computacionais distribuídos” (Agre 1994:105). Nesse momento, o modelo da vigilância – baseado em metáforas visuais (o panóptico e o Big Brother) e territoriais (como a “invasão”), e em uma lógica passiva (não-disruptiva), centralizada e identificada com o Estado – cede espaço ao modelo de captura.

Para Agre (1994), diferentemente de outros usos correntes do termo na computação¹⁴, o modelo de captura se baseia: 1) em metáforas linguísticas (há uma “gramática de ação”); 2) na suposição de que a codificação das atividades humanas envolvem uma intervenção e uma reordenação ativas dessas atividades; 3) em metáforas estruturais (atividade é codificada à partir de um “catálogo”, por exemplo); 4) de uma organização descentralizada e heterogênea, relacionando práticas locais e formações sociais mais amplas; e 5) com objetivos que não são filosóficos ou políticos¹⁵, mas sim resultantes da assimilação de um formalismo matemático virtual ou transcendente (Agre 1994:107). “O modelo de captura descreve a situação que resulta quando gramáticas de ação são impostas sobre atividades humanas e quando as novas atividades reorganizadas são representadas por computadores em tempo real” (Agre 1994:109).

Ao lado da construção de representações sistemáticas de atividades organizacionais e de sua automação, Agre identifica na IA – e, argumentaríamos, sobretudo no seu desenvolvimento como aprendizado de máquina – um momento importante da captura:

Pesquisadoras de IA, mais que seus pares [*counterparts*] em outras áreas de ciência da computação aplicada, começam a pesquisar explicitamente por sistemas ontológicos que permitiriam um computador a representar clara e precisamente [*cleanly and accurately*] uma ampla gama de conhecimento humano – incluindo conhecimento sobre atividades e organizações sociais (Agre 1994:108)

Entretanto, para o autor, a captura tem a ver não apenas com a representação (ainda que esta seja necessária), mas também e principalmente com a reorganização (Agre 1994:110). Esse processo, que poderia ainda ser subdividido em etapas de um ciclo¹⁶, está centrado na noção de gramáticas de ação. “Instituições frequentemente impõe gramáticas em atividades para razões outras que a captura em tempo-real – por exemplo, para segurança, eficiência, potencial para responsabilização [*liability*] e controle simples” (Agre 1994:110). Meios sociais e técnicos são providos com o potencial de uma complexa divisão de trabalho, “depois, os participantes começam, por necessidade, a orientar suas atividades em direção da maquinaria de captura e suas consequências institucionais” (Agre 1994:110). Esse

¹⁴“O discurso computacionalista raramente traz à superfície a conotação de violência na metáfora da captura; não se diz que a informação capturada foge, escapa ou se ressentida de seu aprisionamento”. A captura é utilizada num sentido epistemológico como aquisição de um dado como entrada ou como representação da realidade: “o segundo uso de captura, que é mais comum em pesquisas sobre inteligência artificial, se refere à habilidade de um esquema de representação a expressar totalmente, acuradamente ou ‘de forma limpa’, noções ou distinções semânticas, sem se referir à real tomada de dado” (Agre 1994:106).

¹⁵Acrescentaríamos um adendo, por considerar que os objetivos não são *explicitamente* ou *aparentemente* políticos ou filosóficos, dado que o formalismo matemático já supõe uma filosofia e uma política, concretizadas no seu processo histórico de construção e estabelecimento. Latour (2011, no capítulo 6) mostra como o formalismo matemático é o ponto máximo de pontualização da prática científica e, portanto, seu momento mais social, na medida em que incorpora a agência associada de mais atores.

¹⁶“Análise”, “articulação”, “imposição”, “instrumentação” e “elaboração” (Agre 1994:110).

registro das atividades são um produto da atividade reorganizada e passam a ser guardadas, inspecionadas, auditadas e analisadas estatisticamente. “A acumulação em tempo-real do dado da atividade, em outras palavras, introduz novos elementos para atividade mesma” e “as pessoas cujas atividades estão sendo capturadas provavelmente ajustarão sua conduta baseada no entendimento do que se tornará dado e suas consequências para a suas próprias vidas” (Agre 1994:111).

Desse modo, a captura é tanto uma operação técnica, quanto social. Ao metrificar, registrar e permitir uma identificação precisa da relação entre pessoa e atividade, a captura é particularmente relevante para uma reorganização social baseada na noção de eficiência econômica (Agre 1994:119). “Ao impor uma forma matemática precisa sobre atividades antes não formalizadas, a captura normatiza essas atividades e seus elementos componentes e portanto as prepara (...) para uma eventual transição para uma relação baseada no mercado” (Agre 1994:120). Assim, “o resultado é uma aceleração da atividade econômica cujos benefícios sociais em termos de eficiência produtiva são claros o suficiente, mas cujos custos sociais devem ser um motivo de preocupação” (Agre 1994:121–22).

Entendemos que – criando um diálogo entre Zuboff e Agre, cujos trabalhos possuem elementos complementares para a caracterização que fazemos -, no caso do AdM, podemos descrever os fluxos informacionais tanto do ponto de vista imagético, quanto textual. Primeiramente, os dados são imagens na medida em que visam representar o mundo, mas seu caráter diacrônico e codificado, na forma de uma sequência de zeros e uns, permite que os tratemos como texto. Assim como podemos nos remeter ao texto de uma imagem digital (sua representação binária que permite que a inspecionemos com um editor de texto num computador), mensagens de texto podem ser referidas como imagens textuais (cf. capítulo 4). Na prática, há quase sempre elementos do que Agre (1994) chamou de vigilância e captura, os dois modelos são perspectivas sobre o mesmo processo. A câmera (e outros sensores) frequentemente está conectada ao computador.

Finalmente, Nick Seaver (2019a) localiza um outro deslocamento no contexto das tecnologias digitais contemporâneas: da previsão à captura, ou seja, da ênfase na inferência informacional para a incidência informacional:

Este movimento que Seaver denomina de virada captológica (*captological turn*) tem como principal referência o trabalho desenvolvido por B.J Fogg, fundador do Persuasive Technology lab e criador do campo de pesquisa que ele designou por “captology”, termo derivado da sigla em inglês de computers as persuasive technologies. Ligado à universidade de Stanford, Fogg é um dos precursores no desenvolvimento de modelos e métodos no campo do “behavioral design”, que combina teorias da psicologia behaviorista com a psicologia cognitivo-comportamental, a economia comportamental e as neurociências em aplicações para a economia digital e a indústria computacional (Bruno, Bentes, e Faltay 2019:12–13).

Para Bruno et al. (2019), enquadrando-a em sua noção de “economia psíquica dos algoritmos”, essa virada indica o momento em que as plataformas deixam de privilegiar as avaliações explícitas de usuárias, para focar em métricas implícitas e na “tonalidade psíquica e emocional”: “Por exemplo, ao invés de priorizar notas de classificações, comentários, compartilhamentos, tornam-se cada vez mais relevantes para o funcionamento desses sistemas certas informações como o tempo médio gasto em tipos diferentes de postagens, a pausa em um vídeo, o padrão de navegação, pular uma música recomendada ou um determinado conteúdo” (Bruno et al. 2019:13).

No nosso caso, preocupados em última medida com a ação que o AdM opera sobre a atividade das usuárias, esse é um aspecto relevante, pois relaciona diretamente a prática de captura com aquilo que chamamos de modulação. Embora não haja sempre nas patentes mecanismos explícitos visando aumentar o tempo de uso ou engajamento da usuária, e que, como veremos abaixo na tabela 5.1, nosso recorte não inclua patentes que modelam a emoção especificamente, a própria ideia de sugestões e substituições de ação “não-obstrusivas” nos remete a uma tentativa de ação que visa não impedir a interação contínua da usuária com o sistema algorítmico. Isso mostra a indissociabilidade desses processos, que funcionam juntos no interior das técnicas de AdM. Aqui fica claro também como as distinções entre captura e modulação não são reais e objetivas, sendo, antes, analíticas. Pois a captura é já modulação, enquanto que a a modulação é aparelho de captura; e o mesmo em relação à imagem.

3.2 A informação digital e o aprendizado de máquina como aparelho de captura

Embora Zuboff, Agre e Seaver tenham contribuições importantes para justificar uma mudança de foco do paradigma de vigilância para o de captura, nos ajudando a perceber aspectos do extrativismo presente nas patentes, foi sobretudo em Deleuze e Guattari (2011–2012) e sua noção de “aparelho de captura” que encontramos o modo de mais precisamente conceituar a conversão ativa da dimensão informacional do sujeito para uma realidade numérica apropriada. Mas, antes disso, convém definirmos mais precisamente o que denominamos “informação digital”.

3.2.1 Informação e informação digital

O filósofo Gilbert Simondon, em sua exposição sobre *A individuação à luz das noções de forma e informação* (Simondon 2020b) possui uma potente definição de informação que dialoga, simultaneamente, com a filosofia e com a cibernética. Em suas próprias palavras:

uma informação nunca é relativa a uma realidade única e homogênea, mas a duas ordens em estado de *disparação*: a informação [...] nunca é depositada numa forma que pode ser dada; ela é a tensão entre dois reais díspares, é a *significação que surgirá quando uma operação de individuação descobrir a dimensão segundo a qual dois reais díspares podem devir sistema*; portanto, a informação é um encetante de individuação, uma *exigência de individuação*, jamais uma coisa dada; não há unidade e identidade da informação, pois a informação não é um *termo*; ela supõe tensão de um sistema de ser; só pode ser inerente a uma problemática; a informação é *aquilo pelo qual a incompatibilidade do sistema não resolvido devém dimensão organizadora na resolução*; a informação supõe uma *mudança de fase de um sistema*, pois ela supõe um primeiro estado pré-individual que se individua segundo a organização descoberta; a informação é a formula da individuação, fórmula que não pode preexistir a essa individuação; poder-se-ia dizer que a informação está sempre no presente, atual, porque ela é o sentido segundo o qual um sistema se individua (Simondon 2020b:26–27).

Ou seja, para Simondon o termo “informação” diz respeito a uma “significação que surge de uma disparação”, diferente da forma, que é “termo único” (Simondon 2020b:33). Nesse sentido, as informações são “singularidades reais” subjacentes aos processos de tomada de forma (Simondon 2020b:54), são

eventos relacionais do qual a eficácia depende menos da “fonte”, do que de uma condição de “metaestabilidade” no “receptor” (Simondon 2020a) e, por isso, também estão próximas da definição de Bateson (1979), citado em Garcia dos Santos (2003a:13): são “diferença que faz uma diferença”. É informação não simplesmente como qualidade ou como quantidade, como gostariam respectivamente a “boa forma” da Gestalt (teoria ou psicologia da forma) ou a “entropia” da teoria da informação de Shannon, mas como “intensidade” (Simondon 2020b:360). É por isso também que a máquina, em essência, não possui informação, mas forma, mesmo sinais de informação só aparecem quando a máquina está em relação com um sujeito (Simondon 2020c:211, *passim*). Daí deriva-se o seguinte: “a informação, atividade relacional, não pode ser quantificada abstratamente, mas somente caracterizada por referência às estruturas e aos esquemas do sistema em que ela existe; não se deve confundir a informação com os sinais de informação, que podem ser quantificados, mas que não poderiam existir sem uma situação de informação, isto é, sem um sistema” (Simondon 2020b:346–47, n. 1).

A partir disso, entendemos que informação é esta singularidade real e atual que surge na resolução de um conflito entre heterogêneos no interior de um sistema, propiciando um processo de tomada de forma, uma individuação. Os sinais de informação são sinais quantificáveis que podem significar informação, mas a depender de uma “situação de informação”, isto é, da existência um receptor que está propício a mudar sua situação no sistema¹⁷. A configuração de uma situação de informação, ou sistema, é chamada, por Simondon de “informação primeira” (e.g. Simondon 2020b:27. n. 10). Explicitamos que informação pouco tem a ver aqui com seu uso, talvez mais jornalístico, que a identifica com o conteúdo de uma mensagem, ou com uma notícia sobre um fato.

Mas, a partir disso, o que seria uma informação digital? A expressão trás algum estranhamento, haja vista que o primeiro termo remete a uma singularidade, a uma intensidade, enquanto que o segundo remete justamente à natureza numérica, codificada e, portanto, já quantificável em alguma medida¹⁸. Garcia dos Santos (2000) mostrou bem como a tecnociência foi capaz de decodificar o código genético, tornando-o passível de manipulação, criando uma noção instrumentalizada de informação genética¹⁹. “Decifrado e rompido, numa palavra, decodificado, o código genético foi envolvido numa operação de axiomatização que visa reterritorializá-lo e inscrevê-lo no regime da propriedade privada, transformando a informação enquanto diferença qualitativa numa diferença quantitativa e abstrata” (Garcia dos Santos 2000:418–19). No caso da informação digital, ou mais precisamente: informação eletrônico-digital, o processo é análogo: a reconfiguração de diversas atividades que passam a ser mediadas pelo computador digital faz convergir no substrato microeletrônico uma miríade de processos informacionais. Essa mediação pelo computador digital, do qual a eletricidade e os padrões eletrônicos são, estes sim, verdadeiros sinais de informação quantificáveis, instrumentalizam e condicionam os processos de

¹⁷Essa situação é denominada de “metaestável”, em oposição à situação “estável”, em que não há mudança possível, e “instável”, de mudança absoluta. Uma posição metaestável quer dizer uma situação em que o sistema pode cambiar entre diferentes estados relativamente estáveis.

¹⁸Em geral são dois os sentidos de digital. Por um lado, de forma ampla, o termo digital é entendido como o oposto de analógico: aquilo que é numérico, codificado e discreto em oposição ao incontável, decodificado e contínuo. Por outro lado, o termo também é utilizado de forma mais restritiva para denotar aquilo relativo às tecnologias eletrônicas digitais, como no caso do computador digital, sistemas computacionais digitais, entre outros.

¹⁹“Ora, a manipulação efetuada pela engenharia genética consiste na desarticulação e rearticulação de processos inframoleculares, no rompimento das barreiras entre as espécies dentro de cada espécie, na alteração, embaralhamento e artificialização das sequências genéticas, na produção de seres inéditos, monstruosos, como a mulher-farmácia, animais transgênicos, bactérias que comem petróleo, tomates que resistem ao tempo e não apodrecem” (Garcia dos Santos 2000:417–18).

tomada de forma que passam a ser algorítmicamente mediados. Nessa situação, há um estreitamento entre as informações e as reconfigurações de bits e bytes, ou seja, da informação digital. Desse modo, chamamos de informação digital as singularidades que dão origem a processos de tomada de forma que acontecem em um sistema de computação (“situação de informação”) e que, por isso, possuem hoje uma relativa manipulabilidade.

3.2.2 Aparelho de captura e aprendizado de máquina

Em *Mil Platôs*, Deleuze e Guattari (2011–2012) denominam *aparelho de captura* o processo maquínico²⁰ associado às sociedades com Estado, segundo o qual uma realidade heterogênea seria abstraída e homogeneizada visando a extração de um excedente que comporia a acumulação capitalista. Assim, território, atividade e troca foram, respectivamente, abstraídos em terra (e renda de terra), trabalho (e sobretrabalho) e moeda (e imposto)²¹. Essa abstração ocorre por meio de duas operações que se remetem reciprocamente:

O que forma o aparelho de captura são as duas operações que se encontra a cada vez nos modos convergentes: comparação direta, apropriação monopolista. A comparação sempre supõe a apropriação: o trabalho supõe o sobretrabalho, a renda diferencial supõe a absoluta, a moeda de comércio supõe o imposto. O aparelho de captura constitui um espaço geral de comparação e um centro móvel de apropriação. Sistema muro branco-buraco negro, tal como vimos anteriormente constituindo o rosto do déspota. Um ponto de ressonância circula num espaço de comparação e, circulando, traça esse espaço (Deleuze e Guattari 2011–2012, v. 5, p. 148).

Desse modo, a captura é a diferença ou excesso que constitui o mais-valor, o sobretrabalho ou o sobreproduto (no caso da atividade). Mas a captura só se realiza porque a possibilidade de um excesso já está contido virtualmente naquilo que é objeto de captura. O mecanismo de captura já existe nos conjuntos sobre os quais ela opera (Deleuze e Guattari 2011–2012, v. 5, p. 152). O que quer dizer que o que possibilita a apropriação do território e a extração de uma renda de terra é a possibilidade de sua quantificação (medida da área) que é imanente ao próprio território. O mesmo com a atividade (medida pelo tempo) e com as trocas (medidas pela moeda).

²⁰“Nós definimos as formações sociais por *processos maquínicos* e não por modos de produção (que, ao contrário, dependem dos processos). Assim as sociedades primitivas se definem por mecanismos de conjuração-antecipação; as sociedades com Estado se definem por aparelhos de captura; as sociedades urbanas, por instrumentos de polarização; as sociedades nômades, por máquinas de guerra; as organizações internacionais, ou antes, ecumênicas, se definem enfim pelo englobamento de formações sociais heterogêneas. Ora, precisamente porque esses processos são variáveis de coexistência que constituem o objeto de uma topologia social, é que as diversas formações correspondentes coexistem” (Deleuze e Guattari 2011–2012, v. 5, p. 135).

²¹Em relação à atividade e ao território: “O sobretrabalho não é o que excede o trabalho; ao contrário, o trabalho é o que se deduz do sobretrabalho e o supõe. É só aí que se pode falar de um valor-trabalho, e de uma avaliação que se apoia na quantidade de trabalho social, enquanto que os grupos primitivos estavam num regime de ação livre ou de atividade de variação contínua. No sentido de que ele depende do sobretrabalho e da mais-valia, o lucro do empreiteiro constitui um aparelho de captura, tanto quanto a renda do proprietário: não é somente o sobretrabalho que captura o trabalho, e não é somente a propriedade que captura a terra, mas o trabalho e o sobretrabalho são o aparelho de captura da atividade, como a comparação das terras e a apropriação da terra são o aparelho de captura do território” (Deleuze e Guattari 2011–2012.5, p. 146).

Inspirados nessa análise, propomos chamar de captura a apropriação do potencial de antecipação, classificação, geração, etc. contidos na informação digital pelo aprendizado de máquina. Como no aparelho de captura introduzido por Deleuze e Guattari, a informação é apropriada na medida em que os dados produzidos e extraídos, uma vez processados pelo modelo de aprendizado de máquina, são diretamente comparáveis e podem ser monopolisticamente apropriáveis pela empresa que detém o modelo. Essa captura provê um potencial de inferência e incidência informacional que pode ser utilizado, depois, não só em um contexto de mercado de dados, mas principalmente para a modulação de comportamento de usuárias (Amadeu da Silveira 2019). Uma análise similar é proposta por Zuboff (2018) ao apresentar o “comportamento” como mercantilização da “realidade”, colocando-a ao lado da “terra”, do “trabalho”, e do “dinheiro” como uma “quarta mercadoria fictícia do capitalismo”, segundo a classificação clássica de Karl Polanyi. Couldry e Mejias (2019) também argumentam que o colonialismo de dados expande a apropriação capitalista que passa a se exercer sobre a totalidade da vida humana. Nosso argumento é mais restrito. Não acreditamos que a captura opere sobre a realidade ou sobre a vida humana, mas sobre a informação – como já sugerido por Garcia dos Santos (2003a) isso se dá tanto em relação à informação genética, no caso da apropriação da vida de um organismo em potencial pela patente de seu código genético, por exemplo, quanto em relação à informação digital, por exemplo na criação de um software com licença privada. A informação, entendida como singularidades subjacentes aos processos de tomada de forma, é uma dimensão da realidade e da vida humana, mas não as esgotam.

Laymert Garcia dos Santos, partindo de Simondon e Bateson para mostrar a centralidade da informação para a aproximação entre desenvolvimento tecnocientífico e acumulação capitalista na virada do século XX para o XXI²² e que ampararia sua noção de “virada cibernética”, a define como “a singularidade real por meio da qual uma energia potencial se atualiza, através da qual uma incompatibilidade é superada” (Garcia dos Santos 2003a:13). É exatamente o que ocorre aqui. O que é capturado no modelo de aprendizado de máquina é a virtualidade dos processos digitais que são potencial para disparações, isto é, mudanças e resolução de tensões, nos meios digitais. “Se a técnica é veículo de uma ação que vai do homem [*sic*] ao universo e de uma informação que vai do universo ao homem, [ela] é fator de resolução de um diálogo intenso; o que conta é a interação, o caráter produtivo do agenciamento, e não as partes preexistentes” (Garcia dos Santos 2003a:13–14). Por isso a importância dada ao autor para esse momento, que poderíamos localizar em 1948, ano da publicação simultânea da *Cibernética* de Wiener ([1948] 2007) e da teoria da informação de Shannon (1948): “É preciso, portanto, perceber a virada cibernética como esse ‘movimento comum’ que se dá no campo da ciência e da técnica, a partir do qual se instaura a possibilidade de abrir totalmente o mundo ao controle tecnocientífico por meio da informação”.

Desse modo, para o autor, “a virada cibernética permite que a tecnociência considere tudo o que existe ou existiu como matéria-prima a ser processada por uma tecnologia que lhe agrega valor” e “a informação enquanto diferença que faz a diferença reconfigura o trabalho, o conhecimento e a vida,

²²Sobre a noção de informação como aquela que possibilita pensar a tomada de forma transversalmente ao físico, ao vivo e ao psicossocial, Simondon constata: “Seria preciso definir uma noção que fosse válida para pensar a individuação na natureza física tanto quanto na natureza viva, e em seguida para definir a diferenciação interna do ser vivo que prolonga sua individuação separando as funções vitais em fisiológicas e psíquicas. Ora, se retomamos o paradigma da tomada de forma tecnológica, encontramos uma noção que parece poder passar de uma ordem de realidade a outra, em razão de seu caráter puramente operatório, não vinculado a esta ou àquela matéria, e definindo-se unicamente em relação a um regime energético e estrutural: a noção de informação” (apud Garcia dos Santos 2003a:13).

enquanto a virada cibernética transforma o mundo num inesgotável banco de dados” (Garcia dos Santos 2003a:19, 17). Um diagnóstico similar é feito por Donna Haraway, que entende a informação como fundo comum transversal aos domínios físico, biológico e comunicacional, “aquele tipo de elemento quantificável (unidade, base da unidade) que permite uma tradução universal e, assim, um poder universal sem interferências, isto é, aquilo que se chama de ‘comunicação eficaz’ ” (Haraway 2009:65). Esse olhar sobre a informação e o código vincula fortemente a captura e a ação, de modo que, se aqui comentamos o processo enfocando o primeiro, seus desdobramentos como modulação são explicitados no capítulo 5.

Assim como o território ou a atividade, a informação é heterogênea qualitativamente, é inclusive índice dessa heterogeneidade, da disparação, mas pode ser comparada quantitativamente quando seu processo é reificado na forma de dados (medidos em *bits*). Mais ainda, no interior do modelo de AdM, dados sobre informações muito distintas funcionam conjuntamente como parâmetros numéricos que orientam a operação de funções matemáticas configuráveis (Gonçalves 2022). Especialmente, dados processados por modelos de AdM são apropriados nos parâmetros do modelo. Como em Deleuze e Guattari (2011–2012), é a apropriação monopolista de um excedente que determina a comparação direta da informação. Ambas funcionam juntas formando um aparelho de captura. Portanto, o AdM pode ser visto como um aparelho de captura que compara diretamente a informação digital (das usuárias, sobretudo, mas não só) e apropria monopolisticamente seu excesso, *um excedente informacional* (Parra 2022), *uma imagem* (capítulo 4). Isso é possível porque a informação digital supõe a geração do dado e do parâmetro numérico, do mesmo modo que o território supõe suas possibilidades de cercamento e a atividade humana as de apropriação de sobretrabalho (tabela 3.2).

Tabela 3.2: Comparação dos diferentes aparelhos de captura, ênfase na nossa proposta de aparelho de captura de informação digital.

capturado	aparelho de captura	excedente
território	terra	renda de terra
atividade	trabalho/sobretrabalho	mais-valor
troca	dinheiro	imposto
informação digital	parâmetros numéricos do modelo de AdM	excedente informacional/imagem

3.2.3 Exposição de um aparelho de captura de informação digital

Na patente *Using live data streams and/or search queries to determine information about developing events* (US 11366812) (Carbune e Feuz 2022), por exemplo, o seguinte caso é apresentado:

Na FIG. 2 [figura 3.1], uma primeira usuária muito próxima do incêndio publicou nas redes sociais: “Meu deus [OMG]! Há um incêndio na catedral!”. Outra usuária próxima publica: “Deve haver dez caminhões de bombeiros combatendo o incêndio na igreja!” Outra usuária próximo publica: “Espero que as paroquianas tenham conseguido escapar ao fogo...” Em várias implementações, o monitor de transmissão ao vivo [*live stream monitor*] 134 [referência interna a um diagrama da patente] pode analisar uma pluralidade de transmissões ao vivo, incluindo estas três transmissões, para determinar a probabilidade de ocorrência de

um evento em desenvolvimento. O monitor de transmissão ao vivo 134 pode processar sintaticamente e/ou semanticamente, ou fazer com que outro sistema processe sintaticamente e/ou semanticamente, estas atualizações de transmissão de dados ao vivo para determinar que estão relacionadas. Por exemplo, o monitor de transmissão ao vivo 134 pode gerar embeddings [representações numéricas] dessas atualizações de transmissão ao vivo e determinar que estão relacionadas com base no fato de que as representações se agrupam em um espaço latente [proximidade numérica relacionada a uma proximidade semântica, isso será melhor exposto no capítulo 4] (Carbune e Feuz 2022:14).

Neste caso, a informação é tanto o contexto das usuárias (sua localização, por exemplo, determinada pelo GPS)²³, quanto o conteúdo das mensagens publicadas online. No modelo de AdM elas são comparadas (computacionalmente, alterando o valor dos parâmetros numéricos do modelo) e, finalmente, geram uma imagem. Aqui ela é referida explicitamente como *embeddings* (por vezes traduzidos por vetores densos) no espaço latente. O próximo capítulo vai explorar a natureza do que chamamos de imagem, em relação com a noção “êmica” de espaço latente. Por ora, basta perceber que ela é o potencial de ação que surge como resultado da captura. Notemos que para que o método da patente funcione, a informação sobre contexto e mensagens deve ser captada e transformada em dados. Entretanto, os dados, embora condição necessária para que a captura ocorra, não geram por si só um excedente. Um agregado de dados parado no armazenamento interno de um computador não é capaz de inferir o desenvolvimento de um evento. Por outro lado, uma vez que o monitor de transmissão ao vivo processa esses dados utilizando aprendizagem de máquina, a probabilidade de um evento ocorrer pode ser criada²⁴.

Uma vez calculada a probabilidade, o objetivo final do método patenteado se volta para a ação: “uma ação pode [...] ser executada em resposta à detecção do evento recém-desenvolvido [...] por exemplo [...] um evento como um incêndio ou crime pode ser detectado e os serviços de emergência podem ser alertados” (Carbune e Feuz 2022:10). Considerando que um dos principais serviços da Google é seu buscador online (Google Search), a possibilidade prevista na patente de responder a perguntas para as quais um buscador tradicional não obteria resultados (por exemplo, a definição de um evento em desenvolvimento a partir de perguntas pode gerar respostas para próximas perguntas similares) e que, além disso, funciona na lógica captológica de aumentar o engajamento e a permanência da usuária em seu sistema²⁵, os benefícios que a patente provê para a empresa são evidentes. Ademais, a patente mostra um sentido de aproximação da corporação com instituições estatais (corpo de bombeiros, departamento

²³Por exemplo: “No entanto, outros sinais de informação [*informational signals*] também podem ser considerados, além ou em vez dos mencionados acima, incluindo, mas não se limitando a, informações demográficas sobre as usuárias que postam (e.g., idades, gêneros), preferências ou interesses conhecidos das usuárias que postam (e.g., determinados a partir de seus perfis online), afiliações conhecidas das usuárias que postam (e.g., universidades [*alma maters*], organizações, empregadoras) e assim por diante” (Carbune e Feuz 2022:14).

²⁴Sobre a geração de uma probabilidade a partir do processamento dos dados: “Intuitivamente, essa probabilidade P pode representar uma probabilidade de ocorrência de um evento em um determinado nó x (e.g., do modelo de rede subjacente 570), condicionada a um ou mais recursos da usuária específica e da estrutura-de-dados/embedding (e.g., de uma vizinhança local). Assim, a entrada para o(s) modelo(s) de aprendizado de máquina pode incluir uma estrutura de dados, como um gráfico direcionado ou uma incorporação gerada a partir dele, que modela a difusão de informações sobre o evento em desenvolvimento em uma população, bem como um ou mais recursos ou atributos da usuária, como preferências, interesses, idade, sexo, vários outros dados demográficos, localização, etc.” (Carbune e Feuz 2022:16).

²⁵“Quando um evento em desenvolvimento for considerado provável e relevante para um determinado indivíduo, outras consultas associadas ao evento poderão ser aproveitadas para fornecer a esse indivíduo informações adicionais sobre o evento em desenvolvimento. Por exemplo, suponha que o indivíduo emita uma consulta sobre o evento (“Qual é o tamanho do in-

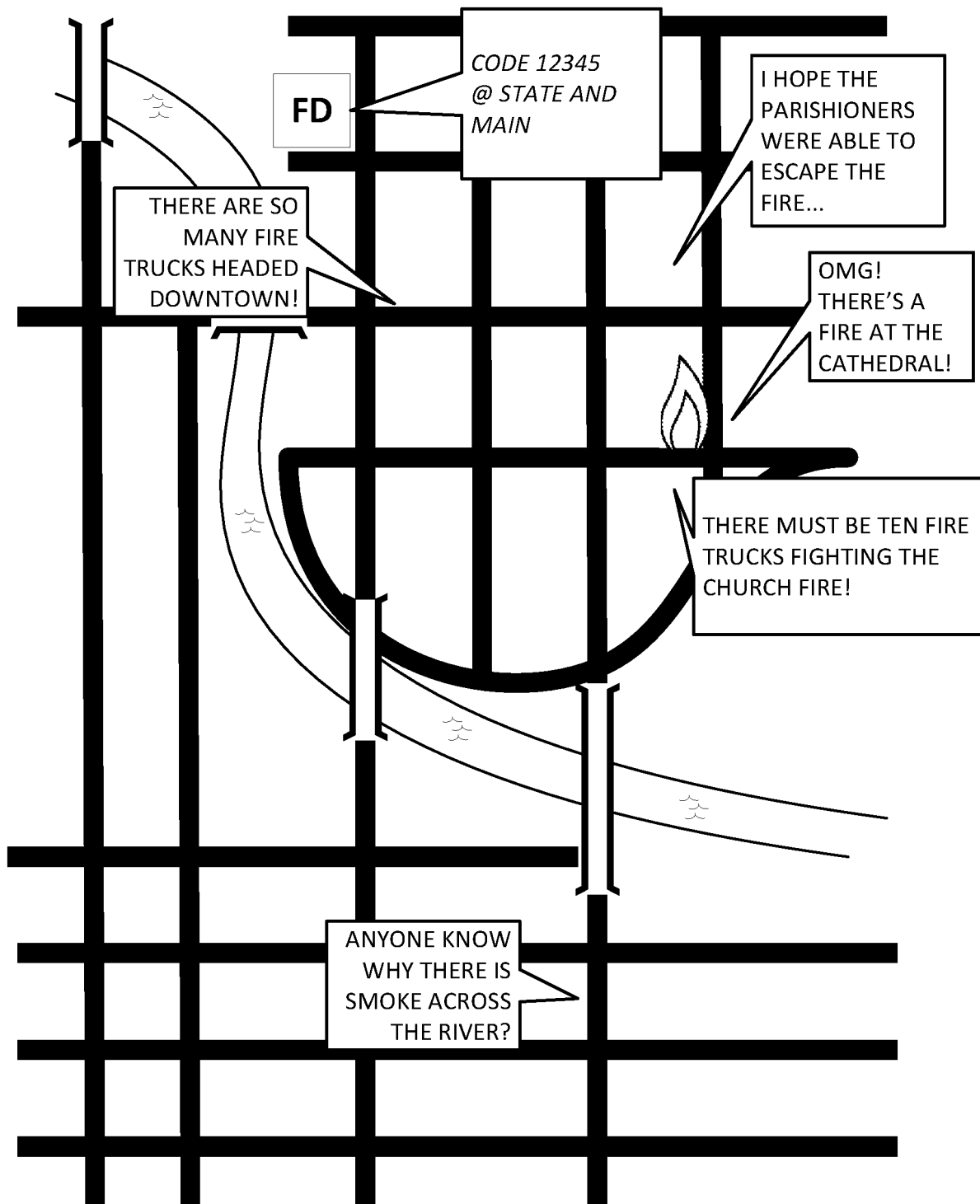


Fig. 2

Figura 3.1: Situação exemplo de evento em desenvolvimento. Aqui o símbolo de uma chama no canto superior direito indica um incêndio na catedral, enquanto que o bloco “FD” representa o departamento de bombeiros e os balões, a localização e o conteúdo de mensagens postadas por usuárias cujas mensagens se relacionam com o evento de incêndio. Respectivamente, da esquerda para a direita e de cima para baixo: “Há muitos caminhões de bombeiro a caminho do centro”, “Código 12345 @ state and main”, “Eu espero que os paroquianos escaparam do incêndio..”, “Meu deus! Há um incêndio na catedral!”, “Deve haver dez caminhões de bombeiro lutando contra o incêndio da igreja”, “Alguém sabe por que há fumaça do outro lado do rio?”. Fonte: Carbune e Feuz (2022).

de polícia) que ressoa com a constatação em Dardot e Laval (2016), por exemplo, do papel conjunto das corporações e dos estados no exercício de poder neoliberal²⁶.

Assim, a própria captura é, já, representação e ação, estando intimamente relacionada com os dois processos analisados nos próximos capítulos: a imaginação e a modulação. Mais um exemplo pode esclarecer o tipo de consequência que tal método poderia ter na sociedade:

Em algumas implementações, o monitor de transmissão ao vivo 134 pode simplesmente monitorar uma seleção aleatória de transmissões de dados ao vivo. Além disso, ou alternativamente, em algumas implementações, o monitor de transmissão ao vivo 134 pode monitorar fluxos de dados ao vivo particularmente ativos, por exemplo, sob a suposição de que essas fontes ativas têm maior probabilidade de publicar atualizações ao vivo sobre eventos em desenvolvimento que testemunharem. Além disso, ou como alternativa, em algumas implementações, *o monitor de transmissão ao vivo 134 pode selecionar fontes de áreas geográficas nas quais é mais provável que ocorram eventos em desenvolvimento. Por exemplo, o monitor de transmissão ao vivo 134 pode favorecer uma primeira fonte que mora em uma área densamente povoada em vez de uma segunda fonte que mora em uma área rural.* Intuitivamente, é mais provável que a primeira fonte seja testemunha de eventos em desenvolvimento do que a segunda, especialmente eventos em desenvolvimento causados por pessoas ou relacionados a elas. Além disso, ou como alternativa, em algumas implementações, o monitor de fluxo ao vivo 134 pode monitorar uma pluralidade de fluxos de dados ao vivo associados a fontes distribuídas em uma grande área geográfica. Isso pode aumentar a probabilidade de que um evento em desenvolvimento em um determinado local seja testemunhado por pelo menos uma das fontes (Carbune e Feuz 2022:17, grifo nosso).

Desse modo, ao intensificar o seu funcionamento em áreas consideradas mais propensas a postarem conteúdos sobre eventos em desenvolvimento, o método patenteado não só reproduz essa desigualdade informacional, como também a naturaliza e contribui ativamente em sua construção. Sendo

cêndio na Catedral?”) para a qual ainda não há informações disponíveis em um mecanismo de busca convencional. *Em vez de ser notificado de que não há informações disponíveis, o indivíduo pode receber sugestões de consultas alternativas para as quais pode haver mais informações disponíveis.* Além disso, ou alternativamente, o indivíduo em particular pode *receber informações que respondam a essas sugestões de consulta alternativas.* / Às vezes, as consultas de outros indivíduos podem fornecer informações adicionais, especialmente quando esses indivíduos são testemunhas em primeira mão do evento. Por exemplo, um primeiro indivíduo pode enviar uma consulta como “quais artefatos estão em risco devido ao incêndio na catedral?” Outros indivíduos com mais conhecimento sobre a coleção de artefatos da catedral podem ter enviado anteriormente consultas como “eles salvaram a do incêndio da catedral?” *Essas consultas enviadas anteriormente podem ser extraídas para fornecer ao primeiro indivíduo, por exemplo, sugestões de consultas alternativas que orientem o indivíduo quanto a uma consulta melhor e/ou forneçam ao indivíduo informações que respondam à sua própria consulta, como o fato de que a provavelmente está em risco”* (Carbune e Feuz 2022:11, grifo nosso).

²⁶“Em algumas implementações, as técnicas descritas neste documento para identificar eventos em desenvolvimento podem ser usadas para acionar automaticamente ações responsivas de várias entidades. Por exemplo, as três primeiras postagens de usuários descritas anteriormente podem ser suficientes para que o monitor de transmissão ao vivo 134 determine, por exemplo, com algum limite de confiança, que há de fato um incêndio ocorrendo na catedral. Ao atingir esse limite, o monitor de transmissão ao vivo 134 ou outro componente do sistema de conhecimento 130 pode *notificar automaticamente o corpo de bombeiros sobre o incêndio na catedral, por exemplo, por meio de uma chamada telefônica automática, alarme automático em um sistema de resposta a emergências etc.* Nesse cenário, a publicação do corpo de bombeiros na FIG. 2 pode, de fato, ser parte da reação do corpo de bombeiros às três atualizações iniciais da transmissão ao vivo indicando a ocorrência do evento de incêndio na catedral (Carbune e Feuz 2022:15, grifo nosso).

as tecnologias da Google importantes *loci* de busca de conteúdos, sendo mesmo pontos de passagem obrigatória para muitas atividades cotidianas, a sub-representação de eventos em áreas menos populosas ou rurais tende a marginalizar estes espaços, pois não só não encontraríamos conteúdos sobre eles, mas também a ação de outras instituições sobre essas áreas ficariam prejudicadas. Privilegiar-se-iam os centros, onde a densidade informacional é maior, em detrimento da periferia.

De modo geral, vemos como informações heterogêneas referentes à usuária, seu contexto e o evento são comparadas no modelo criando, por meio de relações entre elas, novos conhecimentos e, principalmente, possibilidades de ação. Assim, a possibilidade de comparação imanente à informação digital permite a construção de aparatos que visam redirecioná-las para uma apropriação monopolista da Google que passa tanto a saber mais que buscadores tradicionais porventura concorrentes, quanto a ser capaz de agir de modo singular, por exemplo, em se conectando com outras entidades públicas e controlando, em certa medida, os fluxos policiais e relativos ao corpo de bombeiros.

3.3 O processo de captura no aprendizado de máquina

Mas como a captura se desenrola no interior das técnicas analisadas? Tomemos como outro exemplo a patente *Automatically resolving, with reduced user inputs, a set of activity instances for a group of users* (US11108889), cuja ideia geral é prover um método para facilitar a resolução de conflitos na proposição de atividades em grupo (sugerindo atividade/local/data) em um aplicativo de troca de mensagens (Lance e Koonce 2021) (figura 3.2).

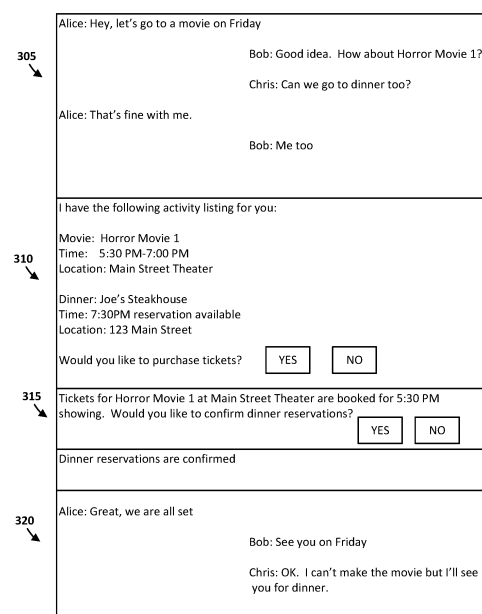


FIG. 3

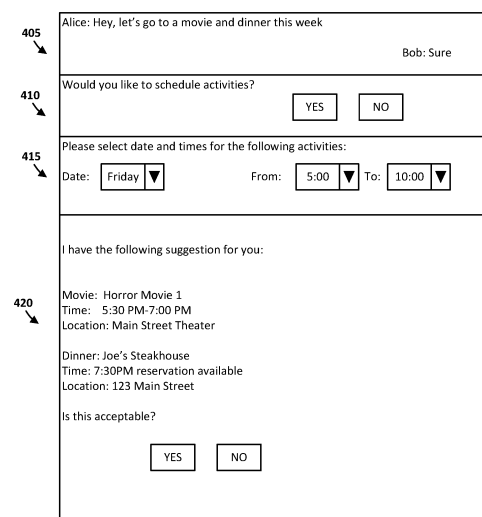


FIG. 4

Figura 3.2: Exemplos do funcionamento dos métodos de resolução de conjunto de atividades. Em ambos os casos, há uma conversa em grupo do ponto de vista de Alice, que conversa com Bob e Chris (apenas na imagem da esquerda). À esquerda, podemos ver que a conversa é interrompida para sugerir atividades que devem ser aceitas ou recusadas por Alice (310 e 315). À direita, a conversa é interrompida para que Alice também inclua explicitamente uma data e horário para o agendamento de uma atividade (415). Fonte: Lance e Koonce (2021).

Nesta, a proposição do método é justificado não só pela facilidade, mas também por prover maior acessibilidade:

Ademais, a determinação automática de conjunto(s) de instâncias de atividades é opcionalmente realizada utilizando um ou mais modelos de aprendizado de máquina que são treinados para otimizar a probabilidade [*likelihood*] de que o grupo de usuárias vai considerar satisfatória a quantidade limitada de conjunto(s). Essa determinação automática de uma quantidade limitada satisfatória de conjunto(s) resulta na redução de entradas de usuárias do grupo para coordenar um conjunto de instâncias de atividades, assim conservando recursos do dispositivo cliente e/ou da rede. Adicionalmente ou alternativamente, a redução de entradas de usuárias pode ser benéfica para usuária(s) com habilidade motora reduzida [*limited dexterity*] (Lance e Koonce 2021:9).

Entendemos que o automatismo pode sim ajudar a resolver problemas de acessibilidade, além de concordar com Simondon (2020b:531) que toda civilização necessita de um nível de automatismo para que seja estável e coesa. Entretanto, neste trabalho nos interessa notar como a máquina e as agências nela concretizadas (no caso, da Google, empresa cujo objetivo final é o lucro) participam ativamente de processos sociais, substituindo parcialmente a agência humana²⁷. Ou seja, estamos mais preocupados com a necessidade que uma civilização tem do dinamismo e da criatividade para não se fechar em uma adaptação estereotipada (Simondon 2020b:531) e suas consequências éticas, estéticas e políticas, por exemplo. Assim, numa assembleia, numa conversa em grupo visando determinar uma atividade que se deseja realizar, o dissenso, as negociações e, em alguma medida, a própria atividade política, são aqui, ao menos em parte, substituídos pela agência objetivada da máquina. Que é também agência de usuárias, de programadoras, etc. É sempre uma agência entre, uma agência maquinaica. Mas que, justamente por isso, também envolve agências específicas da Google.

Em Lance e Koonce (2021), a captura da informação ocorre sobretudo por meio do processamento de linguagem natural (NLP, *natural language processing*) sobre mensagens de texto em um grupo de usuárias. Por exemplo, a partir dele infere-se o tipo de atividade (“*movie* [cinema]”, “*dinner* [jantar]”), a instância de atividade (“Joes Steakhouse”), pontuação, localização, tempo e qualidade global. Além disso, o método pode lançar mão de uma interface de usuária (UI, *user interface*) para determinar a atividade a partir da ação da usuária (por exemplo, data e horário na imagem da direita na figura 3.2). A localização pode ser captada pela patente de diferentes modos (GPS, endereço, caminho para uma atividade). O mesmo ocorre para o tempo (início e fim – de um filme ou museu, por exemplo -, horário de reserva, tempo estimado de uma viagem).

A patente prevê ainda codificações e métricas internas. Por exemplo, uma pontuação de viabilidade supervisionada, isto é, indicação de uma ou mais usuárias sobre a real viabilidade de um conjunto ordenado de atividades. Isso pode ser feito, por exemplo, utilizando sensores que permitam identificar se uma usuária de fato esteve presente nas proximidades de uma atividade sugerida, ou pela indicação no fio de comunicação de que uma usuária é incapaz de ou não se interessa em realizar uma ou mais das atividades sugeridas:

Em algumas implementações, a pontuação de viabilidade supervisionada pode ser baseada em se ao menos uma das usuárias realizou o conjunto de instâncias de atividade ordenadas. Por exemplo, o dispositivo cliente de uma ou mais usuárias pode prover informação de loca-

²⁷ Este argumento foi primeiramente explorado por nós no trabalho “Coletivos maquinaicos e sistemas digitais” (Gonçalves e Padovan 2024b).

lização para a usuária e um ou mais componentes do mecanismo [engine] de atividade em grupo *pode identificar que a usuária não esteve presente na proximidade de uma instância de atividade do conjunto ordenado durante o tempo associado com a instância de atividade* (e.g., não está perto do cinema durante o tempo e exibição [Showtime] de uma instância de atividade de filme de um conjunto ordenado). Em algumas implementações, *uma ou mais usuárias podem indicar no fio de comunicação [communication thread] que a usuária é incapaz ou sem interesse em realizar uma ou mais das atividades de um conjunto ordenado*. Por exemplo, referindo novamente na figura 3 [imagem da esquerda na figura 3.2], no bloco de comunicação, “Chris” indica que é incapaz de ir à instância de atividade de cinema do conjunto ordenado provido. Então, uma pontuação de viabilidade supervisionada pode ser determinada para prover conjuntos ordenados que indicam que o conjunto ordenado provido não satisfaz completamente o grupo. Em algumas implementações, *uma ou mais das usuárias pode ser questionada [prompted]*, ou via aplicação de mensagens ou outro aplicativo, para avaliar [rate] ou caracterizar a pertinência [suitability] do conjunto ordenado de instâncias de atividades providos. As respostas de uma ou mais usuárias podem ser utilizadas para determinar a pontuação de viabilidade supervisionada para o conjunto ordenado provido (Lance e Koonce 2021:17, grifo nosso)

Desse modo, percebemos que o método não serve apenas para sugerir ações, mas também para treinar o modelo de AdM para que seja capaz de sugerir-las. Em outras palavras, criar uma imagem de quais sugestões de atividade-local-tempo são efetivamente realizadas pelas usuárias. Isso pode ser feito de três modos:

1. Pelo feedback explícito dado por usuárias em uma UI que interrompe o fluxo de mensagens;
2. De forma não-obstrusiva, verificando imperceptivelmente se uma usuária realizou a tarefa sugerida baseado, por exemplo, em sua localização;
3. De forma não-obstrusiva, analisando imperceptivelmente o texto das mensagens que poderiam indicar a possibilidade ou não da realização da atividade.

Couldry e Mejias (2019) mostram como a produção de dados a partir do que os autores chamam de relações de dados, “novos tipos de relações humanas [e, acrescentaríamos, não humanas] que possibilitam a extração de dados para mercadorização” (Couldry e Mejias 2019:337) são um momento importante dos novos modos de apropriação capitalistas sobre a informação. Além disso, as bases de dados proprietárias se sustentam, não somente no uso – legal ou ilegal – de uma vasta quantidade de dados já disponibilizados online, mas também na dadoficação de informação privada que circula somente nos circuitos internos de seus produtos (por exemplo, o uso massivo do buscador Google ou do sistema operacional Android, faz com que a Google detenha uma quantidade de dados exclusivos que não estão disponíveis para outras empresas ou usuárias). Entretanto, como vimos, sozinhos, mesmo que agrupados, os dados não formam uma imagem capaz de representar e agir sobre a informação digital. É somente no interior do modelo de aprendizado de máquina que as relações entre os dados são traçadas permitindo uma verdadeira comparação direta e apropriação monopolista. Assim, a dadoficação é um momento intermediário do processo de captura (ver abaixo figura 3.3).

Antes de se tornarem dados, os sensores e a interface captam²⁸ sinais de informação. Aqui, há duas possibilidades. No caso de sensores como o microfone, a câmera, o acelerômetro ou o giroscópio, a captação se dá sobre um sinal de informação analógico, que ainda não foi quantizada (ondas sonoras e eletromagnéticas, temperatura, etc.)²⁹. Desse modo, os sensores não só permitem ler variações exteriores aos sistemas digitais, mas digitalizam-nas, em um primeiro momento de homogeneização. No caso das interfaces de usuária – algoritmos programados especificamente para gerar dados e extraí-los a partir do desenho de telas e outras interfaces entre o computador digital e a usuária –, a captação já age sobre um sinal de informação digitalizado: mensagens de texto codificadas, imagens digitais, movimentos do cursor do mouse em uma tela, etc.

A informação digitalizada é posteriormente tornada dado: parcela congelada do processo de informação digital que deve estar organizada em uma estrutura pré-determinada. Isso é importante porque esses dados serão agrupados em grandes bases de dados. Estas podem ser vistas como grandes tabelas, em que cada linha representa um dado, também chamado de exemplo no contexto de bases de treino de AdM, e cada coluna representa o conteúdo do dado para tal dimensão específica da estrutura pré-determinada, também chamado de característica neste contexto (tbl. 3.3). Exemplos de bases de dados abertas são o Common Crawl, criado a partir da raspagem de informação textual disponível na internet, e a ImageNet, imagens rotuladas para a tarefa de classificação (Crawford e Paglen 2019).

Tabela 3.3: Exemplo fictício de uma tabela de dados com a estrutura: ID (identificação ou número do exemplo ou dado), localização, acelerômetro, giroscópio, última mensagem de texto e se aceitou ou não determinada sugestão.

ID	localização	acelerômetro	giroscópio	mensagem	...	aceita sugestão
0	(1, 2)	(1, -2, 3)	(-4, 5, -6)	“Oi Bob”	...	1
1	(-3, 4)	(-7, 8, 9)	(10, -11, -12)	“Como vai?”	...	1
2	(5, -6)	(13, 14, -15)	(-16, -17, -18)	“Bom dia”	...	0
...

Jose van Dijck (2014), amparada em Mayer-Schönberger e Cukier (2013), define a dadificação como um “novo paradigma na ciência e na sociedade” constituído pela “transformação de ação social em dados online quantificados, portanto permitindo rastreamento em tempo-real e análise preditiva” (Dijck 2014:198). De um ponto de vista crítico, a autora relaciona esse processo com a “vigilância de dados [*dataveillance*]” – “uma forma cada vez mais preferida de monitorar os cidadãos por meio das mídias sociais e das tecnologias de comunicação online” (Dijck 2014:204) – e com o “dadoísmo” – “uma crença na objetividade da quantificação e no potencial de rastrear todos os tipos de comportamento humano e socialidade por meio de dados online” (Dijck 2014:201). Em nosso esquema, a dadificação tem um sentido ainda mais particular, embora intimamente relacionado com a conceitualização de van Dijck: não só a quantificação abstrata de informação, ação e relações sociais, mas seu armazenamento em estruturas predefinidas que constituem a potente instrumentalidade das atuais grandes bases de dados. Ou seja, uma temperatura pode ser medida por um termômetro digital, mas só se torna dado

²⁸Há aqui uma feliz coincidência na língua francesa. Nesta, os sensores são *capteurs*, isto é, captadores, do verbo *capter* [captar] que é também distinto do verbo *capturer* [capturar], por nós utilizado para designar o processo mais amplo.

²⁹No caso do sistema de posicionamento global (GPS, em inglês), a situação é um pouco mais complicada. Embora modele a localização, uma informação analógica, isso é feito pela trilateração de sinais discretos emitidos por satélites, de modo que a informação captada por ele é, já, digital: <https://education.nationalgeographic.org/resource/gps/> (acesso: 08/08/2024).

se essa medida quantitativa é ulteriormente armazenada na forma (temperatura: 36,0), por exemplo. O sensor de uma câmera fotográfica é constituído por elementos eletrônicos fotossensíveis, mas sem um mapeamento de qual posição do sensor se relaciona com qual pixel a ser mostrado em uma tela, formando uma matriz numérica, um dado visual não é constituído.

Fernanda Bruno e colaboradoras desenvolvem a noção de “economia psíquica dos algoritmos” para descrever investimentos tecnocientíficos e econômicos que compõem “uma nova lógica que investe, cada vez mais, em processos algorítmicos de captura, análise e uso de informações psíquicas e emocionais extraídas de grandes volumes de dados” (Bruno et al. 2019:35). Chamamos atenção para a similaridade entre o nosso esquema e aquele apresentado pelas autoras no contexto da análise do funcionamento dos aplicativos de autocuidado psicológicos, os PsiApps³⁰ (Bruno et al. 2019, na figura 6). Embora o foco das autoras seja outro, enfatizando os termos de uso e os *trackers*³¹ como mediadores no interior dos aplicativos, um processo similar é descrito e que, na prática, muito provavelmente envolve aprendizagem de máquina para o processamento dos dados obtidos.

Esse trabalho, que é parte de um programa mais amplo e que pode ser apreciado panoramicamente na linha do tempo da economia psíquica dos algoritmos³², analisa dez PsiApps com foco: 1) no discurso dos aplicativos; 2) nos termos de uso e *trackers*. O primeiro, camada visível desses aplicativos, mostrou uma ênfase na “autonomia individual”, no autocuidado, na autogestão e autoresponsabilização que se identifica com a racionalidade neoliberal analisada por Dardot e Laval (2016). Mas o segundo mostrou, em uma camada invisível, pela análise “[d]os termos de uso e políticas de privacidade, de um lado; e, de outro, os rastreadores (*trackers*) presentes nos aplicativos”, mediadores informacionais operando a captura de dados.

Nos interessa sobretudo as ressonâncias encontradas em seu trabalho e no nosso em relação à captura operada pelos *trackers*, naquele, e pelo aprendizado de máquina, neste. Ressaltamos, por exemplo, como sua análise mostrou, por um lado, que *a finalidade dos trackers está profundamente atrelada ao mercado capitalista*, sendo as principais funções desempenhadas as de “tracker de anúncios” (33,3%) e “trackers analíticos e de perfilização” (44%)³³; e, por outro lado, *o papel central das big techs nesses “circuitos de captura”*, em particular da Google e da Facebook, pontos de passagem obrigatória e detentoras de 66,6% dos *trackers* analisados (Bruno et al. 2021:46 et seq).

No AdM, as bases de dados continuamente geradas nesse processo são utilizadas em processos iterativos simultâneos de ordenamento ou cálculo dos parâmetros numéricos de um modelo de AdM (Gonçalves 2022). Mesmo Bruno et al. (2021) já destacam a natureza “relacional”, “heterogênea” e “assi-

³⁰Aplicativos que “vão desde técnicas de meditação, passando por controle do sono ou de outras funções emocionais e fisiológicas, até a realização de testes psicológicos e terapia guiada” e que “têm o potencial de oferecer aos usuários informações e serviços, possibilitando acesso a dados em tempo real, orientações sobre condições clínicas e incentivo a hábitos benéficos à saúde psíquica, emocional e física” ao mesmo tempo que “ampliam potencialmente a coleta de informações psicológicas e emocionais, bem como as predições que se pretendem extrair dessas coletas que são valiosas não apenas para fins comerciais, mas também para a pesquisa médica, clínica e científica” (Bruno et al. 2021:35).

³¹“Um *tracker* é um *software* cuja tarefa é coletar informações a partir do uso de um aplicativo. Ele pode informar ao desenvolvedor como o usuário utiliza o aplicativo e o dispositivo móvel, e geralmente é distribuído pelas empresas no formato SDK (Software Development Kit)” (Bruno et al. 2021:40).

³²Disponível em Medialab.UFRJ (2024).

³³“Enquanto os primeiros são utilizados para veicular automaticamente anúncios segmentados por públicos-alvo, gerando receita com sua aplicação; o segundo grupo se destina a capturar automaticamente uma série de eventos e dados do usuário, visando a construção de perfis, extração de métricas, segmentação de audiência, análise de comportamento dos usuários, relatórios de desempenho do aplicativo, correlação de eventos etc.” (Bruno et al. 2021:47).

métrica” da captura, enfatizando como seu processo é “cíclico”, “recursivo” e “performativo”³⁴. No AdM, a captura culmina com o treinamento de modelos, gerando uma imagem (capítulo 4) e possibilitando a modulação (capítulo 5). Um esquema final desse processo é apresentado na figura 3.3.

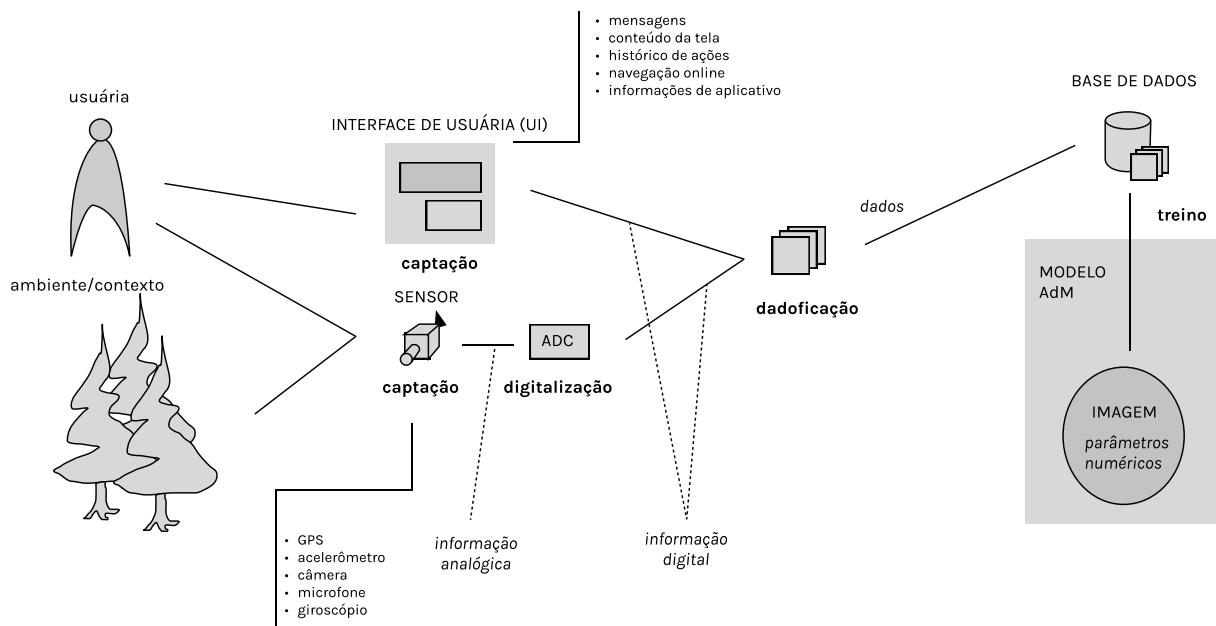


Figura 3.3: Esquema do processo de captura de informação digital. Fonte: elaboração própria.

³⁴ “[O]s dados capturados, analisados, perfilizados e segmentados são continuamente utilizados para testar, modificar e atualizar os aplicativos, gerando novos dados para análise num ‘circuito de retroalimentação’ contínuo, para utilizar um termo caro à cibernética” (Bruno et al. 2021:48).

Capítulo 4

Imaginação

O aprendizado de representação é um conjunto de métodos que permite que uma máquina seja alimentada com dados brutos e descubra automaticamente as representações necessárias para sua detecção ou classificação. Os métodos de aprendizagem profunda são métodos de aprendizagem de representação com vários níveis de representação, obtidos pela composição de módulos simples, mas não lineares, que transformam a representação em um nível (começando com a entrada bruta) em uma representação em um nível mais alto e ligeiramente mais abstrato. Com a composição de um número suficiente dessas transformações, funções muito complexas podem ser aprendidas.

Deep Learning

Yann LeCun, Yoshua Bengio e Geoffrey Hinton

Em vez de evocar lendas de cognição alienígena, é mais razoável considerar o aprendizado de máquina como um instrumento de ampliação do conhecimento que ajuda a perceber características, padrões e correlações em vastos espaços de dados além do alcance humano. Na história da ciência e da tecnologia, isso não é novidade: já foi buscado por instrumentos ópticos ao longo das histórias da astronomia e da medicina. Na tradição da ciência, o aprendizado de máquina é apenas um Nooscópio, um instrumento para ver e navegar no espaço do conhecimento (do grego *skopein* “examinar, olhar” e *noos* “conhecimento”).

The nooscope manifested

Matteo Pasquinelli e Vladan Joler

Desde alguns anos, o uso de IA nas artes tem se proliferado com enorme rapidez, mostrando uma aceleração não apenas no uso dessas técnicas, mas também de suas capacidades: o que gerou muitas dúvidas e questionamentos, por exemplo sobre as possibilidades criativas da inteligência artificial e da dificuldade de separação entre imagens por ela geradas, fotos e ilustrações. Algumas obras recentes mostram a questão exemplarmente: em 2023, o trabalho *Pseudomnesia: The Electrician* [Pseudominésia: o eletricista]¹, do artista alemão Boris Eldagsen, venceu a categoria criativa do Prêmio Mundial de Fotografia da Sony (SWPA, *Sony World Photography Award*), importante concurso internacional de

¹Disponível em: <https://www.eldagsen.com/pseudomnesia/> (acesso em: 23/10/2024).

fotografia. Já em 2024, *FLAMINGONE*², fotografia de Miles Astray, venceu o People's Vote e se classificou em terceiro no Jury Award da categoria destinada a imagens geradas por inteligência artificial (IA) do concurso internacional 1839 Awards.

The Electrician é uma das imagens que compõe a série *Pseudomnesia*, “memórias falsas” em grego, que, segundo Eldagsen (2023a), remete a “uma reminiscência espúria de eventos que nunca ocorreram, em oposição a uma memória que é meramente imprecisa”. Desse modo, não é uma surpresa que tal imagem não seja uma fotografia. A obra, na realidade, é uma construção realizada por meio de inteligência artificial gerativa (IAg), um processo que o autor chama de “promptografia”. Nas palavras do artista, a imagem foi “imaginada por linguagem e mais reeditada entre 20 e 50 vezes através de IAs gerativas, combinando técnicas de ‘inpainting’, ‘outpainting’ e ‘prompt whispering’” (Eldagsen 2023a), respectivamente: interpolação da imagem que visa preencher seu interior, extrapolação da imagem que a expande em alguma dimensão e geração de imagens novas a partir de comandos de texto. Segundo Eldagsen (2023b), *The Electrician* é a primeira imagem gerada por IA a vencer um concurso de fotografia:

O trabalho que o SWPA escolheu é o resultado de uma complexa interação de engenharia de prompt, inpainting e outpainting que se baseia na minha riqueza de conhecimento fotográfico. Para mim, trabalhar com IAs gerativas é uma cocriação, em que eu sou o diretor. Não é sobre apertar um botão – e está pronto. É sobre a exploração da complexidade desse processo, começando com o refinamento de prompts de texto, então o desenvolvimento de um fluxo de trabalho complexo, e a mistura de várias plataformas e técnicas. Quanto mais você cria um fluxo de trabalho e define parâmetros, mais criativa sua parte se torna (Eldagsen 2023b).

Já o trabalho *FLAMINGONE*, cujo título mescla as palavras em inglês para “flamingo” e “desaparecido [*gone*]”, se revelou como a primeira fotografia a vencer um concurso destinado à imagens gerada por IA. Segundo Astray (2024), “não não há muito o que contar” sobre a foto: “quando em Aruba [...] acordando às 5 da manhã para chegar antes das multidões em uma praia em que flamingos vagam livremente [...] eu queria algumas fotos sem pessoas as estragando”. A fotografia que “convenceu tanto o júri, quanto o público” é, na verdade, uma foto “tão real quanto o simples arranhar da barriga com que o pássaro está ocupado [na sua foto]” (Astray 2024). Para Astray (2024), a fotografia, que aparenta ter capturado na lente um flamingo sem cabeça, “era o candidato perfeito porque é uma foto surreal e quase inimaginável [*a surreal and almost unimaginable shot*], e ainda assim completamente natural”.

Em circunstâncias diferentes, e com opiniões distintas sobre o lugar da IA na arte³, nenhum dos dois artistas recebeu o prêmio. Eldagsen o recusou, advogando, entretanto, por mais diálogo sobre o tema na comunidade de fotografia, pedido que não foi atendido pela Sony (Eldagsen 2023b); já Astray foi desqualificado ao revelar a natureza da foto. Mas o que este exemplo de falha recíproca na discriminação do suporte da imagem indica com uma interessante simetria é o atual desafio em continuar a encarar imagens fotográficas e geradas por IA como possuindo naturezas absolutamente distintas.

²Disponível em: <https://www.milesastray.com/news/newsflash-reclaiming-the-brain> (acesso em: 23/10/2024).

³Enquanto que Eldagsen entende que é inevitável que “IA substituirá a fotografia” (Eldagsen 2023a), embora entenda que elas sejam de naturezas distintas e não devam competir entre si (Eldagsen 2023b), Astray (2024) diz que “submeteu esta foto real [*actual photo*] na categoria de IA do 1839 Awards para provar que conteúdo produzido por humano não perdeu sua relevância” e que “a Mãe Natureza e seus intérpretes humanos ainda podem vencer a máquina [*beat the machine*]”, pois “sua criatividade e emoções são muito mais que apenas uma sequência de dígitos [*string of digits*]”.

O caso do uso de IA nas artes nos ajuda, sem muita delonga, a passar ao largo de questões relativas à pretensa falseabilidade das imagens geradas por IA, permitindo, ao contrário, partir do pressuposto de sua aparente convergência com outros modos de criação de imagem – isto é, da premissa de que não há como diferenciar uma imagem gerada por IA e por uma câmera fotográfica apenas pelo resultado final produzido. Mas outros exemplos relativos à IA revelam ainda mais questões. Em tempos de negacionismos, chamam atenção as práticas relacionadas com o que tem sido chamado “desinformação”, ou seja, a criação de “notícias falsas [*fake news*]”, por exemplo a partir da geração de imagens e vídeos (e.g. “*deep fake*”). Uma forma comum em que isso ocorre é com a exposição de personalidades conhecidas em situações artificiais, como evidencia a relativa exposição que imagens geradas por IA do atual líder católico, Papa Francisco, receberam nas mídias⁴. De forma divergente, outras obras de arte que incluem a exploração da chamada “alucinação” em modelos de AdM ou a criação de vídeos que evoluem imagens e remetem a cenas oníricas apontam para um potencial que é menos reprodutor, no sentido de uma captura objetiva de imagens, e mais criativo desses modelos⁵. Um exemplo desse caso é o trabalho *Machine hallucinations: nature dreams* de Refik Anadol (2024).

Essa capacidade imagética, ou de representação, do AdM – a qual, mostraremos abaixo, não se limita à visualidade – é um aspecto que faz convergir tanto a literatura técnica, quanto a literatura crítica. O artigo “Deep learning [Aprendizado profundo]” escrito por LeCun, Bengio e Hinton (2015) e publicado na revista *Nature* que é citado na primeira epígrafe é considerado um marco consolidador do aprendizado profundo e das redes neurais profundas, principal vertente de aprendizagem de máquina contemporânea. Nele, tais algoritmos são considerados como “aprendizado de representação”, enfatizando ainda que a potência do aprendizado profundo está relacionado com sua capacidade de manter representações em múltiplos níveis para uma mesma entrada. Já o artigo-manifesto ou texto-mapa de Pasquinelli e Joler (2021) da segunda epígrafe é uma das descrições mais precisas do funcionamento da aprendizagem de máquina de uma perspectiva crítica. Nele, similarmente, o AdM é considerado um instrumento de representação, um “nooscópio” que opera como ampliação da percepção e do conhecimento sobre um “espaço de conhecimento”, atualmente disponível na forma de dados⁶. Dessa forma, considerando que a operação de representação, e todas as controvérsias relativas a esse termo, é central no AdM, analisaremos neste capítulo a aprendizagem de máquina do ponto de vista da operação de *imaginação*, isto é, de criação de imagens, de representação do mundo.

⁴Por exemplo: <https://www.nytimes.com/2023/04/08/technology/ai-photos-pope-francis.html> (acesso: 02/10/2024).

⁵Gisele Beiguelman e Refik Anadol são dois artistas que exploram usos alternativos da IA em suas produções visuais. Um exemplo da primeira é a exposição *Botannica Tirannica*, que esteve em cartaz no Museu Judaico de São Paulo de 28 de maio a 18 de setembro de 2022 (<https://museujudaicosp.org.br/exposicoes/botannica-tirannica-giselle-beiguelman/>, acesso: 29/10/2024), do segundo, um dos vários trabalhos é o *Unsupervised*, também parte da série *Machine hallucinations*, que esteve em cartaz entre novembro de 2022 e abril de 2023 no Museu de Arte Moderna de Nova York (MoMA, _Museum of Modern Art): <https://www.moma.org/calendar/exhibitions/5535> (acesso: 29/10/2024).

⁶Num diálogo com o conceito de “inconsciente óptico” em Benjamin (1993), Byung-Chul Han (2023:94) ainda sugere, em outra aproximação com instrumentos de percepção, que o AdM permitiria ampliar a visão sobre as ações humanas – “como uma lupa digital” – revelando parcelas de ação – “microações” – inconscientemente estruturadas – um “psicograma do inconsciente” (Han 2023:36) –: “[a]nalogamente ao ‘inconsciente óptico’, a inter-relação microfísica ou micropsíquica também poderia ser chamada *inconsciente digital*” (Han 2023:94).

4.1 Imagens: entre representação e performatividade

4.1.1 O caráter social das imagens

Quando Deleuze (1996) constata que “[t]odo atual rodeia-se de uma névoa de imagens virtuais”, não é de uma noção visual de imagem que ele se referia: todo atual tomado como objeto tem correlativamente a si um campo virtual que diz respeito a suas possibilidades de aceção e desenvolvimento. Essas imagens são reais, pois existem efetivamente, embora não sejam atuais, pois existem como que em potência, falta-lhes algo que as faça devir realidade atualizada. Mas, por vezes, essas imagens se atualizam de um modo peculiar, não com uma substituição dela pela atualidade do objeto que ocupa sua posição potencial, mas como imagem para outrem. É este o caso quando estendemos a mão intuitivamente ao achar que um copo de vidro vai cair da mesa: não é que o copo se atualizou naquilo que era, para o próprio objeto, sua imagem; mas sim que ele se tornou imagem para nós, em nossa percepção do objeto, no sujeito, se quisermos. Não faremos aqui uma diferenciação e caracterização de noções de imagem virtual e atual, mas gostaríamos apenas de notar um aspecto que se evidencia no exemplo da mão que vai em direção à imagem do copo: essa imagem não é visual, não vemos realmente o copo em outra posição, elas são de outra ordem. Quando reagimos intuitivamente à imagem de um copo que cai, não o imaginamos como visualidade, mas como algo totalmente distinto, que nos impele à ação. É imagem que modela para nós a posição do copo – e, portanto, tem um sentido de “imagem de mundo”, muito mais do que de “imagem de um objeto” – , isto é, é uma representação, mas não uma representação visual.

Em seu *Imagination et invention*, Simondon (2008) explora essa concepção de imagem que não está centrada no olho e na visão, mostrando processos de objetivação e subjetivação das imagens que nos ajudam a explicar a performatividade das imagens no AdM⁷. Ainda na introdução, partindo de uma pesquisa da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO, *United Nations Educational, Scientific and Cultural Organization*) sobre os estereótipos nacionais, o autor nota que “o francês, para o inglês, é um comedor de sapos e de lesmas”, enquanto que “os franceses, em certas regiões dos Estados Unidos, são percebidos como ‘Don Juan’ ” (Simondon 2008:11). Simondon constata que as imagens participam de processos de objetivação e de subjetivação que dão a elas uma verdadeira existência social: “Na verdade, essas imagens [os estereótipos] exprimem diferentes níveis da distância social [...] bem como atitudes e crenças” (Simondon 2008:11). Mais importante ainda é quando essas imagens “puramente mentais e subjetivas”, em um processo de “causalidade cumulativa”, se tornam uma “atitude real com estado social objetivo”:

é o que mostra Gunnar Myrdal na importante pesquisa sobre a situação dos negros nos Estados Unidos; Os preconceitos que empregadores ou proprietários brancos têm sobre as qualidades e falhas dos negros predeterminam a possibilidade ou a impossibilidade de tal ou tal conduta (por exemplo, profissões); por sua vez, e desta vez objetivamente, as escolhas profissionais predeterminam um certo método de criação dos filhos, um certo nível de educação e ideais definidos; depois de alguns ciclos de trocas recorrentes que vão da

⁷Neste livro, Simondon apresenta seu “ciclo genético das imagens”, separando-as em quatro fases progressivas e cíclicas: imagens motoras, imagens perceptivas, imagens afetivo-emotivas (ou imagens-símbolo) e, por meio da invenção, objetos-imagem.

imagem à realidade e da realidade à imagem por percepção, a imagem primitiva é realizada e encontra justificativa suficiente no estado social para se estabilizar (Simondon 2008:11)

Ou seja, os estereótipos, como imagens, agem socialmente, condicionando as possibilidades daquilo que eles visam representar⁸. Em seu argumento, qualquer elemento mediador entre sujeito e objeto, intermediário entre polos abstrato e concreto, pode ser entendido como imagem. “Com efeito, a imagem [...] não é somente mental: ela se materializa, devém instituição, produz riqueza, é difundida pela rede comercial tanto quanto a ‘*mass media*’ difunde a informação” (Simondon 2008:13). As imagens possuem uma “intensa capacidade de propagação”, “impregnam as civilizações e as carregam com sua força”, elas “exprimem os fatos sociais e econômicos [...] mas assim que são materializadas e objetivadas, também constituem um fardo e introduzem uma tensão que determina parcialmente o futuro social” (Simondon 2008:13). Desse modo, há tanto um movimento que vai do mental ao real objetivo, quanto do real objetivo ao mental. “Qualquer imagem é suscetível a incorporar-se a um processo de recorrência de materialização ou idealização; depositados na moda, na arte, nos monumentos, objetos técnicos, a imagem torna-se fonte de percepções complexas despertando movimento, representação cognitiva, afetos e emoções” (Simondon 2008:13).

Simondon fala também da importância da imagem na incorporação do passado e sua disponibilização para o trabalho prospectivo, de patente importância para nosso estudo: “Previsão no domínio coletivo (empresas, mesmo nações) corresponde à função de projetos e expectativas racionais de curto, médio ou longo prazo... Este esforço de racionalização coletiva da visão lançada sobre o futuro é uma das características do mundo contemporâneo” (Simondon 2008:16). Desse modo:

As necessidades de previsão de ação a longo prazo introduziu a dimensão do futuro e dissipou o mito, pelo menos em domínio econômico e demográfico; o tempo começa a se organizar como o espaço; *o futuro é anexado pelo saber, ou seja, não é mais o campo privilegiado do optativo, do desejo ou do querer*. E, no entanto, a imagem recupera a sua densidade e a sua força que a leva à antecipação do futuro coletivo, fora e acima das racionalizações prospectivas, que são principalmente extrapolações, mas não invenções reais (Simondon 2008:17, grifo nosso).

Essa imagem prospectiva, como força coletiva de racionalização que passa a organizar o tempo tanto quanto o espaço, anexando o futuro ao saber (em oposição ao optativo), não é para Simondon uma real antecipação, pois para isso falta-lhe a dimensão qualitativa que a ficção científica, por outro lado, soube incorporar (criando uma complementaridade inventiva ao trabalho de prospecção): “O que o que falta à previsão para ser uma verdadeira antecipação, é este poder qualitativo, esta *physis* que dá ao futuro a sua verdadeira dimensão como desenvolvimento contínuo. Pois prever, não é apenas uma questão de ver, mas de inventar e viver: a verdadeira previsão é, em certa medida, práxis, tendência ao desenvolvimento do ato já iniciado” (Simondon 2008:17).

Essa constatação reafirma a sugestão de Garcia dos Santos (1981) que mostra como obras de ficção científica do século XX (1984 de George Orwell, *Admirável mundo novo* de Aldous Huxley e sobretudo

⁸A citação não deixa de lembrar a ideia de “imagens de controle”, expressão utilizada por Patricia Hill Collins em relação ao racismo perpetrado por meio de estereótipos em relação aos negros (Collins 2016). Conferir também abaixo as seções sobre [Estereotipia](#) e [Racismo e sexismo algorítmico](#).

Nous autres de Eugène Zamiatine) anteciparam de maneira mais potente as problemáticas relativas ao progresso e à tecnologia do século seguinte que outras obras das ciências sociais.

Para nós, interessa-nos sobretudo sua caracterização não restritivamente visual de imagem, bem como sua descrição sobre o caráter social que advém de processos de objetivação e subjetivação de imagem. Principalmente na medida em que as imagens criadas por algoritmos de AdM podem constituir processos de causalidade cumulativa que condicionam os desenvolvimentos coletivos.

4.1.2 As imagens técnicas

Já no que diz respeito à intersecção das teorias da imagem com a tecnologia, um debate central é aquele travado em torno das “imagens técnicas”. Para Vilém Flusser (2018), filósofo tcheco-brasileiro, imagem técnica é “imagem produzida por aparelho”⁹. Desse ponto de vista, a fotografia pode ser entendida como um marco paradigmático da contemporaneidade. O “primeiro aparelho” e “primeiro objeto pós-industrial”, pois seu valor não está em si como objeto ou em sua posse, mas na informação que ela suscita, o que depende de sua “programação” (Flusser 2018:21, 64 *et seq.*). Para o autor, do mesmo modo que a invenção da escrita marcou a passagem de uma sociedade focada em imagens (que representavam o real), para uma focada no texto; a fotografia – entendida como produto de um texto científico aplicado que só poderia ser gestado em uma sociedade texto-centrada – marca a passagem da escrita (que representa imagens), para um novo tipo de imagens. Essas tecnoimagens, por sua vez, representam textos e constituem portanto um pensamento conceitual (Flusser 2018, *sem data*).

Para Flusser (2018), a “imagem tradicional” seria uma abstração do fenômeno concreto, ou a cristalização de uma ideia para uma superfície bidimensional. Diferentemente, a “imagem técnica” é uma reconstituição bidimensional de textos: “transcodificam”, na superfície, um pensamento conceitual, ainda mais abstrato em relação às chamadas “imagens tradicionais”. Há aqui não só uma diferença histórica – a imagem antes e depois do advento do texto –, mas também e principalmente de natureza, de ontologia¹⁰.

Embora divirjamos da rígida linearidade do esquema proposto por Flusser, em que as imagens técnicas marcariam uma ruptura absoluta em relação às imagens tradicionais e ao texto, consideramos as ideias de “aparelho” e “programa” bastante interessantes e pertinentes para nossa discussão. Para Rodrigo Petronio, “Flusser descrevia a gênese dos algoritmos ao descrever os aparelhos” (Flusser 2018:8). É que, como “caixa-preta”, do “complexo aparelho-operador” ou da relação câmera-fotógrafa só se pode ver a entrada e a saída, “*input e output*” (Flusser 2018:23). O ato fotográfico como operação “mágica” é

⁹Sendo, ainda, “imagem” uma “superfície significativa na qual as ideias se interrelacionam magicamente”, isto é, sincronicamente, e aparelho um “brinquedo”, no sentido de que é algo que interage em um jogo, “que simula um tipo de pensamento” (Flusser 2018).

¹⁰Historicamente, as imagens tradicionais precedem os textos, por milhares de anos, e as imagens técnicas sucedem os textos altamente evoluídos. Ontologicamente, a imagem tradicional é abstração de primeiro grau: abstrai duas dimensões do fenômeno concreto; a imagem técnica é abstração de terceiro grau: abstrai uma das dimensões da imagem tradicional para resultar em textos (abstração de segundo grau); depois, reconstituem a dimensão abstraída, a fim de resultar novamente em imagem. Historicamente, as imagens tradicionais são pré-históricas; as imagens técnicas são pós-históricas. Ontologicamente, as imagens tradicionais imaginam o mundo; as imagens técnicas imaginam textos que concebem imagens que imaginam o mundo” (Flusser 2018:21). Sobre as diferenças históricas e ontológicas entre “imagens tradicionais”, “texto” e “imagens técnicas” no autor, conferir também Flusser (*sem data*).

levado a cabo por seu programa¹¹. Estar programado é o que caracteriza o aparelho, sendo os programas, tal qual na máquina de Turing, “compostos de símbolos permutáveis” (Flusser 2018:34, 36). É por isso que embora a câmera esteja “programada para produzir fotografias”, seguindo ainda o programa da indústria fotográfica, do aparelho socioeconômico, político-cultural, etc., ela não é totalmente automatizada, há sempre abertura, margem de indeterminação¹²: “O aparelho fotográfico funciona em função dos interesses da fábrica, e esta, em função dos interesses do parque industrial. E assim *ad infinitum*. [...] O decisivo em relação aos aparelhos não é quem os possui, mas quem esgota o seu programa” (Flusser 2018:39). Assim, “[o] fotógrafo exerce poder sobre quem vê suas fotografias, programando os receptores”, mas “[o] aparelho fotográfico exerce poder sobre o fotógrafo” (Flusser 2018:40). Há uma composição hierarquizada de agências que explica o funcionamento da câmera.

Nos chama atenção não apenas a proximidade estabelecida pelo autor entre a fotografia e conceitos utilizados em caracterizações do computador digital – por exemplo, nos usos de termos como “programa” –, mas também as vias por quais sua agência é caracterizada, nos remetendo à ideia la-touriana de “ação-rede” [*actor-network*], principalmente no que diz respeito à ação que é sempre ação conjunta de fotógrafa-câmera, ou fotógrafa-câmera-indústria-etc.

Já Arlindo Machado (2004), partindo de intuições semelhantes – mas superando o esquema evolucionista e a ideia de uma ruptura absoluta entre as imagens tradicionais e as imagens técnicas em Flusser –, argumenta que, não existindo no humano um modo direto de expressar suas imagens internas, toda imagem seria técnica, pois “toda imagem materializada em algum tipo de suporte é o resultado da aplicação de algum tipo de técnica de representação pictórica” (Machado 2004:223). Assim, levando-se em conta que há sempre um sistema de pensamento que informa a artista ou instrumentos pictográficos que mediam sua prática, pode-se dizer que *toda imagem é técnica*. Apesar disso, a fotografia e o computador digital representem momentos importantes de automatização do gesto criador de imagem, em que a mediação técnica é indubitável, e portanto ele propõe, “apenas para fins operativos”, uma definição mais estrita de imagem técnica como aquelas em que a mediação técnica “produz uma diferença no universo das imagens, marcando muitas vezes uma cisão, uma distância em relação às imagens do homem, às suas imagens interiores” (Machado 2004:224). Já não há mais uma mudança qualitativa que promove uma ruptura entre imagens pré-técnicas e técnicas, mas sim uma diferença de grau que confere às imagens geradas por câmera fotográfica e pelo computador uma maior distância em relação às imagens subjetivas, o que faz com que elas ganhem estatuto de exterioridade e objetividade¹³.

¹¹“A magia pré-histórica ritualiza determinados modelos, mitos. A magia atual ritualiza outro tipo de modelo: *programas*” (Flusser 2018:25).

¹²“Enquanto *objeto duro* [hardware], o aparelho fotográfico foi programado para produzir automaticamente fotografia; enquanto *coisa mole* [software], impalpável, foi programado para permitir ao fotógrafo produzir fotografias deliberadas automaticamente. São dois programas que se coimplicam. Por trás destes, há outros. O da fábrica de aparelhos fotográficos: aparelho programado para programar aparelhos. O do parque industrial: aparelho programado para programar indústrias de aparelhos fotográficos e outros. O socioeconômico: aparelho programado para programar o aparelho industrial, comercial e administrativo. O político-cultural: aparelho programado para programar aparelhos econômicos, culturais, ideológicos e outros. Não pode haver um ‘último’ aparelho, nem um ‘programa de todos os programas’. Isso porque todo programa exige metaprograma para ser programado. A hierarquia dos programas está aberta para cima” (Flusser 2018:38).

¹³Para Latour (2017), o adjetivo “técnico” estaria ligado a uma delegação de ação. Nesse sentido, uma mediação técnica seria aquela em que um elemento não humano mediador congrega ações de diferentes espaços, tempos e agências, concretizando em si um novo “programa de ação” (ver capítulo 7). Justamente por isso, quanto mais técnico é algo, mais social ele o é: entendemos que é daí que surge a defasagem em relação à imagem interna individual. Nesse ponto, acreditamos haver uma convergência possível entre o sistema flusseriano exposto acima, a descrição machadiana e a conceitualização de Latour.

Em relação às imagens digitais contemporâneas, novas questões se colocam. Para Giselle Beiguelman (2021), se baseando em Flusser (2008), “as imagens técnicas são as imagens eletrônicas e digitais”, que, “[p]or serem constituídas por pixels, são ‘zerodimensionais’, apesar de o suporte a que aderem ser bidimensional ou tridimensional”. Portanto, “[o] desafio que impõem é o de transcender o ‘nível ontológico das imagens tradicionais’, indo do ‘abstrato ao concreto’” (Beiguelman 2021:17, n. 10). Por outro lado, Cesar Baio (2022) também traz contribuições interessantes à partir da análise das imagens técnicas contemporâneas: digitais e algoritmicamente medidas. Este, seguindo mais os passos de Machado, recusa uma singularidade ontológica das tecnoimagens, ou melhor, a expande:

A crítica ao modelo normalizador antropocêntrico conduz à discussão sobre a materialidade dos processos de produção de imagem como base para a hipótese de uma pluriontologia da imagem, segundo a qual, a ação artística é compreendida em sua potência de produção de novas maneiras de conceber a imagem. O estabelecimento dessa hipótese se fundamenta na discussão do papel da materialidade dos algoritmos nas imagens, considerando seus processos de produção, atualização e circulação. Tal hipótese sugere que o reconhecimento da materialidade dos processos de produção e circulação da imagem é parte fundamental para a compreensão do aspecto performativo que a imagem assume na cultura contemporânea, destacando a necessidade de pesquisas mais aprofundadas nesta direção (Baio 2022:83).

Partindo de Flusser e Machado, mas em diálogo com autoras como Haraway (2023b), e focando no que ele chamou de “arte dos aparatos” na produção artística contemporânea, constata que “cada obra de arte passa a ser a materialização de uma maneira singular de conceber a imagem, uma projeção especulativa de realidades, mundos e sujeitos [... estabelecendo], assim, um conceito sobre o mundo, algo que se torna parte da poética do artista” (Baio 2022:90). Isto é, em oposição a uma diferença ontológica clássica entre imagens tradicionais e técnicas, o autor propõe que cada obra tem sua própria ontologia, ou ao menos um processo ontogenético singular, plasmando certas formas específicas de ver que são, simultaneamente, representação e ação sobre o mundo¹⁴. Isso leva a uma constatação de que toda imagem possui, para além de seu aspecto representativo mais evidente, uma performatividade. E especificamente em relação às imagens algorítmicas: “Analisados desta perspectiva” que aproxima a noção de algoritmo da computação ao de programa em Flusser, “os processos de produção de imagens acontecem por meio de algoritmos computacionais, de software e hardware, que cristalizam nas máquinas de imagem práticas culturais que estruturam as formas de ver e existir dadas pelos fluxos industriais, pela cultura visual ou pelas estruturas que organizam as relações de poder na sociedade (Baio 2022:96). E ainda:

A mediação digital amplifica os tipos de dados captados – de sensores que leem dados diversos dos corpos às práticas de *profiling* nas redes digitais –, aumentam a capacidade de processamento de informações com tecnologias como o aprendizado de máquinas – inteligência artificial –, enquanto entrecruzam informações de tempos e espaços diversos por

¹⁴O autor ainda constata que: “A hipótese da pluriontologia não necessariamente exclui o modelo moderno de representação, mas o incorpora como mais um modo de conceber a imagem dentre uma quantidade inumerável de possibilidades. Cada artista, em cada obra de arte, pode criar uma nova maneira de conceber a imagem, de modo que esta possa cristalizar modelos epistemológicos ou mesmo cosmológicos de mundo completamente diferentes” (Baio 2022:90).

meio das múltiplas redes pervasivas de computação. Nesse contexto, compreender adequadamente a cultura visual emergente, cada vez mais, requer a análise do aspecto performático da imagem (Baio 2022:97).

Assim, o debate sobre as imagens técnicas está fortemente entrelaçado com uma investigação sobre as agências, tanto do aparato que cria a imagem, quanto da imagem propriamente dita. Partiremos dessa perspectiva que entende que toda representação, por ser operação situada em uma temporalidade, espacialidade e ponto de vista específico, é também uma ação: pois ela cria imagens e formas de ver que participam ativamente da nossa vida social.

4.1.3 Imagens operativas, pós-fotográficas e máquina-para-máquina

Partindo da premissa esboçada em Baio (2022) de que toda imagem é tanto representação, quanto performatividade, acreditamos que a perspectiva gestada na obra do cineasta Harun Farocki em relação às imagens operativas e pós-fotográficas adicionam elementos importantes para a caracterização de sua ação, que podem ser estendidas para explicar a operatividade presente nas imagens não visuais e de tipo máquina-para-máquina (m2m, *machine-to-machine*) do AdM.

Desde a década de 1980, Farocki se propôs a investigar as câmeras eletrônicas e as imagens produzidas por elas, que “conferiam algum tipo de ação bastante específica” e “que tinham o traço comum de dispensar a presença humana na operação do dispositivo” (Genaro 2015:135). Segundo Genaro (2015), o documentário *Reconhecer e perseguir* (2003) é sua obra de maior vitalidade sobre as imagens operativas, que são definidas no ensaio “Influências transversais” (2002) publicado na revista *Trafic* como imagens que partem de sua própria operação técnica para constituição de sua realidade, “imagens subtraídas a um operativismo em si mesmas, dispensando ou eliminando qualquer elo que ativa um sentido de transindividuação com outras” (Genaro 2015:135). Por isso, são imagens que saem do campo estético e de lazer, para integrarem o campo da comunicação e da política. Nesse sentido, a obra do cineasta explora “distintas imagens operativas, estabelecidas como dispositivos de controle temporal e espacial na produção/destruição em fábricas, prisões, ruas, exércitos etc.” (Genaro 2015:135).

Já as imagens pós-fotográficas se referem àquelas surgidas com a tecnologia digital eletrônica, em que não há um referente separado no mundo, mas são sintetizadas por computador, feitas por simulação, produto de cálculos e de operações eletrônico-digitais (Genaro 2015). Considerando que este tipo de imagem seria “virtual” – embora entendamos que elas são tão atuais quanto as imagens produzidas por emulsão fotoquímica, sendo mais corretamente chamadas de “eletrônicas” ou “digitais” (cf. capítulo 7) – Genaro (2015) mostra como os filmes de Farocki as exploram de forma heterogênea: do ponto de vista da “apresentação do real a partir do virtual” – *Serious Games* (2009-10) –, da “interpretação do real a partir do virtual” – *O que há?* (1991) –, do “prolongamento do real no virtual” – *Criadores do mundo de consumo* (2001) –, da “injeção do real no virtual” – *Imagens do mundo e inscrições de guerra* (1988), da “visualização do virtual por uma janela real” – *Paralelo* (2012-4) – e da “tele-presença do real no virtual” – *Reconhecer e Perseguir* (2003) – (Genaro 2015:136).

Nos interessa como na visão do cineasta “aquilo que uma imagem é ou produz depende das mediações ou da ausência destas no ambiente em que elas emergem” (Genaro 2015:137). Diríamos mais: a própria imagem operativa é, no sentido de Latour (2017), uma mediação técnica. Desse modo, em *Imagens da prisão* (2000), por exemplo, a substituição da “intervenção humana direta” pela máquina explicita um deslocamento da conhecida vigilância centralizada para uma vigilância distribuída. Dessa

forma, as câmeras-armas¹⁵ utilizadas na prisão e apresentadas no filme são paradigmáticas¹⁶, pois “com as imagens produzidas por estas câmeras, o espectador se reduz a um inspetor ou policial, interessando-se apenas em encontrar os ‘desvios’ ocorridos nelas”, o próprio cineasta, citado por Genaro, constata que “[o] que é interessante nas imagens de câmeras de vigilância é que elas são usadas de um modo puramente indicial, não se referem a impressões visuais, mas apenas a certos fatos: o carro ainda estava no estacionamento às 14h23min? [etc.]” (Genaro 2015:142). Mas essa automatização da visão-ação pode ir ainda mais longe: “as cintas eletrônicas no corpo dos prisioneiros [...] mudam completamente o assunto: elas dispensam a própria arquitetura da prisão” (Genaro 2015:142). O mapeamento da localização de prisioneiras que portam uma tornozeleira eletrônica, ou a simulação de consumidoras de um supermercado são exemplos de imagens pós-fotográficas apresentadas no filme.

Para Genaro (2015), se baseando sobretudo no documentário *Reconhecer e perseguir*, que apresenta uma arqueologia da imagem operativa a partir do míssil teleguiado, são duas as características da imagem operativa: *manter-se à distância e imitar-suplantar o olho humano*. No contexto do míssil, ela serve de “controle de qualidade da ação à distância do míssil”, é “uma imagem-informação, sem nenhuma intencionalidade estética” (Genaro 2015:145). Entretanto, sua operatividade não se restringe à guerra: “[m]anter-se à distância, a partir de imagens que objetivam acompanhar um míssil, é, evidentemente, um assunto de alto interesse militar”, mas, também, “[c]âmeras permitem manter o trabalho à distância do calor e do ruído” em uma empresa de laminação (Genaro 2015:145). Neste caso, “[s]ão câmeras que servem simultaneamente para visualizar e controlar à distância diferentes ângulos e ambientes de laminação”. Já em relação ao segundo aspecto, Genaro (2015:145) nota como “[e]m determinado caso, o olho humano é incapaz de ver certos defeitos na produção e, assim, uma imagem em diferentes tons de vermelho pode ressaltar a conformidade ou inconformidade do material laminado”. Raios-X, coloração artificial, zooms, filtros de imagens e pós-processamento computadorizado são formas de partir da visão humana, superando-a operacionalmente.

Nesse mesmo sentido, apontamos que há hoje uma tendência que parte das características das imagens operativas e pós-fotográficas, mas que deixam de ser visuais e passam cada vez mais a operar em fluxos m2m¹⁷. Desse modo, imagens biométricas captadas por sensores de impressão digital ou de

¹⁵Interessante notar como essa analogia da câmera com a arma não é recente: Flusser (2018:43) constatara que “[q]uem observar os movimentos de um fotógrafo munido de um aparelho (ou de um aparelho munido de fotógrafo) está observando movimento de caça”. O cineasta Jean-Luc Godard também sugere a aproximação da câmera de fotografia e da arma de fogo em seu *Les carabiniers [Os fuzileiros]* (1963), que termina mostrando as imagens “capturadas” pelo fuzileiro após retornar das batalhas. Até mesmo os slogans da Kodak, “*You Press the Button, We Do the Rest* [você aperta o botão, nós fazemos o resto]” e “*Clic, clac, merci Kodak* [Clic, clac, obrigado Kodak]” não escondem a similaridade com a automatização da violência (cf. Latour 2011, capítulo três). É importante dizer, entretanto, que esta não é a única apresentação possível da agência da câmera: a cineasta Agnès Varda em seu *Les glaneurs et la glaneuse [Os catadores e a catadora]* (2000), traduzido em português como “Os catadores e eu”, aproxima o gesto fotográfico à “catação”, em uma oposição interessante à “captura” que descrevemos no capítulo 3.

¹⁶Como comenta Genaro (2015:142), “[a]o final [do documentário], assistimos as cenas do caso de uma briga de gangue dentro do presídio de Corcoran (Califórnia), que provocou a morte do prisioneiro Martinez. Do centro de vigilância da prisão, a polícia disparou a arma de fogo acoplada à câmera de vigilância do pátio, acertando letalmente o prisioneiro”.

¹⁷“A visão humana deixa de ser substancial e passa a ser acidental”, comenta Garcia dos Santos em relação ao livro *La machine de vision* de Paul Virilio: “Não há como resumir aqui o mapeamento impressionante que Virilio faz do processo de transferência do olhar ao longo dos séculos XIX e XX até que a ‘máquina de visão’, capaz de realizar não só um reconhecimento das formas, mas principalmente uma interpretação completa do campo visual, inaugura a possibilidade de ‘uma visão sem olhar’, na qual uma câmera de vídeo acoplada a um computador assumirá a função de analisador do ambiente e, com ela, a capacidade de interpretar automaticamente o sentido dos acontecimentos, tanto no campo da produção industrial quanto da

íris nunca são vistas por uma operadora humana, sendo exclusivamente processada por uma máquina ou sistema algorítmico. No filme *Imagens da prisão*, a simulação de consumidoras de um mercado e a posição das detentas portanto torneleira eletrônica se tornam imagens nas telas, mas isso não é completamente necessário. Hoje em dia essas imagens (de consumidoras, de detentas) podem alimentar outros softwares que vão gerir produtos ou pessoas algorítmicamente. Igualmente, fluxos de dados como sinais de GPS e IPs geram mapas informacionais que não são visuais, mas que podem servir de imagem no interior de aplicações como o Google Maps ou Uber. E assim por diante. Durante nossa investigação, constatamos que a maioria das imagens mobilizadas pelo AdM são desse tipo, embora algumas das patentes mencionem explicitamente imagem da câmera ou da tela de um smartphone. Desse modo, partindo de uma perspectiva que entende imagens como representações intimamente ligadas com a realidade social – seja por processos de objetificação e subjetivização; seja por sua performatividade que cria formas de ver; seja por sua operacionalidade material – e que não são necessariamente visuais, podemos olhar para os algoritmos de AdM com foco em sua capacidade de representação, entendendo nisso a criação de uma imagem.

4.2 O filo maquínico e os agenciamentos de representação: uma genealogia do aprendizado de máquina

Em diálogo com ideias precedentes, como a noção de “tendência tecnológica” do arqueólogo e antropólogo André Leroi-Gourhan e a de “linhagem tecnológica” do filósofo Gilbert Simondon, em *Mil platôs*, Deleuze e Guattari propõem explicar a relação entre tecnologia e sociedade pelo conceito de “filo maquínico”¹⁸. Para Deleuze e Guattari (2011–2012, v. 5, p. 93), “[é] possível falar de um *phylum maquínico*, ou de uma linhagem tecnológica, a cada vez que se depara com um *conjunto de singularidades, prolongáveis por operações, que convergem e as fazem convergir para um ou vários traços de expressão assinaláveis*”. Isto é, singularidades abstratas ou virtuais que se concretizam e se atualizam por meio de operações convergentes. Ou ainda, um maquinismo virtual que devém mecanismo(s) atual(is) (Ferreira 2004). No caso de uma divergência de operação, há, simetricamente, divergência do filo, como no caso punhal-espada de ferro e faca-sabre de aço (Deleuze e Guattari 2011–2012, v. 5, p. 93). Mas, não é só o material (ferro e aço), nem a forma (um ou dois gumes) que definem o filo, “[c]ada *phylum* tem suas singularidades e operações, suas qualidades e traços, que determinam a relação do desejo com o elemento técnico (os afectos ‘do’ sabre não são os mesmos que os da espada)” (Deleuze e Guattari 2011–2012, v. 5, p. 93-4).

Como apontado por Ferreira (2004:5), o termo “filo” remete tanto ao desejo, como em sua aceção como “amigo”/“amante” (e.g. “filosofia”), quanto a “linhagem”/“família” como em “filogênese”: “Atração, simpatia, afinidade, o filo maquínico é a linhagem e o desejo dos maquinismos”. Desse modo, o filo maquínico pode ser visto como um contínuo, mas um contínuo de variações que, portanto, se diferencia constantemente:

sempre é possível instalar-se no nível de singularidades prolongáveis de um *phylum* a outro, e reunir ambos. No limite, não há senão uma única e mesma linhagem filogenética, um único e mesmo *phylum* maquínico, idealmente contínuo: o fluxo de matéria-movimento,

robótica militar” (Garcia dos Santos 2003b:175). Conferir ainda a discussão sobre captura e vigilância no capítulo 3.

¹⁸Seguindo Ferreira (2004), optamos por utilizar a grafia “filo” em “filo maquínico”, embora tanto o texto original “*phylum* machinique”, quanto a tradução brasileira “*phylum* maquínico”, utilizem a grafia “*phylum*”, tal qual utilizada em francês.

fluxo de matéria em variação contínua, portador de singularidades e traços de expressão. Esse fluxo operatório e expressivo é tanto natural como artificial: é como a unidade do homem com a Natureza. Mas, ao mesmo tempo, não se realiza aqui e agora sem dividir-se, diferenciar-se (Deleuze e Guattari 2011–2012, v. 5, p. 94).

É por isso que, paralelo ao filo, contínuo e real, mas abstrato e virtual, é preciso definir suas formas de atualização concretas. É isso o que Deleuze e Guattari chamam de agenciamento. “Denominaremos *agenciamento* todo conjunto de singularidades e de traços extraídos do fluxo — selecionados, organizados, estratificados — de maneira a convergir (consistência) artificialmente e naturalmente: um agenciamento, nesse sentido, é uma verdadeira invenção” (Deleuze e Guattari 2011–2012, v. 5, p. 94). Uma espécie de realismo construtivista centrado na atividade inventiva. “Os agenciamentos recortam o *phylum* em linhagens diferenciadas distintas e, ao mesmo tempo, o *phylum* maquinico os atravessa todos, abandona um deles para continuar num outro, ou faz com que coexistam” (Deleuze e Guattari 2011–2012, v. 5, p. 94).

Como acontece com outros conceitos deste “autor dual”, para falar com Viveiros de Castro (2018), não há relação opositiva entre filo e agenciamentos, mas recorrências e retroalimentações, processos de atualização que vão do filo ao agenciamento e processos de virtualização deste em relação àquele: “Há, com efeito, um *phylum* maquinico em variação que cria os agenciamentos técnicos, ao passo que os agenciamentos inventam os *phylums* variáveis. Uma linhagem tecnológica muda muito, segundo seja traçada no *phylum* ou inscrita nos agenciamentos; mas os dois são inseparáveis” (Deleuze e Guattari 2011–2012, v. 5, p. 95).

É importante enfatizar, aqui, que o adjetivo “maquinico” não exclui a dimensão humana do filo, mas, ao contrário, remete à noção operativa de *máquina* presente na obra dos autores. Pouco depois de escreverem *O Anti-édipo* (Deleuze e Guattari 2011b), Deleuze e Guattari elaboram-na da seguinte maneira:

Já não se trata de confrontar o homem e a máquina para avaliar as correspondências, os prolongamentos, as substituições possíveis ou impossíveis entre ambos, mas de levá-los a comunicar entre si para mostrar como o homem compõe peça *com a máquina*, ou compõe peça com outra coisa para constituir uma máquina. A outra coisa pode ser uma ferramenta, ou mesmo um animal, ou outros homens. Portanto, não é por metáfora que falamos de máquina: o homem *compõe máquina* desde que esse caráter seja comunicado por recorrência ao conjunto de que ele faz parte em condições bem determinadas. O conjunto homem-cavalo-arco forma uma máquina guerreira nômade nas condições da estepe. Os homens formam uma máquina de trabalho nas condições burocráticas dos grandes impérios. O soldado de infantaria grego compõe máquina com suas armas nas condições da falange. O dançarino compõe máquina com a pista nas condições perigosas do amor e da morte... Não foi de um emprego metafórico da palavra máquina que partimos, mas de uma hipótese (confusa) sobre a origem: a maneira como elementos quaisquer são determinados a compor máquinas *por recorrência e comunicação*; a existência de um “*phylum* maquinico” (Deleuze e Guattari 2011a:508–10).

Essa noção de máquina, que desde *Kafka* (Deleuze e Guattari 2018) se desdobra na de agenciamento, enquadra relações de funcionamento entre heterogêneos traçadas anteriormente ao corte su-

jeito/objeto ou natureza/cultura, observando, antes, sua realidade híbrida (cf. [Latour 2019](#)). Em *O anti-édipo*, encontramos máquinas desejanças, técnicas e sociais. Se as primeiras visam ser suporte de um conceito materialista e coletivo de desejo, visando dar conta sobretudo de funcionamentos e movimentos inconscientes e moleculares, as segundas se relacionam com funcionamentos tecnológicos e as últimas, sociais. Fazendo-as convergir, dizemos que um agenciamento seria a “unidade real mínima”, e também “co-funcionamento”, “simbiose” e “simpatia” entre entidades heterogêneas ([Deleuze e Parnet 1998](#)). Os agenciamentos podem ainda se bifurcar em duas faces: *agenciamentos coletivos de enunciação*, sistemas semióticos com regimes de signos que enfatizam os enunciados (ou a expressão); e *agenciamentos maquínicos*, sistemas pragmáticos que mobilizam ações e paixões e que enfatizam o estado de coisas (ou o conteúdo). Desse modo, emprestando a expressão de Haraway ([2023a](#)), o conceito de agenciamento desemboca em uma perspectiva material-semiótica de descrição da realidade. É nesse contexto que propomos enquadrar a *representação* como uma tendência técnica específica no *filó*, e o AdM como um *agenciamento* correspondente.

Desse modo, gostaríamos de propor não uma história definitiva do aprendizado de máquina como máquina de representação, mas uma genealogia possível que traça uma relação de continuidade no *filó* maquínico, que pode ser tomada em composição com outras propostas que descrevem a história da IA como, por exemplo, a da passagem de máquinas dedutivas para máquinas indutivas ([Cardon et al. 2018](#)) ou ainda do AdM como uma técnica de auto-organização da informação, legatária da organização do trabalho coletivo no capitalismo incipiente ([Pasquinelli 2023](#)). Nossa proposta segue a pista de Pasquinelli e Joler ([2021](#)), que entendem o AdM como “*um instrumento de ampliação do conhecimento que ajuda a perceber características, padrões e correlações através de vastos espaços de dados para além do alcance humano*” ([Pasquinelli e Joler 2021:1263](#)). Como ressaltado na epígrafe que abre o capítulo, tal qual o microscópio ou o telescópio inauguraram novas formas de ver e, portanto, de representação científica do mundo, o AdM seria um “nooscópio”, um “instrumento para navegar no espaço do conhecimento”. Desse modo, ao lado da perspectiva renascentista e da fotografia moderna, buscamos propor uma história do AdM que o enquadre como um agenciamento de representação no seio dos debates sobre as imagens.

4.2.1 Renascimento

Seguindo Machado ([2004:226](#)), podemos considerar o livro *Da pintura* de Leon Battista Alberti ([2014](#)), publicado originalmente em 1443, não só como o primeiro registro sistematizado sobre a pintura, como indica Leon Kossovitch em seu prefácio da obra de Alberti, mas também “a mais antiga reflexão que se conhece sobre a imagem técnica”. Nesse sentido, apresentaremos a descrição contida nesse tratado sobre a imagem renascentista – que congrega os ideais de objetividade e de antropocentrismo, de naturalismo e de figurativismo – como ponto de partida de nossa história do AdM como um agenciamento de representação.

No ensaio, Alberti propõe discorrer sobre a pintura “não como matemático, mas como pintor”, embora à pintora, “antes de tudo[, lhe] deseje que saiba geometria”, pois “quem não conhecer geometria não entenderá nem [suas] regras nem regra alguma de pintura” ([Alberti 2014:71, 128](#)). O texto, que é dividido em três livros, se inicia com os “rudimentos” que visam apresentar conceitos básicos de óptica e, principalmente, de matemática, derivados da geometria euclidiana e de releituras árabes medievais, como explicitado nas notas de rodapé do livro ([Alberti 2014](#)). O foco desta parte é apresentar a “pirâmide

visual”, o centro da pintura renascentista¹⁹ e do que hoje entendemos como perspectiva linear. Convém notar que, na Renascença, chamava-se *perspectiva artificialis* o método de construção de imagens gráficas em oposição à *perspectiva naturalis* dedicada ao estudo da visão humana (Domingues 2020). Desse modo, Alberti explica que “[a] pirâmide é a figura de um corpo no qual todas as linhas retas que partem da base terminam em um único ponto. A base dessa pirâmide é uma superfície que se vê. Os lados da pirâmide são aqueles raios que chamei extrínsecos. O vértice, isto é, a ponta da pirâmide, está dentro do olho, onde está o ângulo das quantidades” (Alberti 2014:77–78). Como comenta Machado (2004), a ideia dessa pirâmide é que o prolongamento de nosso olho até o objeto representado se interseccionaria em um plano (“um vidro”) intermediário imaginariamente posicionado entre nós e o objeto, prefigurando a imagem a ser reproduzida. Desse modo, explicita-se que se trata de uma técnica que visa ativamente ilusionar a natureza, como constata Ana Clara Domingues (2020).

Esse ideal de mimese da natureza, simbolizado pela coincidência do termo perspectiva para a visão humana e para a pintura, é apresentado extensivamente no *Da pintura*. Por um lado, destaca-se que a natureza deve ser exaltada como bela e tomada como objeto e inspiração da pintura (Alberti 2014:100, 108, 130, 134). Por outro lado, mostra-se como a natureza também deve influenciar e determinar as próprias técnicas pictóricas (Alberti 2014:101, 104, 124, 130, 133). Desse modo, “imitação da natureza”, com o estatuto de objetividade advindo dela, e “beleza” entendida como “configuração ideal”, são consideradas por Machado (2004:226) como as “duas máximas da iconografia renascentista”. Sobre a aparente contradição dessa díade, ele comenta:

A imagem obtida por intermédio de tal dispositivo [a pirâmide visual e sua concretização em diversos instrumentos] era a seguir “corrigida” por uma aplicação do código da perspectiva *artificialis*, cuja função básica era sugerir uma ilusão de profundidade sobre a tela plana, mas pressupunha também toda uma ideologia da objetividade e da verossimilhança decorrentes de sua base “científica”. De fato, a perspectiva renascentista [...] era encarada pelo homem do *Quattrocento* como um sistema de representação plástica baseado nas leis “objetivas” do espaço formuladas pela geometria euclidiana e, como decorrência de tal fato, acreditava-se que ela deveria nos dar a imagem mais justa e fiel da realidade visível (Machado 2004:225).

“Natureza e beleza não se opõem: como eleita, a beleza é natural, e a natureza, bela” constata Kossovitch (Alberti 2014:64), comentando que a “categoria fundamental do pensamento pictórico [renascentista] é o ‘tirar’”, como no “tirar do natural” que é, como continua Kossovitch, “retratar”, “representar”: “devemos tirar da natureza o que queremos pintar e sempre escolher as coisas mais belas” (Alberti 2014:133). É que a representação é ato negativo: representa-se um referente, justamente porque ele não está lá, diferentemente da performance, que é ato positivo de quem performa. Entretanto, à revelia do ideal “especular” da pintura renascentista, Domingues (2020:52) argumenta que “tal técnica desencadeou uma série de produções artísticas que modificou os modos de perceber e enxergar o mundo por produzir um efeito muito próximo daquilo visto na natureza”, mas de tal maneira que “a perspectiva atua nas imagens de modo que o que vemos não é natural, e sim fruto de uma série de intervenções que os artistas elaboraram para ilusionarem os espectadores”. Esse esforço estabelece uma

¹⁹Alberti comenta que: “não será, pois, a pintura outra coisa que a intersecção da pirâmide visual representada com arte por linhas e cores numa dada superfície, de acordo com uma certa distância e posição do centro e o estabelecimento de luzes” (Alberti 2014:83).

“cultura visual ocidental” ou moderna²⁰ que, embora contestada em momentos como as vanguardas do século XX, continua a existir influenciando a fotografia e o cinema (Domingues 2020:53). No mesmo sentido, Maria Salviano (2023), analisando o “estatuto da imagem na arte e na tecnologia”, apresenta a perspectiva renascentista como uma tentativa de sistematizar uma visão de uma “janela para o mundo”, se amparando em Hans Belting²¹ para apontar o caráter antropocêntrico da imagem renascentista:

Por muito tempo, entendeu-se este processo como uma forma de reproduzir objetivamente o modo como víamos o mundo, uma maneira de construir o olhar que seria válida para qualquer pessoa que fosse dotada de visão. Belting (2012), porém, demonstra como o tipo de imagem feita a partir do cálculo da perspectiva no Renascimento não era universal, mas fortemente localizado. De acordo com o autor, ainda que a técnica matemática que inspirou o cálculo fosse de origem árabe, a imagem em perspectiva do Renascimento era algo completamente inconcebível de ser feito na cultura vizinha. E isso não se dava devido a uma falta de condições técnicas, mas pelo fato de que as imagens árabes tinham um estatuto ontológico completamente diferente do europeu, não podendo ser entendidas a partir de uma associação direta com a visão (Salviano 2023:2–3).

Enfatizamos que o caráter antropocêntrico da imagem renascentista está ligado com o privilégio do modo de percepção humano: na imagem renascentista o olho humano é tido como ideal de representação pictográfica. Esse antropomorfismo da visão é muito mais forte para a construção tecnicamente mediada de tais imagens do que, por exemplo, o uso de braços (ou cabeças) humanas para a medição das distâncias no quadro (Alberti 2014:91, 108–9, *passim*). É principalmente como modo de ver que o humano é tomado como “a dimensão e medida das coisas” (Alberti 2014:88).

Do segundo livro, dedicado à “pintura” propriamente dita, há um aspecto que nos interessa sobremaneira. Se trata da apresentação de um exemplo singular de instrumento utilizado para a criação de imagens, que no texto é denotado simplesmente como “véu”:

Para bem aprendê-lo [o exercício da boa circunscrição, considerada a etapa em que se deve concentrar o “trabalho principal” da pintura], nada mais adequado, a [seu] ver, que o véu, o qual, entre os amigos, costum[a] chamar de intersecção. É da seguinte maneira: é um véu muito fino, de tecido pouco fechado, tinto com a cor que se quiser, com fios mais grossos formando quantas paralelas se queiram. Coloc[a-se] esse véu entre o olho e a coisa vista de modo que a pirâmide visual penetre pela tela do véu. [...] (Alberti 2014:102–3).

O uso do dispositivo se justifica na medida em que é “impossível imitar uma coisa que não continua a manter uma mesma aparência”; assim, o véu plasma “o verdadeiro vértice da pirâmide”, fazendo com que “ao ver uma coisa, ela será sempre a mesma” (Alberti 2014:103). Como explica Mineo (2014), o véu projeta sobre o objeto da pintura uma grade quadriculada de paralelas horizontais e verticais sobre a superfície observada, de modo que o véu funciona como uma intersecção da pirâmide visual, pois

²⁰Um exemplo de conceitualização não moderna da imagem está presente na cosmologia Yanomami (e.g., Kopenawa e Albert 2019). Uma interessante discussão sobre aproximações e divergências do uso Yanomami em relação ao ocidental/moderno é apresentado em Senra (2011).

²¹A referência é: BELTING, Hans. La perspectiva y la cuestión de las imágenes: caminos entre Oriente y Occidente. In: Florencia y Bagdad: *Una historia de la mirada entre Oriente y Occidente*. Madrid: Ediciones Akal, 2012, p. 22-92.

ele se interpõe entre o olhar do artista e o objeto, funcionando como “uma prefiguração da própria superfície da pintura (que também é uma intersecção da pirâmide visual)” (Mineo 2014:111). De tantos outros aparatos concebidos naquele momento²², Machado enfatiza outro, a “camera obscura”, de uso generalizado no Renascimento e que era “dispositivo destinado a reproduzir o mundo visível da forma mais exata possível” (Machado 2004:226). Neste caso,

[a] objetividade da imagem obtida por esse dispositivo parecia inquestionável ao homem daquele período: afinal, era a própria realidade que se fazia projetar de forma invertida na parede da câmara oposta ao orifício por onde entrava a luz, enquanto o papel do artista consistia apenas em fixar essa imagem com pincel e tinta. Ou seja, a imagem se originava da própria realidade representada e não da imaginação do artista (Machado 2004:226).

Entretanto, apesar das vantagens de tal dispositivo para a imitação da natureza, ele também “apresentava sérios problemas de definição, além de exibir uma curvatura nas partes mais afastadas do centro, conforme se pode constatar ainda hoje nas câmeras de ‘buraco de agulha’ [pinhole]” (Machado 2004:226). Para corrigir essas deformações, o advento de outro aparato foi necessário, o das lentes “objetivas”, inventadas por Daniele Barbaro no século XVI que, apesar do nome, funcionava refletindo e refratando, isto é, deformando ainda mais, os raios de luz (Machado 2004:226). Desse modo, constata Machado (2004:226), resolvera-se nos séculos XV e XVI “todos os problemas ópticos necessários para a produção ‘automática’ de imagens”, argumentando que essa “tecnologia” é responsável não só por “boa parte da iconografia desse período” como também pela “diretriz metodológica e construtiva” da produção de imagens plásticas.

O uso de tais ferramentas, amparadas em um “primado do intelecto sobre a mão” (Alberti 2014:65) e na busca por uma “imagem cientificamente verossímil”, que balizaram a prática renascentista, são, para Machado (2004:225), “a própria essência” da imagem técnica. É a partir do Renascimento que “a imagem se torna cada vez mais calculada, arquitetada, conceitualizada, construtiva, encarnando a própria utopia de um total controle do visível” e quando “a paisagem visualizada no quadro advém cada vez mais sóbria, encorpada, matematicamente controlada, regida por conceitos de simetria e de funcionalidade” (Machado 2004:224–25). Ou seja, encontramos já na pintura renascentista o ponto de vista antropocentrado e a aparente objetividade que vão conferir à fotografia – tanto quanto à IAg contemporânea –, utilizando a expressão de Machado (1984), sua “ilusão especular”.

4.2.2 Fotografia

Por conta do acima exposto, podemos considerar, com Machado (2004:227), que “[a] fotografia é filha legítima da iconografia renascentista”, não só por utilizar “recursos tecnológicos dos séculos XV e XVI”,

²²Outro exemplo oferecido por Machado (2004:225) é o seguinte: “Era comum, no Renascimento, encontrar no ateliê do artista aparelhos de pintar baseados no princípio da tavoletta de Brunelleschi e constituídos basicamente de um ponto de referência para o olho do pintor (apenas um olho; o outro deveria ser tapado, pois, como se sabe, a imagem renascentista era monocular) e um vidro translúcido para a projeção das imagens. Olhando do ponto de referência, o artista ‘copiava’ sobre o vidro translúcido à sua frente os modelos e objetos colocados no lado posterior. Tal era a materialização da célebre ‘pirâmide visual’ que definia o ato pictórico no Renascimento: o ponto de referência era o vértice ou ‘centro visual’ da pirâmide, ao passo que o vidro correspondia a uma intersecção vertical do objeto geométrico, personificando o quadro (a tela) onde deveriam ser projetadas as imagens das coisas compreendidas dentro da pirâmide. O artista obtinha assim um esboço da imagem, com dados tomados do próprio objeto representado, de modo que bastava transferi-lo depois para a tela e cobri-lo de cores para que a obra ganhasse forma acabada”.

mas principalmente por “dar continuidade ao modelo de imagem construído no Renascimento, modelo esse marcado pela objetividade, pela reprodução mimética do visível e pelo conceito de espaço coerente e sistemático, espaço intelectualizado, organizado em torno de um ponto de fuga”. Para Machado, o empenho do Renascimento foi o de “colocar *fora do homem* a produção de imagens”, no sentido de transferi-la da subjetividade para a objetividade, remetendo-as primeiramente à natureza, e, depois, a “máquinas e instrumentos derivados da investigação científica”²³. Desse modo, opondo-se à imagem interna subjetiva, há uma imagem objetiva exteriorizada, que “não vindo diretamente do homem, pressupõe sempre uma mediação técnica para exteriorizá-la” (Machado 2004:222).

Em continuidade com o ideal renascentista, a invenção dos métodos químicos de fixação da imagem na película fotográfica no século XIX é considerada por Machado (2004:227) como a retirada do “último gesto artesanal”, pois abre a possibilidade “de uma produção inteiramente automática e tecnológica da imagem, dando nascimento, portanto, a uma imagem da qual a intervenção do homem pode ser excluída”. Hoje, entretanto, percebemos que este gesto não era ainda o último, pois a imagem gerada por IA prescinde inclusive do dedo do fotógrafo, mesmo que, como também percebemos, exija muitos outros para efetivar o seu programa. Um automatismo, para gerar significação, não pode prescindir totalmente do gesto humano (Simondon 2020c).

No livro *A ilusão especular*, Arlindo Machado (1984) tece uma contundente crítica ao estatuto dominante da fotografia que a estabelece como representação objetiva e inquestionável do mundo (como na expressão: “uma imagem vale mais do que mil palavras”), mostrando que isso não passa de uma “ilusão especular”, isto é, de uma ideologia que apresenta a fotografia como mimese objetiva, sem uma mediação ativa.

Um aspecto particularmente interessante de sua abordagem é sua crítica à “mística da homologia automática”, que privilegiaria o referente em relação às operações técnicas envolvidas no ato fotográfico. Em outras palavras, é uma visão da fotografia que enfatiza mais a reflexão da luz, que sua refração. Contrapondo-se a visões importantes do campo, como as de Susan Sontag e André Bazin, Machado (1984:37) questiona a natureza indicial da foto: “nada é mais *subjetivo* do que as objetivas fotográficas, porque o seu papel é personificar o olhar do *sujeito* da representação.” Comentando a ideia de que “sem referente não há fotografia”, argumenta que:

Se não existir câmara escura, a lente com seu poder organizador dos raios luminosos, um diafragma rigorosamente aberto como manda a análise da luz operada pelo fotômetro, um obturador com velocidade compatível com a abertura do diafragma e a sensibilidade da película, se não houver ainda uma fonte de luz natural ou artificial modelando o referente e um operador regendo tudo isso, também não haverá fotografia, muito embora o candidato a referente possa estar disponível (Machado 1984:39).

Para o autor, “[a] ênfase no referente, a concepção de fotografia como reflexo bruto da ‘realidade’ só pode se justificar como postura estratégica, isto é, ideológica” (Machado 1984:39). O livro conti-

²³“A imagem ‘objetiva’ é a imagem que vem de fora (‘da natureza’), aquela que se pode apreender com máquinas e instrumentos derivados da investigação científica, e a sua principal virtude é estar imune à subjetividade humana, às imagens interiores, que deformam e adulteram a realidade visível. [...] A perspectiva corrige os dados imagéticos diretamente tomados da realidade (‘da natureza’) e os conforma a modelos matemáticos de beleza, não esquecendo que ‘beleza’, para o homem renascentista, significa reconhecimento da configuração ideal dos objetos, pressupondo portanto o conhecimento de sua estrutura e do seu relacionamento harmônico com todos os outros objetos da cena” (Machado 2004:226–27).

nua apresentando uma série de argumentos sobre a realidade construída da imagem fotográfica, por exemplo mostrando efeitos de permanência da luz, das fotografias em escala radicalmente distintas daquela do olho humano e das deformações provocadas pelo uso de “lentes bizarras” (Machado 1984). Um exemplo particularmente interessante é da imagem que temos, e que é constantemente reproduzido em animações, de uma roda que se molda em uma diagonal baixo-cima, traz-frente em um carro que acelera:

A célebre fotografia que Jacques-Henri Lartigue tirou durante o Grande Prêmio Automobilístico da França, em 1912, é o melhor exemplo dessa dilatação temporal própria da imagem fotográfica: como resultado de um tempo de exposição muito longo em relação ao movimento dos carros e de uma modalidade de obturação que “varre” horizontalmente o quadro da câmera, o referente resulta visivelmente distorcido, com os volumes comprimidos e as linhas verticais inclinadas para frente (as rodas do automóvel) ou para trás (os espectadores do fundo). Isso ocorreu porque à medida que o obturador ia “varrendo” o quadro, o motivo ia se deslocando, de sorte que a cada posição do primeiro correspondia um avanço na posição relativa do segundo. [...] Durante a época em que a maioria das câmeras eram dotadas de obturadores de fenda horizontal, esse efeito de achatamento e inclinação era tão comum nas fotos de objetos velozes que, por transferência semiótica, os próprios desenhistas e artistas gráficos passaram a usar essa convenção para representar a velocidade, como ainda hoje se vê nas histórias em quadrinho (Machado 1984:46–47).

Assim, há uma linha de continuidade que liga os ideais de objetividade e antropocentrismo renascentistas até a maneira dominante de enxergar a fotografia²⁴. Tal qual a perspectiva renascentista, a fotografia como espelho representa objetivamente a natureza e da maneira como o humano a enxerga. O uso de lentes que provocam uma deformação exagerada e a fotografia abstrata, como a série *Fotoformas* de Geraldo de Barros, são frequentemente tidos como exceções que causam estranhamento justamente por apontarem para a natureza construída da aparência de objetividade e de cópia da visão humana que a fotografia geralmente reproduz.

4.2.3 Do computador digital à inteligência artificial

Com o advento das tecnologias digitais, a imagem fotográfica passa por uma diferenciação que resulta, simultaneamente, no “vídeo” – que, como as vanguardas modernas em relação à imagem renascentista, realiza um afastamento em relação ao ideal de representação objetiva do mundo – e na “síntese numérica”, que Machado (2004:232) denomina de imagem digital: “a imagem digital aparece como uma verdadeira hipertrofia dos postulados estéticos do século XV, na medida em que ela realiza hoje o sonho renascentista de uma imaginação puramente conceitual”, a computação gráfica e as abstrações matemáticas “dão origem a imagens ainda mais calculadas, coerentes e formalizadas do que a pintura do *Quattrocento*”. É sobretudo isso o que nos interessa, embora o argumento do autor vá alhures ao se atentar também às mudanças divergentes advindas com o vídeo.

²⁴“Pode-se mesmo dizer que a imagem fotográfica encontra-se marcada por uma fatalidade figurativa que a conecta irremediavelmente com a iconografia renascentista e que só mesmo um gesto extremo de radicalidade, com repercussões na própria engenharia da câmera, pode subverter de forma consequente” (Machado 2004:230).

O diagnóstico do autor é interessante, percebendo na ambiguidade do potencial do computador digital, uma possibilidade de convergência entre essas duas tendências: a maleabilidade do vídeo²⁵ e da homogeneização, metrificacão e comparabilidade da síntese numérica. Se este se conforma aos “cânones oficiais de objetividade e coerência”, “o modelo renascentista, mantido vivo pela imagem técnica”; o primeiro é legatário das práticas anamórficas e da “aventura da arte moderna” na “desconstrução dessa positividade” (Machado 2004:229). Assim, com o vídeo, “em vez da exploração da imagem consistente, estável e naturalista da figura clássica, ela se definirá resolutamente na direção da distorção, da desintegração das formas, da instabilidade dos enunciados e da abstração como recurso formal”, chegando ao ponto em que “a descoberta do efeito de feedback, que permitiu gerar imagens sem necessidade de registro fotográfico, e a invenção dos sintetizadores de vídeo, que possibilitaram criar toda uma iconografia informal sem necessidade do concurso da câmera” fez a video-arte criar uma “iconografia resolutamente contemporânea e a lograr uma reconciliação das imagens técnicas com a produção estética de nosso tempo” (Machado 2004:230–31). Por outro lado, a síntese numérica – apesar de seu realismo matemático desvinculado do referente²⁶ – “reconcilia novamente, senão com o efeito ilusionista da figuração renascentista, pelo menos com o seu caráter construtivo e a sua resoluta fusão de arte e ciência” (Machado 2004:233). Na prática, entretanto, ainda assistimos uma dominância do realismo próprio dos ideais renascentistas:

Na fotografia, no cinema, na televisão e mesmo nos novos produtos audiovisuais propostos pela informática, há uma predominância quase absoluta da imagem especular consistente do século XV, da qual não conseguimos nos desprender mesmo depois de quase um século de desconstrução dessa imagem pela chamada arte moderna. [...] O mais sofisticado spot publicitário exibido na televisão, apesar de construído com recursos tecnológicos de última geração, nos quais se incluem captação em película cinematográfica, pós-produção em vídeo de alta definição e inserções de imagens modeladas e animadas em computador, em geral, nada mais faz que celebrar uma iconografia historicamente datada, tomada como modelar e repetida até a exaustão pelas sucessivas gerações (Machado 2004:228–29)

Acreditamos que o que ocorre na aprendizagem de máquina é similar, pois ainda que a possibilidade de geração ou derivação de imagens a partir de relações calculadas nas bases de dados – como o exemplo das alucinações maquínicas já citado – insiste-se em sua utilização de modo a conformar com uma aparência de objetividade e de antropomorfismo da visão. Nesse sentido, o projeto “*This person*

²⁵Sobre a novidade vídeo, o autor comenta que: “[d]iferentemente da imagem fotoquímica, a imagem eletrônica é muito mais maleável, plástica, aberta à manipulação do artista, resultando portanto mais suscetível às transformações e às anamorfoses. Pode-se nela intervir infinitamente, alterando suas formas, modificando seus valores cromáticos, desintegrando suas figuras. Não por acaso, a arte do vídeo, que se constitui tão logo os recursos técnicos se tornam disponíveis, definir-se-á rapidamente como uma retórica da metamorfose: em vez da exploração da imagem consistente, estável e naturalista da figura clássica, ela se definirá resolutamente na direção da distorção, da desintegração das formas, da instabilidade dos enunciados e da abstração como recurso formal” (Machado 2004:230).

²⁶“Há, todavia, uma diferença: a partir do computador, o ‘realismo’ resulta visivelmente desencarnado, sem qualquer vinculação direta com a paisagem registrada. O realismo praticado na era da informática é um realismo essencialmente conceitual, elaborado com base em modelos matemáticos e não em dados físicos arrancados da realidade visível” (Machado 2004:232). Já Kittler (2016:328), a esse respeito, conclui que “[o] processamento digital de imagens, justamente pelo fato de – ao contrário das artes tradicionais – não pretender ser representação, coincide com o real”, o que afirma a materialidade e performatividade da imagem digital.

does not exist [Essa pessoa não existe]” (*Anón sem data*) é um exemplo cabal. Segundo o repositório de código do software utilizado na página (*Rico [2019] 2024*), ela utiliza a técnica StyleGAN (*Karras et al. 2020*) para gerar rostos realistas de pessoas que não existem a partir de bases de dados de rostos de pessoas. O realismo das imagens nos surpreende pelo fato de ainda atribuímos maior capacidade de mimese passiva à câmera fotográfica do que ao computador, mas sua realização é perfeitamente coerente com a história apresentada aqui dos agenciamentos de representação. Como em outras aplicações de IA – sendo o próprio chatGPT um exemplo disso através de “imagens textuais” – a técnica se propõe a gerar imagens verossímeis ou aparentemente reais, reproduzindo os ideais de objetividade e antropocentrismo.

Mesmo assim, acreditamos que Machado tem parcialmente razão ao constatar uma tendência de aproximação entre ambos os paradigmas:

A tendência atual é encarar o registro fotográfico efetuado pela câmera como a mera obtenção de uma matéria-prima que deverá ser posteriormente trabalhada e transformada por algoritmos de tratamento da imagem. “Fotografia” agora é o nome que se dá ao resultado de um processo de edição e não à marca deixada pela luz sobre uma superfície fotossensível. Em resumo: enquanto certos produtos da computação gráfica aspiram ao (antigo) poder de convicção da fotografia fotoquímica, a fotografia se converte ela própria em vídeo (as próprias câmeras fotográficas já são agora eletrônicas), como que anunciando uma era de indiferenciação fenomenológica entre imagens técnicas e artesanais, objetivas e subjetivas, internas e externas (*Machado 2004:234*).

Com o aumento de situações mediadas por algoritmos, a incorporação de aspectos da imagem digital pela fotografia e por outros modos de criação de imagens é inevitável. Gisele Beiguelman, comentando a interatividade de obras de arte contemporâneas, afirma que “[a]s imagens digitais são, sobretudo, mapas informacionais que contém uma série de camadas, o que permite que sejam relacionadas entre si e com outras mídias, a partir de atributos matemáticos. São esses atributos que vão, por exemplo, relacionar determinada coordenada de uma imagem a um texto ou um comportamento (como um movimento ou ativação de escurecimento, por exemplo)” (*Beiguelman 2021:18*).

No contexto de utilização de smartphones e outras interfaces gráficas mediadas por algoritmos, a constatação da autora extrapola o campo da arte. As imagens digitais em redes sociais, aplicativos, e sites, são dessa natureza, simultaneamente imagens e sinais de informação, são imagens operativas, de modo que elas “nos olham”, da mesma forma que as olhamos, na medida em que elas são, cada vez mais, responsivas a nossas ações – nos leem com sensores, nos classifica por meio de técnicas de perfilamento e, por isso mesmo, cria uma imagem de nós²⁷.

Um último exemplo mostra de maneira particularmente interessante o imbricamento entre fotografia e inteligência artificial: a divulgação da primeira imagem de um buraco negro produzida pelo Telescópio do Horizonte de Eventos (EHT, *Event Horizon Telescope*). A imagem que captura o centro da galáxia M87, localizada à 55 milhões de anos-luz da Terra, foi produzida utilizando oitos telescópios espalhados pela superfície terrestre (*Baio 2022*). Na descrição de Baio:

²⁷“Seja nas redes digitais ou em uma exposição de arte, cada vez mais a imagem nos olha antes de olharmos para ela, assim como se modifica de modo a responder às informações por ela recolhidas” (*Baio 2022:96*).

As imagens apresentadas pela equipe do EHT surgiram de medições de radiação luminosa emitida pelo fenômeno cósmico, mas a transformação desses raios em uma imagem somente foi possível com a aplicação de uma série de procedimentos computacionais. Dentre os procedimentos descritos nos artigos científicos publicados pela equipe estão diferentes métodos de reconstrução de imagens a partir de dados brutos, uso de dados sintéticos, pesquisa de parâmetros-chave e marcação fiducial (AKIYAMA et al., 2019). Antes mesmo da etapa de reconstrução da imagem por software, foram utilizados algoritmos correlacionadores para alinhar os sinais de cada um dos oito telescópios espalhados pelo globo terrestre a uma referência de tempo e espaço comum. Tal procedimento visa corrigir as diferenças nos dados captados de pontos distintos da geometria terrestre, criando um “ponto de vista único”, e pela distância temporal entre os múltiplos momentos em que os dados brutos foram captados, criando o “instante fotográfico” da imagem reconstruída. Porém, para conseguir revelar a fotografia, os pesquisadores precisavam de informações além daquelas efetivamente captadas. Para preencher os pontos cegos dos telescópios, foram utilizadas tecnologias de aprendizagem de máquina (ramo da inteligência artificial), treinadas para completar a ausência de dados, de modo a dar a ver o objeto cósmico em um campo visual que preenchesse por completo o quadro fotográfico (Baio 2022:91–92).

A descrição de Baio, que explicita os bastidores do processo de criação da imagem, nos permite ver com muita clareza o aspecto operativo, construído e performativo da foto, mas que, para o autor, “corresponde a características reconhecidamente herdadas da cultura visual da ontologia clássica” (Baio 2022:92). Isso porque tanto o uso de aparatos técnicos, a explicitação dos métodos em uma publicação em artigo científico e a preocupação com reprodutibilidade mostram uma construção de objetividade para a imagem; mas também, porque, tal qual apontado pelas próprias cientistas envolvidas, o processo de criação dessa fotografia permitiria a geração de uma multiplicidade de imagens, de forma que a imagem divulgada publicamente envolveu uma série de decisões que construíram uma foto em detrimento de outras: ao escolher divulgar uma única imagem do processo de captação dos raios luminosos da galáxia, “a partir de um determinado ponto de vista e com um quadro completamente preenchido” (Baio 2022:92), a fotografia reproduz, em algum nível, um antropomorfismo da visão remanescente do Renascimento.

Segundo Baio (2022), essa imagem produzida pelo EHT é da mesma natureza que as imagens produzidas por câmeras profissionais ou por smartphones, pois em todos esses casos, trata-se de construções visuais algoritmicamente mediadas – especialmente, o processamento de sinais digitais produzidos por sensores de ondas eletromagnéticas e que, cada vez mais, envolvem aprendizado de máquina em alguma etapa. Assim, tanto na foto do EHT, quanto em câmeras profissionais e smartphones, há uma série de elementos técnicos que participam da construção da imagem e compõem a sua performatividade. Havendo exposto um caso em que o AdM integra a construção de uma imagem, passemos para uma investigação mais específica sobre a imagem no AdM.

4.3 Aprendizado de máquina e espaço latente: imagem, representação e performatividade para além da visualidade

Para descrever a imagem no aprendizado de máquina, partimos sobretudo das redes neurais profundas, que são endogenamente consideradas como um caso ímpar de um *aprendizado de representação* (LeCun et al. 2015), mas cujo funcionamento gostaríamos de generalizar. Nelas, explicita-se sobretudo o ideal “prospectivo” de criação de imagens de mundo a partir da análise de dados passados que geram classificações, previsões ou gerações de padrões. Examinemos mais detidamente o que quer dizer considerar um algoritmo de aprendizagem de máquina como um de aprendizado de representação e o que significa dizer que há, nele, uma imagem, mesmo que ela escape à visualidade.

Amparando-se na literatura técnica (e.g. Goodfellow et al. 2016), percebe-se que as RNAs podem ser vistas como funcionando em dois momentos distintos: “treinamento” e “inferência”, embora na prática nem sempre esses processos sejam distintos no tempo (exemplo na figura 4.1). Em uma primeira etapa, no *treinamento*, o algoritmo processa exemplos de uma base de dados e calcula automaticamente parâmetros numéricos (também chamados de “pesos sinápticos” no caso das RNAs) que possibilitam-na resolver determinada tarefa (é o processo mais explícito de “aprendizagem” a partir de dados). Além dos parâmetros numéricos, calculados dinamicamente no treinamento, dois outros elementos determinam a estrutura do modelo e da aprendizagem: sua *arquitetura* – isto é, quantidade, disposição e tipo de neurônios no caso de uma RNA (um exemplo é a importante arquitetura “transformer” utilizada em MdLs –, e a *função objetivo* – “uma descrição [matemática] da finalidade da máquina” (Gonçalves 2022:12, por exemplo a minimização do “erro quadrático médio”, ou da “entropia cruzada”) – (conferir também Goodfellow et al. (2016)). Em seguida, na etapa de *inferência* ou, como também entendemos, *incidência*, novas entradas podem ser inseridas no algoritmo que, em um processo de cálculo computacional, produz uma saída no sentido de resolução da tarefa. Dessa forma, o algoritmo é capaz de generalizar um conhecimento da base de dados para entradas inéditas através de processos de interpolação e extrapolação (Pasquinelli e Joler 2021), isto é, de síntese de resultados que ocupam lugar, respectivamente, intermediário a dados da base, ou para além deles.

Em termos mais abstratos, uma rede neural pode ser conceituada como uma função matemática flexível: ela recebe uma entrada x e produz uma saída y . O que determina a tarefa que a função $y = f(x)$ resolve são os parâmetros numéricos automaticamente calculados a partir da base de dados (Gonçalves 2022). Essa analogia é especialmente valiosa, pois possibilita uma aproximação da nossa ideia de imagem com aquela utilizada na matemática. Nesta, a imagem (ou mais precisamente o “conjunto imagem”) é o conjunto de *todos os resultados possíveis de uma dada função*, ou ainda, seus resultados potenciais, seu campo de valores virtuais. Formalmente, dizemos que em uma função $f : X \rightarrow Y$ (função f que relaciona elementos do conjunto domínio X a um conjunto contradomínio Y), sua imagem é o subconjunto Im do contradomínio Y tal que:

$$\text{Im}(f) = \{y \in Y \mid y = f(x) \text{ para todo } x \in X\}$$

Ou seja, a imagem da função f contém elementos y pertencentes a Y tal que y é o resultado da avaliação da função $f(x)$ para todo x pertencente a X . Por exemplo, se a função f é aquela que representa a potenciação ao quadrado $f(x) = x^2$, tomando como domínio e contradomínio o conjunto de todos os números reais (\mathbb{R}), então, sua imagem é o conjunto de todos os resultados possíveis, isto

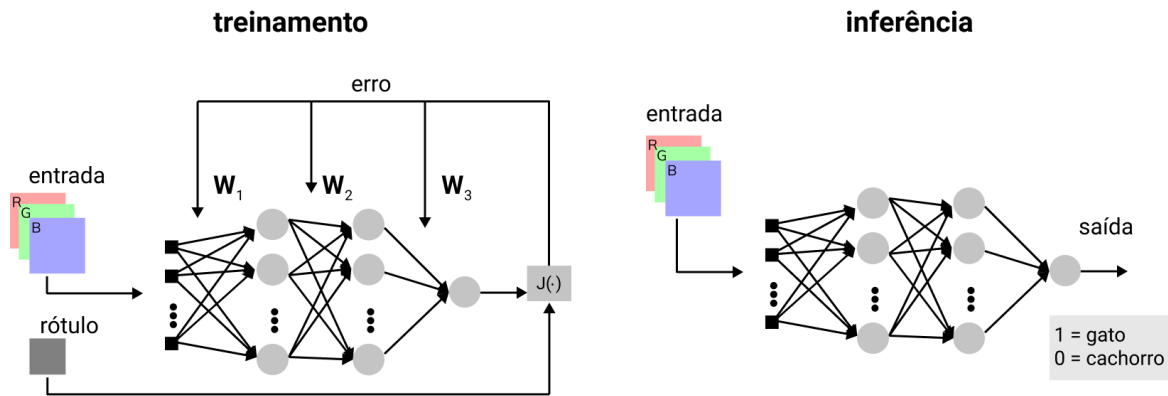


Figura 4.1: Exemplo esquemático de uma rede neural utilizada para classificar imagens visuais (em um computador digital representadas como matrizes de pixels RGB – vermelho, verde e azul) de cachorros e de gatos. São representados os dois momentos do algoritmo: treinamento e inferência. Na figura, W se refere aos parâmetros numéricos que caracterizam a saída da rede neural, enquanto que $J(\cdot)$ é a função objetivo, métrica utilizada para comparar o resultado real e esperado na etapa de treinamento. O rótulo é o resultado esperado para a saída do modelo em cada iteração do treinamento, isto é, zero ou um. Fonte: Gonçalves (2022).

é, apenas os números reais positivos (\mathbb{R}_+), pois não é possível obter um número negativo ao se operar uma potência de dois sobre um número real. Uma representação gráfica do contradomínio, domínio e imagem está apresentada na figura 4.2.

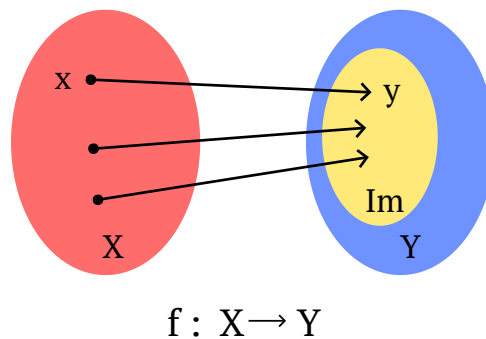


Figura 4.2: Domínio (X) em vermelho, contradomínio (Y) em azul e imagem (Im(f)) em amarelo. Fonte: elaboração própria, baseado em <https://commons.wikimedia.org/wiki/File:Codomain2.SVG> (acesso: 02/07/2024).

Tendo isso em vista, o que é chamado de *espaço latente*, *espaço de características* ou *representação interna* no AdM é também uma imagem, repositório de virtualidades, que existe como efeito dos parâmetros numéricos no interior de uma rede neural, ou ainda, reservatório potencial de todos os resultados possíveis dessa rede-função. Percebamos que, embora esses valores sejam virtuais, os parâmetros que os condicionam são atuais, isto é, existem no aqui e agora, de modo que, existindo por meio da informação digital, esse campo virtual já está, de certa forma, capturado (cf. Garcia dos Santos 2003a). Em geral o espaço latente é um espaço matemático-computacional abstrato e multidimensional formado a partir de variáveis internas (latentes) ao modelo de AdM (cf. Goodfellow et al. 2016). Essas variáveis, os próprios parâmetros numéricos, ou efeitos deste, não são feitas para serem observadas visualmente por humanos, mas participam do processo de cálculo do resultado final da rede neural (em outras pa-

lavras, podem ser entendidas como imagens m2m). Exemplos desses espaços são o espaço de vetores de palavras (*word embeddings*) em algoritmos de modelagem de linguagem (figura 4.3); o espaço de características de imagens em redes neurais convolucionais (figura 4.4); e o espaço latente da camada intermediária de uma rede AE que pode ser visualizado como um *manifold* (figura 4.5).

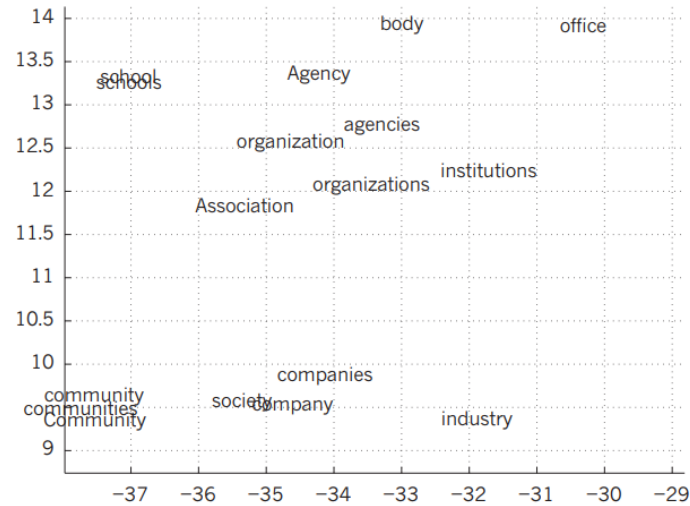


Figura 4.3: Representação bidimensional (via t-SNE, algoritmo de redução de dimensionalidade) de vetores de palavras multidimensionais de uma rede neural de modelagem de linguagem. Os eixos X e Y são duas dimensões numéricas abstratas. Para nós, o que importa desta representação é que uma proximidade entre palavras no gráfico indica uma proximidade entre os valores numéricos dos vetores que representam essas palavras no espaço latente do modelo. Desse modo, é possível perceber que existe uma relação entre uma proximidade numérica e uma proximidade semântica. Fonte: Le Cun, Bengio e Hinton (2015).

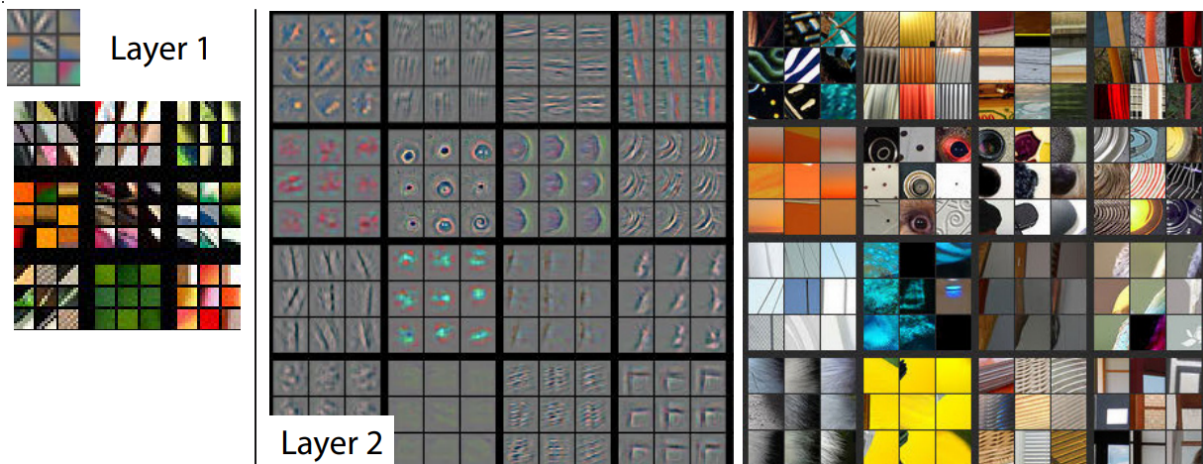


Figura 4.4: Visualização de duas camadas internas de uma rede neural convolucional e da projeção do espaço latente na forma de pixels via desconvolução. A ideia implícita neste trabalho é mostrar que o espaço latente nas redes neurais convolucionais registraria informações visuais que vão de estruturas mais simples (linhas e degradês na primeira camada) até estruturas mais complexas (quinas e curvas na segunda camada). Fonte: Zeiler e Fergus (2014).

Dois postagens de Christopher Olah – ex pesquisador de interpretabilidade na OpenAI e que também já trabalhou na Google Brain –, nos ajudam a entender esse processo. Segundo o autor, a aprendizagem de máquina tem a particularidade de que, “[e]m algum nível fundamental, ninguém entende” embora seu funcionamento seja “fundamentalmente muito simples” o que possivelmente tem

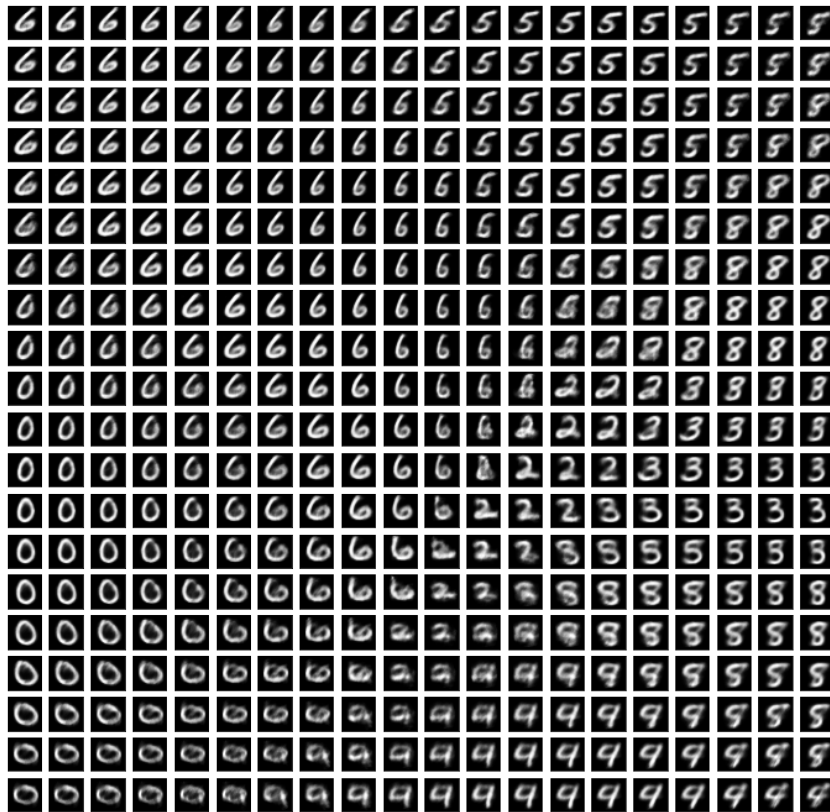
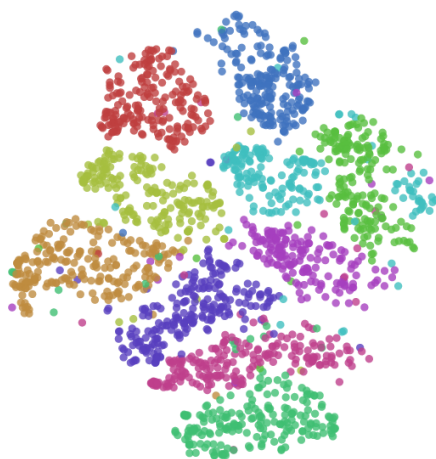


Figura 4.5: Manifold para uma rede neural de tipo autocodificador treinada na base de dados MNIST. Cada imagem representa a saída produzida pelo modelo após um deslocamento de posição no espaço latente de um autocodificador treinado para gerar dígitos escritos a mão. A posição de um resultado na grade do manifold se relaciona com sua posição relativa no espaço latente. Percebe-se como há uma continuidade do espaço que, por sua vez, captura informação: características estruturais dos dígitos escritos a mão. Assim, os dígitos gerados na primeira fileira vão de representações claras do número 6 até o 5 passando por formas intermediárias que não são claramente um 6 ou um 5. Fonte: gerado pelo autor.

a ver com a limitação humana para trabalhar com “milhares”, “dezenas de milhares” ou “milhões” de dimensões que esses algoritmos mobilizam (Olah 2014b). No primeiro artigo (Olah 2014b), a base de dados MNIST é explorada sob a “hipótese do manifold” e da “análise topológica de dados”: “MNIST é um manifold de baixa dimensão, varrendo e curvando-se através de seu espaço latente [embedding space] de alta dimensão”. Sua investigação mostra como, por meio de transformações não-lineares de redução de dimensionalidade, os dígitos da base de dados mencionada se tornam separáveis em um espaço bidimensional (figura 4.6).

Na segunda publicação, Olah (2014a) nos mostra como as transformações operadas por uma rede neural podem ser vistas como transformações não-lineares aplicadas sobre o espaço inicial dos dados (figura 4.7) o que, em dimensões adequadas, é capaz de separar conjuntos de dados que não seriam separados em dimensões menos elevadas (figuras 4.8 e 4.9). Juntando ambas as ideias, entendemos mais claramente o que ocorre na aprendizagem de máquina: produz-se uma representação dos dados a partir da qual tarefas como a classificação ou a predição se tornam simples ou evidentemente resolvíveis para uma máquina. A grande potência do aprendizado de máquina não está na resolução de tarefas por vias extremamente complexas, mas, ao contrário, na criação de uma *boa representação* em que prever a próxima palavra, classificar um ponto, ou inferir um número, seja tarefa particularmente simples. O AdM é, principalmente, um agenciamento de representação.



A t-SNE plot of MNIST

Figura 4.6: Redução de dimensionalidade via t-SNE dos dígitos da base de dados MNIST. Respectivamente: zeros (vermelho), seis (azul), dois (amarelo), cinco (azul claro), três (verde claro), uns (laranja), setes (roxo), oitos (lilás), noves (magenta) e quatros (verde). Fonte: Olah (2014b).

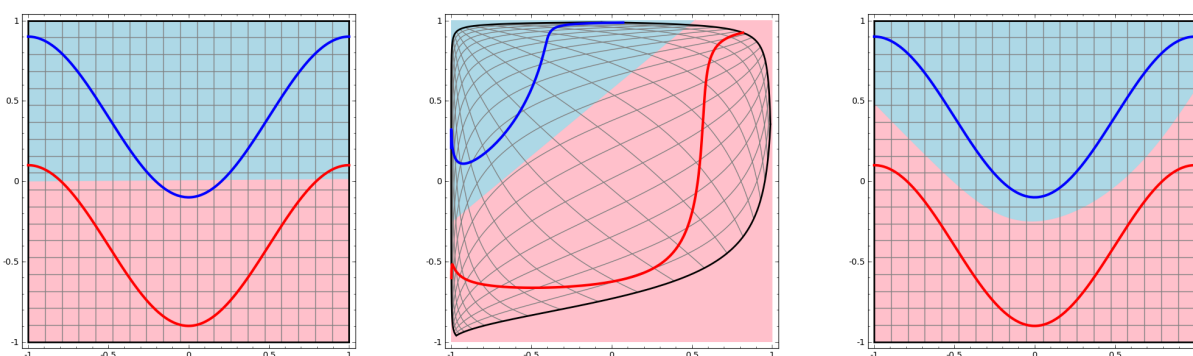


Figura 4.7: Exemplo de dois conjuntos não separáveis linearmente (esquerda), de uma transformação não-linear no conjunto que o faz separável linearmente (centro) e do resultado dessa transformação como uma classificação não-linear no conjunto de entrada (direita). Desse modo, a operação de um algoritmo de aprendizado de máquina não-linear pode ser visto tanto como uma classificação não-linear quanto uma deformação não-linear de um espaço que passa a ser linearmente separável. No AdM, o treinamento visa calcular números que vão parametrizar essa transformação. Fonte: Olah (2014a).

O que há de comum nos três exemplos de espaços latentes apresentados acima e que, esperamos, podem ser melhor entendidos à luz das análises de Olah, é que esse espaço condensa a informação digital proveniente de uma grande quantidade de exemplos da base de dados – uma série diacrônica – em uma representação sincrônica mais eficaz para a inferência e incidência de sinais de informação. Essa condensação não é feita sem problemas e possui limites como, por exemplo, a estrutura do modelo que permite representar os dados em mais ou menos dimensões (cf. Gonçalves 2022, e.g. “2.1 Limitação na representação dos dados de entrada”). Desse ponto de vista, a criação do espaço latente pode ser entendida como um processo de compressão de informação, como argumentado em Pasquinelli e Joler (2021). Mas, para além disso, também pode ser dito que algo novo é produzido nesse processo: nos algoritmos de modelagem de linguagem, por exemplo, a proximidade dos vetores pode ser correlacionada com proximidades semânticas (Mikolov et al. 2013), o que permite que o algoritmo seja capaz de gerar textos com sentido ou operar traduções (é o que ocorre no caso do recente chatGPT); as camadas

Figura 4.8: Exemplos animados ([gif1](#),[gif2](#)) de um caso de sucesso e de falha na separação dos dados por transformação não-linear. Fonte: Olah (2014a).

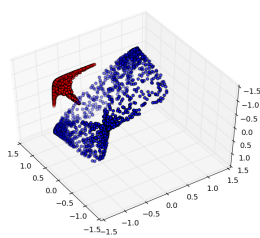


Figura 4.9: Exemplo animado ([gif](#)) de um caso de falha de separação dos dados em um plano bidimensional (esquerda), exemplo estático de sua separação em três dimensões (centro) e exemplo animado ([gif](#)) do sucesso de separação observando um corte bidimensional do caso tridimensional. Fonte: Olah (2014a).

internas de uma rede neural convolucional fabricam estruturas parciais que são utilizadas para compor uma miríade de imagens complexas (como em tecnologias de visão computacional); e o espaço latente de um autocodificador se apresenta como um espaço computacional contínuo que permite a geração de algarismos e até mesmo de intermediários – representação gráfica entre um dígito e outro – (processo similar ao que acontece em tecnologias como dall-E e Stable Diffusion para geração de imagens). O espaço latente do modelo de AdM, entendido como uma imagem que condensa as virtualidades extraídas do processamento de dados, é justamente o excedente que é capturado da digitalização até a criação dos parâmetros numéricos (capítulo 3); ao mesmo tempo que é o potencial de ação que permite ao modelo operar uma modulação sobre a usuária (capítulo 5). É importante notar que esses exemplos dados são acepções parciais do espaço latente, na medida em que ele se constitui como uma imagem com um componente virtual, embora capturado na forma de parâmetros atuais, isto é, ela possui um potencial de produções diversas que não se esgota em uma execução, mas que está limitado às possibilidades de modelagem quantitativa do modelo de AdM. Cada um desses exemplos mostra atualizações parciais que, esperamos, permitem ilustrar a ideia geral do que chamamos da “imagem” no AdM.

Além disso, apontamos que a construção dessa imagem – que de certa forma é imagem da atividade

da usuária no que diz respeito às possibilidades de modulação – é feita com base na composição de uma multiplicidade de outras imagens heterogêneas. O computador digital e, em especial, o smartphone, permitem novas formas de representação imagética que não se limitam ao campo visual: microfones captam voz (imagem sonora), GPS a geolocalização (imagem geográfica) e uma miríade de sensores e algoritmos numa interface visual proliferam dados cada vez mais heterogêneos sobre a realidade. São essas imagens parciais que constituem a imagem geral que define o espaço latente de um algoritmo de AdM.

4.3.1 Exemplo na patente

A patente *Performing subtask(s) for a predicted action in response to a separate user interaction with an automated assistant prior to performance of the predicted action* (US 11222637) (Mirelmann et al. 2022) visa prever interações entre uma usuária e um assistente automatizado baseado em dados sobre interações anteriores. A justificativa para a proposição do método é minimizar latência na interação usuária-assistente, o que também pode reduzir tempo de tela da usuária ou outro recurso (internet, por exemplo). Desse modo “subtarefas identificadas podem ser realizadas antes da usuária prover requisições subsequentes relacionadas com uma ou mais ações previstas” (Mirelmann et al. 2022:14). Essas ações são aquelas que a usuária “é propensa a [*is likely to*] requerer em seguida”, como um pedido para desligar as luzes da cozinha, ler mensagens de e-mail, mostrar a temperatura, entre outras (Mirelmann et al. 2022:19).

Isso é possível por meio da criação de uma imagem das ações realizadas pela usuária, o que também significa a construção de uma capacidade de antecipar suas ações subsequentes. Essa capacidade contida na imagem é consequência direta do processamento de dados:

Em algumas implementações, a entrada para o mecanismo de previsão de ação 206 pode ser processada usando um ou mais modelos treinados de aprendizado de máquina, que podem ser treinados de acordo com uma variedade de dados diferentes. Por exemplo, os dados usados para treinar um modelo de aprendizado de máquina empregado pelo mecanismo de previsão de ação 206 podem incluir *dados de interação histórica entre a usuária 218 e o assistente automatizado 204, dados de interação histórica entre uma ou mais outra usuária e uma ou mais outras instâncias do assistente automatizado, dados contextuais caracterizando figuras [pictures] de um contexto no qual a usuária forneceu a declaração falada 222, recursos operacionais do dispositivo de computação 202, um local na usuária 218 e/ou no dispositivo de computação 202, conteúdo de linguagem natural da declaração falada 222, dados de rede caracterizando propriedades de uma rede local (por exemplo, uma rede Wi-Fi fornecida por um roteador Wi-Fi 214) e/ou qualquer outra fonte de dados que pode ser usada para fazer previsões sobre ações que uma usuária solicitará a execução* (Mirelmann et al. 2022:16, grifo nosso).

Embora a interação com o assistente por meio da voz esteja prevista, é interessante notar como no método o áudio é primeiramente convertido para texto, com a justificativa de que a identificação semântica é mais fácil ali: um algoritmo de reconhecimento de fala pode utilizar redes neurais ou outros modelos estatísticos para identificar porções de áudio que corresponderiam a palavras ou frases, a partir disso o “dado textual pode ser utilizado para gerar e/ou identificar frase(s) de comando, intenção(ões),

ação(ões), valor(es) [*slot values*] e/ou qualquer outro conteúdo especificado pela usuária” (Mirelmann et al. 2022:18). A partir desses dados, uma imagem é continuamente redefinida para modelar, em termos de ações e probabilidades, o padrão de interação usuária-assistente:

O modelo de aprendizado de máquina treinado pode ser usado para processar a entrada a fim de gerar as ações previstas e, opcionalmente, gerar uma probabilidade correspondente para cada ação. *Uma probabilidade para uma ação prevista pode caracterizar uma probabilidade de que a usuária direcionará o assistente automatizado para inicializar o desempenho da ação prevista. Quando a usuária solicita que uma ação prevista seja executada, o modelo de aprendizado de máquina treinado pode ser opcionalmente modificado para refletir a previsão confirmada. Alternativamente, ou adicionalmente, quando a usuária ignora a solicitação de uma ação prevista dentro de um período limite de tempo, o modelo treinado de aprendizado de máquina pode ser opcionalmente modificado para refletir a ação prevista incorretamente*(Mirelmann et al. 2022:13, grifo nosso).

Nota-se aqui também como o treino e a inferência ocorrem de maneira simultânea (via atualização da imagem do modelo). Um exemplo concreto do funcionamento do método patenteado seria a criação de uma imagem que incluisse a preferência de duas usuárias distintas em relação a um serviço de streaming de música. Se os dados contextuais indicarem que a usuária requereu uma ação desse tipo, o mecanismo de predição de ação pode identificar uma playlist para ser transmitida em uma ação de streaming de música. Entretanto, se uma segunda usuária for identificada, o mesmo mecanismo pode identificar uma playlist diferente para ser sugerida ou ainda um outro serviço de streaming. “A identificação da primeira aplicação para a primeira usuária pode ser baseada nos dados contextuais da primeira usuária que indiquem que um primeiro serviço de streaming é utilizado mais frequentemente (ou mesmo exclusivamente) pela primeira usuária para a ação de ‘stream de música’ ” (Mirelmann et al. 2022:15). O mesmo pode ser feito para a segunda usuária:

Consequentemente, o mecanismo de previsão de ação 106 pode primeiro gerar previsão(ões) gerais de ação, opcionalmente utilizando um modelo de aprendizado de máquina treinado usando exemplos de treinamento semissupervisionados rotulados com base nas previsões gerais de ação que correspondem às ações mais específicas realmente realizadas. Além disso, o mecanismo de previsão de ação 106 pode utilizar os dados contextuais para determinar ação(ões) mais específica(s) para que dados de avanço de ação [*action advancement data*] possam ser gerados e/ou subtarefa(s) realizada(s), que é(são) adaptada(s) a essa(s) ação(ões) mais específica(s). Por exemplo, uma conexão com o primeiro serviço de streaming pode ser iniciada para a primeira usuária, enquanto uma conexão com o segundo serviço de streaming é iniciada para a segunda usuária. Dessas e de outras maneiras, um modelo de aprendizado de máquina treinado utilizado pelo mecanismo de previsão de ação 106 pode ser compacto (e armazenável e utilizável de forma eficiente em dispositivos clientes) e/ou pode ser treinado de forma eficiente por meio da utilização das previsões gerais de ação. (Mirelmann et al. 2022:15)

A patente nos mostra, além disso, que essa determinação de preferências e probabilidades não é, senão, o resultado de um conjunto muito específico e complexo de cálculos realizados no modelo, incluindo, por exemplo, métricas distintas para diferentes atividades inferidas:

Em algumas implementações, os atributos da usuária e/ou seu dispositivo de computação que são incluídos como entradas e processados usando o modelo de aprendizado de máquina treinado podem incluir métricas para a usuária de um ou mais tipos específicos de ações. *Como um exemplo não limitativo, as métricas para a usuária podem incluir uma métrica correspondente para cada um dos tipos de ações cujas probabilidades são previstas utilizando o modelo de aprendizado de máquina treinado. Cada uma das métricas pode ser baseada em uma quantidade de desempenhos, para um tipo correspondente de ações, com base na entrada da usuária e/ou no dispositivo de computação.* Opcionalmente, os desempenhos podem ser desempenhos que ocorreram dentro de um limite de tempo (por exemplo, nos últimos três meses) e/ou desempenhos de ações mais recentes podem ser ponderados mais fortemente do que desempenhos de ações menos recentes. *Como um exemplo específico, suponha que uma primeira usuária apenas emprega o assistente automatizado para controlar seu sistema de alarme e transmitir música, e executa ambas as ações em quantidades iguais. Em tal exemplo, esses dois tipos de ações podem receber uma probabilidade de 0,5, enquanto todos os outros tipos de ações receberiam 0. Como outro exemplo, se uma segunda usuária empregar apenas o assistente automatizado para assistir a filmes, controlar luzes, obter direções e ver sua agenda, e executar cada uma dessas ações em quantidades iguais, cada um desses tipos de ações pode receber uma probabilidade de 0,25, enquanto todos os outros tipos de ações podem receber 0.* Note-se que, em tais exemplos, as probabilidades geradas com base no processamento de tais métricas (junto com a(s) ação(ões) que a usuária solicitou) não estarão em conformidade estrita com as métricas em muitas (e talvez todas) situações. No entanto, as probabilidades serão influenciadas pelas métricas de diferentes usuárias, de modo que, por exemplo, diferentes probabilidades serão geradas para as usuárias A, B e C, cada um com métricas diferentes, mas que solicitaram a realização da mesma ação. (Mirelmann et al. 2022:19, grifo nosso).

Desse modo, pudemos localizar em Mirelmann et al. (2022) a descrição detalhada da criação de uma imagem, em sentido amplo, nos algoritmos de AdM. Esta imagem é aquela derivada do processamento de exemplos de entrada no treinamento que cria, em uma síntese comprimida, uma representação da atividade que o método visa modelar: no caso da patente, uma interação de uma usuária com um assistente automático. Mas esta imagem não é apenas atualização material – na forma de parâmetros numéricos, bits e ordenamentos eletrônicos – da atividade virtual das usuárias que é capturada na informação digital. É também, e talvez principalmente, imagem virtual, potencial de se atualizar na forma de incidências e inferências informacionais que vão compor a ação final desempenhada pela usuária-dispositivo-de-computação: no nosso caso, a execução de subtarefas que facilitam, ou atrapalham, uma determinada execução visada pela usuária²⁸.

²⁸Novamente, aqui fica claro o aspecto performativo da imagem que pode penalizar a usuária, por exemplo, no caso de uma demora de execução em relação ao tempo esperado, modulando a usuária para que se conforme com o padrão comportamental que faz melhor funcionar a máquina: “em algumas dessas implementações, o método pode incluir ainda a atribuição, com base em uma obrigação computacional estimada específica [*particular estimated computational obligation*] determinada para a ação fornecida, de um tempo de expiração para a usuária e/ou outra usuária solicitar a execução da ação fornecida” (Mirelmann et al. 2022:22).

4.4 O processo de imaginação no aprendizado de máquina

A patente *Message suggestions* (US11252114) (Braun e Patel 2022) também nos permite analisar esse processo de criação de imagem, de imaginação do AdM, de forma privilegiada, inclusive do ponto de vista de uma imagem que não é visual (mas sim textual) (figura 4.10). Com a justificativa da eficiência²⁹, este método envolve a captura de textos de mensagens recebidas, a determinação de conteúdo específico nas mensagens analisadas por meio da conversão do texto em vetores matemático-computacionais (na imagem do AdM) e a geração de sugestões que são utilizadas simultaneamente como substituição da agência da usuária e como exemplos de treino que constrói o potencial do AdM:

O método de exemplo 100 processa as comunicações associadas a uma conta de usuária (e.g., mensagens de texto enviadas e recebidas, e-mails, IMs, PMs, chamadas telefônicas, correios de voz e similares) para gerar dados de treinamento que um sistema de computação pode utilizar para treinar um aplicativo de aprendizado de máquina [*machine-learning application*]. O sistema de computação pode então usar o aplicativo de aprendizado de máquina treinado para processar uma mensagem recebida e fornecer uma ou mais mensagens prontas relevantes e personalizadas [*relevant and personalized canned responses*] que podem ser selecionadas para serem incluídas em uma resposta à mensagem recebida (Braun e Patel 2022:17).

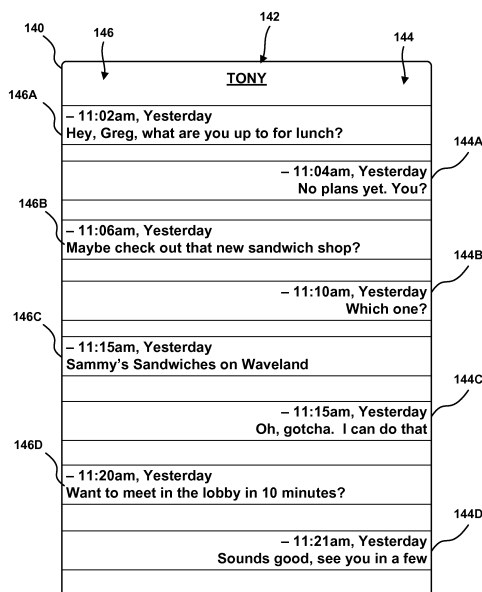


FIG. 2

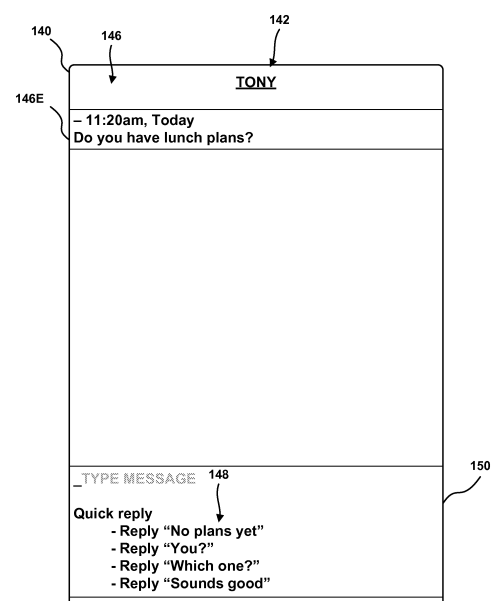


FIG. 3

Figura 4.10: Exemplos do funcionamento dos métodos de sugestão de mensagens. À esquerda, um fluxo de mensagens entre Tony e Greg é apresentado como série histórica a ser capturada e utilizada pelo método da patente. À direita, as mensagens já processadas proveem sugestões à Greg para a mensagem de Tony. Fonte: Braun e Patel (2022).

²⁹“Em alguns exemplos, ao fornecer aos aplicativos acesso a técnicas de adaptação de máquina projetadas para operar e aprender sobre o comportamento de uma usuária de uma plataforma de computação, a aplicação de aprendizado de máquina pode tornar os sistemas de computação mais fáceis de usar, mais eficientes do ponto de vista da usuária e economizar o tempo e o esforço da usuária na utilização da variedade de aplicativos disponíveis no sistema de computação” (Braun e Patel 2022:17).

Ressalta-se que, para o treinamento, o modelo pode ainda associar uma determinada mensagem tanto com o histórico passado, quanto futuro³⁰. Isto é, o vetor gerado para determinada mensagem pode incluir o contexto de mensagens passadas e posteriores àquela mensagem. Na imagem, tal qual em uma superfície visual, as informações estão sincronicamente dispostas para que relações entre elas sejam estabelecidas. A codificação das mensagens, na patente, se dá na forma de vetores numéricos que podem representar tanto uma contagem de palavras, quanto relações semânticas:

No bloco 104 do método 100, o sistema de computação processa, para uma respectiva mensagem da pluralidade de mensagens enviadas, um conjunto de uma ou mais mensagens na mesma conversa que a respectiva mensagem para criar um vetor de características [*feature vector*] que inclui um ou mais características [*features*] do conjunto de uma ou mais mensagens. Por exemplo, o vetor de características pode contar instâncias de um ou mais características do conjunto de uma ou mais mensagens. O conjunto de uma ou mais mensagens pode incluir ou não a respectiva mensagem enviada. Em outro exemplo, o processamento no bloco 104 pode operar em um conjunto de uma ou mais mensagens do mesmo remetente, embora a(s) mensagem(ns) possa(m) não estar necessariamente associada(s) à mesma conversa que a respectiva mensagem enviada. Em um exemplo, as primeiras características incluem uma ou mais palavras no conjunto de uma ou mais mensagens e, no bloco 104, *o sistema de computação cria o vetor de características usando um modelo de saco de palavras [bag-of-words] ou um modelo n-gram. No modelo de N-gram, N pode ser um número inteiro, como 1 no modelo de saco de palavras, ou N pode ser 2, 3, 4 e similares, e cada característica pode representar um grupo de N palavras consecutivas da(s) mensagem(ns). O vetor de características resultante é geralmente um vetor esparsa que conta as ocorrências de cada característica encontrada no subconjunto de mensagens (Braun e Patel 2022:18, grifo nosso).*

Conforme descrito nesse trecho, um modo simples de representar um texto numericamente é criando um vetor de características que representa a frequência de cada palavra ao longo do texto: para cada palavra, o valor do vetor em determinada dimensão equivale a quantidade de aparições dela no texto (modelo saco de palavras). Outra possibilidade é estender esse modelo para sequência de duas ou mais palavras, contando a ocorrência de duplas ou triplas, por exemplo (modelo N-gram). Um dos problemas desse método, apontado no texto, é que o vetor resultante será muito provavelmente um vetor esparsa, isto é, em que a maioria de suas dimensões valem zero (o que ocorre pois a maioria dos textos de entrada não possuem todas as palavras do dicionário). Isso não é muito eficiente em termos de utilização de recursos, esbarrando em limites computacionais em relação ao número de palavras computadas no vetor. Esse problema relativo à dificuldade de representação numérica de textos, que rapidamente implica a criação de vetores com muitas dimensões, é frequentemente referido como a “maldição da dimensionalidade”, o crescimento exponencial do uso de recursos com o aumento das dimensões dos dados de entrada (Pasquinelli e Joler 2021:1273). Embora neste caso, o modelo aponte

³⁰Por exemplo: “o processamento no bloco 104 pode ser realizado à medida que cada mensagem é enviada da conta da usuária e, portanto, o processamento opera em mensagens prévias. Entretanto, em alguns exemplos, o processamento no bloco 104 pode operar em mensagens anteriores e/ou subsequentes. Além disso, o processamento no bloco 104 pode operar em um conjunto de N (e.g., 3) mensagens que precedem imediatamente (e/ou vêm depois) uma respectiva mensagem enviada” (Braun e Patel 2022:18).

o uso de um modelo de saco de palavras, uma técnica mais clássica, gostaríamos de apontar que o uso de técnicas mais complexas como a word2vec apresentada por Mikolov et al. (2013) permitem uma representação numérica das características semânticas ainda mais precisa e eficiente. A técnica word2vec cria dinamicamente vetores densos para as entradas, com dimensões que não se referem claramente a palavras, mas que criam um espaço abstrato em que palavras semanticamente similares se aproximam matematicamente:

De forma surpreendente, essas questões [que visam estabelecer relações entre pares de palavras] podem ser respondidas pela realização de operações algébricas simples com a representação vetorial da palavra. Para descobrir uma palavra que é similar a pequeno [*small*] no mesmo sentido que maior [*biggest*] é similar a grande [*big*], nós podemos simplesmente computar o vetor $X = \text{vetor}(\text{"biggest"}) - \text{vetor}(\text{"big"}) + \text{vetor}(\text{"small"})$. Então, nós procuramos no espaço vetorial pela palavra mais perto de X medida pela distância de cosseno, e usamos ela como resposta para a questão (nós descartamos as palavras da questão de entrada durante essa busca). Quando os vetores de palavras estão bem treinados, é possível encontrar a palavra correta (palavra menor [*smallest*]) usando esse método. (Mikolov et al. 2013:5)

Se do ponto de vista da representação é evidente que a possibilidade de operar matematicamente sobre dados que não são originalmente numéricos é de grande valor, notamos que também do ponto de vista da compressão, a representação vetorial é vantajosa. Esse tipo de representação diminui a complexidade computacional em relação às técnicas precedentes (Mikolov et al. 2013), contornando parcialmente a chamada “maldição da dimensionalidade”. Dessa forma, o artigo é pioneiro em propor uma técnica capaz de representar milhões de palavras no vocabulário, enquanto tentativas anteriores usariam vocabulários de no máximo centenas de milhares de palavras. A ideia de vetores de palavras, ou word2vec (Mikolov et al. 2013), é um importante elemento dos MdL como o BERT da Google e os GPTs da OpenAi. Entretanto, ressaltamos, como Pasquinelli e Joler (2021) argumentam, que à redução de dimensionalidade pode levar a “redução de categorias, o que pode ter um impacto na representação da diversidade social”. Além disso, ela “pode encolher taxonomias e introduzir vieses, normalizando ainda mais a diversidade no mundo e obliterando identidades únicas” (Pasquinelli e Joler 2021:13). Esse aspecto será retomado no capítulo 8.

Voltando à patente, percebemos que a informação digital capturada não se restringe ao conteúdo da mensagem, sendo características da identidade da usuária e metadados das mensagens, outras possibilidades de dados a serem captados pelo método:

Além disso, o sistema de computação também pode incluir a identidade da remetente como uma característica (e.g., na forma de um identificador cifrado gaia [*a hashed gaia ID*, um identificador único para todos os produtos da Google], como o número de telefone do remetente). Outros dados que podem ser incluídos como característica incluem a hora do dia em que a mensagem foi enviada ou recebida. Em outro exemplo, no bloco 104, o sistema de computação também pode processar chamadas que foram recebidas (ou perdidas) de uma destinatária da respectiva mensagem enviada e incluir dados relacionados a essas chamadas como características no vetor de características. Os dados relacionados às chamadas podem ser incluídos como um ou mais recursos que indicam, por exemplo, a usuária que fez

a chamada, se a chamada foi recebida ou perdida, o horário em que a chamada foi recebida ou perdida e assim por diante. (Braun e Patel 2022:18)

Nota-se novamente, o caráter heterogêneo da imagem. O vetor de características pode agregar simultaneamente o conteúdo da mensagem e o contexto relativo ao qual este dado foi produzido. A patente ainda explicita o método de treino, isto é, de criação da imagem no interior do modelo, o que aconteceria de forma “não-obstrusiva” durante o uso de um aplicativo de trocas de mensagens:

No bloco 106 do método 100, o sistema de computação gera dados de treinamento (ou conjuntos de dados de treinamento) que *vinculam uma ou mais frases extraídas da respectiva mensagem e o respectivo vetor de características*. Esses dados de treinamento podem assumir a forma de uma tupla [sequência de elementos separados por vírgula] para cada frase. Em um exemplo, uma respectiva tupla pode assumir a forma [(frase), (ID da remetente), (característica1) (contagem da característica1), (característica2) (contagem da característica2)...] Referindo-se novamente ao exemplo acima, em que a mensagem 144B é a respectiva mensagem, a respectiva tupla pode assumir o formato [which one?, (Tony ID), hey 1, greg 1, lunch 1, plans 1, check 1, sandwich 1, shop 1]. Em geral, no contexto de aplicativos de aprendizado de máquina, cada frase se torna um rótulo [isto é, resultado esperado na saída] e as características (e.g., palavras, ID da remetente, representações Unicode de emoticons, hora do dia, chamadas de voz e similares) nas mensagens anteriores formam o vetor de características [ou seja, as entradas] (Braun e Patel 2022:18, grifo nosso).

A partir disso, o método pode “mapear frases particulares com o vetor de característica de mensagens recebidas para identificar as frases que podem ser mais relevantes como respostas em potencial” (Braun e Patel 2022:18). Dessa forma, mensagens pré-definidas também podem ser sugeridas à usuária, mostrando mais uma vez o acoplamento treino–inferência/incidência:

Em uma modalidade, um sistema de computação treina um aplicativo de aprendizado de máquina com dados de treinamento *para determinar mensagens “prontas” inteligentes que podem ser selecionadas e inseridas em uma mensagem a ser enviada*. O sistema de computação pode desenvolver os dados de treinamento processando o histórico de mensagens da usuária, por exemplo. Em um exemplo, *quando a usuária recebe uma comunicação ou mensagem em um sistema ou dispositivo de computação (e.g., um telefone celular), o aplicativo de aprendizado de máquina pode fornecer à usuária uma ou mais respostas predefinidas contextualmente relevantes e personalizadas para a comunicação recebida*. Uma lista de mensagens predefinidas pode ser fornecida a uma usuária em uma exibição gráfica. A usuária pode então selecionar uma resposta predefinida para enviar em resposta à comunicação recebida (Braun e Patel 2022:17, grifo nosso).

Assim, ressaltamos como a captura é uma operação importante para a criação de uma imagem, neste caso, uma representação das trocas de mensagens de uma usuária em particular. Mas, além disso, também fica claro como a imagem serve, em última instância, para modular a atividade da usuária, sugerindo mensagens na interface gráfica e, portanto, facilitando um fluxo que se adeque a reproduzir as sugestões e desincentivando a espontaneidade e a divergência. Captura-imaginação-modulação funcionam conjuntamente criando um mecanismo de realimentação que funciona de uma maneira em que

a usuária é apenas um de seus componentes. A imaginação é a operação que define a aprendizagem de máquina como um agenciamento de representação: como a perspectiva renascentista ou a fotografia, o AdM se apresenta como representação objetiva do mundo, embora tanto um quanto outro sejam tanto uma representação do mundo, quanto uma construção performativa dele. A imaginação é a criação de uma imagem que conecta a captura à modulação: dados capturados em parâmetros numéricos são relacionados no modelo criando uma capacidade de inferência e de incidência informacional que são utilizados na modulação da usuária.

Capítulo 5

Modulação

O controle da realidade nascente pela realidade antiga define o modo social do processo informacional: o modelo triódico [*triodique*] é o análogo funcional de uma estrutura social.

A amplificação nos processos de informação

Gilbert Simondon

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

Post scriptum sobre as sociedades de controle

Gilles Deleuze

5.1 Molde e modulação, disciplina e controle

De forma ampla, modulação é uma ação de mudança entre modos, podendo expressar distintos efeitos relativos à variação. Em seu conhecido “Post scriptum sobre as sociedades de controle”, Deleuze (2013), inspirado provavelmente na filosofia da técnica de Gilbert Simondon, utiliza o termo modulação para descrever a maneira de exercício de poder própria do final do século XX e que pode ser estendida para a contemporaneidade (Lazzarato 2006; cf. Souza, Silveira, e Avelino 2021). O breve ensaio é dedicado a caracterizar (apresentando “histórico”, “lógica” e “programa”) as chamadas “sociedades de controle”, formas contemporâneas de coletivo social que sucederiam o que Foucault chamou de “sociedades disciplinares” (e.g. Foucault 2021). Por conta disso, sua caracterização da modulação, operação própria às sociedades de controle, se baseia fortemente em uma comparação dela com a operação de molde, própria às sociedades disciplinares¹.

¹Entretanto, não há nisso uma pretensão de opor controle e disciplina, modulação e molde: as sociedades de controle seriam aquelas em que, embora o foco e a novidade do exercício de poder recaiam sobre o controle, ainda subsistem processos

Em Simondon (2020b), modulação e molde são tomados como os termos da operação mais geral de modelagem: “O molde e o modulador são casos extremos, mas a operação essencial de tomada de forma cumpre-se do mesmo jeito; ela consiste no estabelecimento de um regime energético, durável ou não. *Moldar é modular de maneira definitiva; modular é moldar de maneira contínua e perpetuamente variável.*” (Simondon 2020b:52, grifo nosso). Exemplo do primeiro, é a feitura de tijolos de barro a partir de um molde de madeira; enquanto que, do segundo, seria o funcionamento de válvulas eletrônicas, em que um sinal de informação controla uma grade que faz variar continuamente o fluxo de elétrons². Mas Simondon constata que muitas outras operações técnicas estão em uma situação intermediária, exemplos citados são uma feira e um laminador, que são moldes contínuos operados em etapas: “os laminadores que produzem a chapa estriada *modulam* a matéria, ao passo que um laminador liso somente a *modela*. *Moldagem e modulação* são os dois casos-limite, e a *modelagem* é o caso médio” (Simondon 2020b:52)³.

Já para Deleuze, a modulação é descrita como uma forma aparentemente sutil de poder ou de potência que age influenciando continuamente, fazendo tender ou mesmo forçando imperceptivelmente para uma determinada situação. Em suas palavras: “Os confinamentos são moldes, distintas moldagens, mas os controles são uma modulação, como uma moldagem auto-deformante que mudasse continuamente, a cada instante, ou como uma peneira cujas malhas mudassem de um ponto a outro” (Deleuze 2013:221). Em oposição aos internatos, os diversos meios de confinamento próprios da disciplina, em que o indivíduo é separadamente tratado como variável independente⁴; “[...] os diferentes modos de controle, os controlatos, são variações inseparáveis, formando um sistema de geometria variável cuja linguagem é numérica [*numérique*, digital] (o que não quer dizer necessariamente binária).” Portanto, há uma noção de funcionamento interligado de diferentes esferas de controle que operam por meio de uma linguagem digital unificada. Essa linguagem unificada, em nosso entendimento, se relaciona com a

de soberania e de disciplina. Também por isso, não tentaremos resolver a controvérsia que se apresenta sobre se o controle seria algo de natureza distinta da disciplina ou uma nova forma de disciplina (cf. Dardot e Laval 2016:215, *et seq*). Nosso foco é derivar do conceito deleuziano uma caracterização positiva (e não opositiva/negativa) de modulação para a sociologia da tecnologia e em especial para a descrição do funcionamento do AdM.

² “[A] diferença entre os dois casos reside no fato de que, para a argila, a operação de tomada de forma é finita no tempo: ela tende bem lentamente (em alguns segundos) a um estado de equilíbrio, e daí o tijolo é desmoldado; ao desmoldá-lo, utiliza-se o estado de equilíbrio quando este é atingido. No tubo eletrônico, emprega-se um suporte de energia (a nuvem de elétrons num campo) de uma inércia bem fraca, tanto que o estado de equilíbrio (adequação entre a repartição dos elétrons e o gradiente do campo elétrico) é obtido num tempo extremamente curto relativamente ao anterior (alguns bilionésimos de segundos num tubo de grande dimensão, algumas dezenas de bilionésimos de segundo nos tubos de pequena dimensão). Nessas condições, o potencial da grade de comando é utilizado como *molde variável*; a repartição do suporte de energia segundo esse molde é tão rápida que ela se efetua sem atraso apreciável para grande parte das aplicações: o molde variável serve, então, para fazer variar no tempo a atualização da energia potencial de uma fonte; não se detém quando o equilíbrio é atingido, mas continua-se modificando o molde, isto é, a tensão na grade; a atualização é quase instantânea, nunca há parada para desmoldagem, pois a circulação do suporte de energia equivale a uma *desmoldagem permanente*; um modulador é um *molde temporal contínuo*” (Simondon 2020b:51–52).

³ Consideramos a analogia com os algoritmos que são chamados de modelos, como no caso do AdM, profícua: modelos clássicos de IA, como os sistemas especialistas e outros baseados em árvores ou que produzem resultados determinísticos, isto é, sem que uma mesma entrada possa resultar em valores diferentes, se aproximariam mais com *moldes*. Já os algoritmos de AdM, por possuírem um dinamismo, realimentação com a usuária e envolverem uma estocasticidade na etapa de treinamento, podem resultar em valores diferentes para uma mesma entrada, portanto se aproximando mais com *moduladores*. Comparar também com a discussão feita por Rouvroy e Berns (2015) sobre as aproximações e divergências entre a estatística clássica e a algorítmica.

⁴ As variáveis independentes são as condições, ou as entradas de uma função, que influenciam o resultado de uma variável dependente, a saída de uma função.

noção transversal de informação e sua captura na forma de códigos discretizados: informação genética e informação digital (Bruno e Rodríguez 2022; Garcia dos Santos 2007). Acreditamos que o “Manifesto ciborgue” de Haraway (2009), complementa o p.s. deleuziano de uma maneira bastante interessante nesse aspecto. Para a autora, há uma ordem mundial emergente, a “informática da dominação”, em que:

[a]s estratégias de controle irão se concentrar nas condições e nas interfaces de fronteira, bem como nas taxas de fluxo entre fronteiras, e não na suposta integridade de objetos supostamente naturais. [...] As estratégias de controle serão formuladas em termos de taxas, custos de restrição, graus de liberdade. Os seres humanos, da mesma forma que qualquer outro componente ou subsistema, deverão ser situados em uma arquitetura de sistema cujos modos de operação básicos serão probabilísticos, estatístico. Nenhum objeto, nenhum espaço, nenhum corpo é, em si, sagrado; *qualquer componente pode entrar em uma relação de interface com qualquer outro desde que se possa construir o padrão e o código apropriados, que sejam capazes de processar sinais por meio de uma linguagem comum*. A troca, nesse mundo, transcende à tradução universal efetuada pelos mercados capitalistas, tão bem analisada por Marx (Haraway 2009:62 grifo nosso).

Com o advento da noção de informação, “as ciências da comunicação e as biológicas modernas são construídas por uma operação comum – *a tradução do mundo em termos de um problema de codificação*, isto é, a busca de uma linguagem comum na qual toda a resistência ao controle instrumental desaparece e toda a heterogeneidade pode ser submetida à desmontagem, à remontagem, ao investimento e à troca” (Haraway 2009:64). Neste contexto, o humano também se torna atravessado por novos modos de controle. A modulação se constitui como uma malha geral e variável de afetação (“moldagem auto-deformante”) e de ação singularizada (“peneira cujas malhas mudassem de um ponto a outro”). Ela é um modo de poder que opera por comunicação instantânea e controle contínuo (Deleuze 2013:220), modulando de forma particular cada indivíduo, afetando-o sutilmente, influenciando processualmente inclusive a sua tomada de forma. *Controle de informação, mais que disciplina de forma*.

Por esse motivo, além da comunicação instantânea e controle contínuo, acreditamos que a ação diretamente exercida sobre o meio (e indireta sobre o indivíduo) é outra característica da modulação. Na disciplina, a palavra de ordem e a anatomopolítica agem diretamente sobre o corpo, enquanto que a modulação age mais como um controle de fluxos ou de acessos, como mostra o caso da tornozeleira eletrônica ou do cartão dividual da epígrafe, metáforas que, hoje, podem ser estendidas para se referir aos smartphones. No caso do salário por mérito, exemplo apresentado por Deleuze, esse aspecto também fica claro: trata-se sobretudo de estruturar um campo concorrencial que induz uma autorresponsabilização neoliberal, tal qual analisada por Dardot e Laval (2016). Essa situação de “perpétua metaestabilidade” significa salários flexíveis e singularizados, com uma lógica muito mais imperceptível e muito menos explicável e que ressoa com o sistema de competição neoliberal (Deleuze 2013). No caso das atuais trabalhadoras plataformizadas ou uberizadas a situação é ainda mais evidente.

Já no que diz respeito à relação indivíduo-coletivo, ela também se altera nas sociedades de controle: a *cifra* (biometria, usuária/senha, PIN, *hashes* e até mesmo o perfil) representa indivíduos “dividuais” e divisíveis (pensar nos múltiplos perfis para uma mesma pessoa, mas também em múltiplas pessoas que acessam um mesmo perfil online. Os padrões estatísticos de perfilamento captam dimensões parciais de indivíduos, nunca ele totalmente) enquanto que em relação à coletividade, ela se expressa como

amostras e bancos (estatísticas, que substituem a ideia de massa nas sociedades disciplinares). O aprendizado de máquina pode ser visto como uma materialização dessa lógica, na medida em que já funciona operando sobre sujeitos individuais e de bases de dados.

Além do dinheiro, que perde seu lastro metálico e se torna virtual⁵. E do humano que passa de um “produtor descontínuo de energia” para um “feixe contínuo”, “ondulatório”, não mais “confinado”, mas sim “endividado”. Há uma mudança das próprias máquinas, da tecnologia. Das máquinas simples (alavancas, roldanas) das sociedades de soberania, passamos para máquinas energéticas da disciplina. “[A]s sociedades de controle operam por máquinas de uma terceira espécie, máquinas de informática e computadores, cujo perigo passivo é a interferência, e o ativo a pirataria e a introdução de vírus” (Deleuze 2013:223). Embora seja questionável que ainda hoje os vírus e a pirataria sejam os principais “perigos” do computador digital, esta sendo em larga medida derrotada por leis de propriedade intelectual e novos modos de capitalização da cultura⁶, consideramos que o mais importante são as possibilidades de relações sociais que tais máquinas propiciam ou combatem.

Antes de seu p.s., em 1980, Deleuze e Guattari já haviam notado, à luz das transformações tecnológicas recentes, que a dominação capitalista seria composta não apenas de uma exploração econômica via sujeição social, mas de novas formas de servidão maquínica⁷:

Se as máquinas motrizes constituíram a segunda idade da máquina técnica, as máquinas da cibernética e da informática formam uma terceira idade que recompõe *um regime de servidão generalizado*: “*sistemas homens-máquinas*”, reversíveis e recorrentes, substituem as antigas relações de sujeição não reversíveis e não recorrentes entre os dois elementos; a relação do homem e da máquina se faz em termos de comunicação mútua interior e não mais de uso ou de ação. Na composição orgânica do capital, o capital variável define um regime de sujeição do trabalhador (mais-valia humana) tendo por quadro principal a empresa ou a fábrica; mas, quando o capital constante cresce proporcionalmente cada vez mais, na automação, encontramos uma nova servidão, ao mesmo tempo que o regime de trabalho muda, que a mais-valia se torna maquínica e que o quadro se estende à sociedade inteira. Dir-se-á mesmo que um pouco de subjetivação nos distanciava da servidão maquínica, mas que muito nos reconduz a ela. *Sublinhou-se recentemente a que ponto o exercício do poder moderno não se reduzia à alternativa clássica “repressão ou ideologia”, mas implicava processos de normalização, de modulação, de modelização, de informação, que se apoiam na linguagem,*

⁵Laymert Garcia dos Santos, em suas “Considerações sobre a realidade virtual”, escreve sobre a financeirização: “Nas telas dos computadores, as posições dos derivativos mudam a cada minuto, o futuro modelando o presente. A evolução dos valores reflete evidentemente os riscos e recursos que transitam no espaço econômico, no mundo das mercadorias; mas reflete tanto, ou mais, o que se passa nas redes digitais, porque estas produzem *uma imagem* dos mercados em constante mutação, imagem que, por sua vez, afeta os valores dos papéis, ao permitir a imediata comparação da *performance* de cada um com a de todos os outros, suscitando reavaliações, depreciações e valorizações” (Garcia dos Santos 2003b:119).

⁶Penso aqui nas plataformas de filmes (como a Netflix) e músicas (como a Spotify), por exemplo, que, de uma maneira ainda mais alinhada com a descrição de Deleuze, lucram por meio do controle de acesso de uma usuária a determinadas mídias.

⁷Diferenciando ambas, os autores escrevem: “Há servidão quando os próprios homens são peças constituintes de uma máquina, que eles compõem entre si e com outras coisas (animais, ferramentas), sob o controle e a direção de uma unidade superior. Mas há sujeição quando a unidade superior constitui o homem como um sujeito que se reporta a um objeto tornado exterior, seja esse objeto um animal, uma ferramenta ou mesmo uma máquina: o homem, então, não é mais componente da máquina, mas trabalhador, usuário..., ele é sujeito à máquina, e não mais submetido pela máquina” (Deleuze e Guattari 2011–2012, v. 5, p. 167-8). Lazzarato (2010) também explora o conceito deleuze-guattariano construindo sobre essas noções e suas implicações para o capitalismo contemporâneo.

na percepção, no desejo, no movimento, etc, e que passam por microagenciamentos (Deleuze e Guattari 2011–2012, v. 5, p. 169-70, grifo nosso)

É justamente esse regime de servidão maquínica, em que o humano funciona com a máquina, como componente desta, que explica a posição que a usuária ocupa no AdM. Em um sentido, é a usuária que mobiliza um algoritmo para conseguir uma sugestão, classificação ou predição; mas, em outro sentido, é a máquina que mobiliza a usuária no processo de treinamento, a usuária funcionando para a máquina. O exemplo oferecido por Deleuze e Guattari acerca da TV é de uma atualidade surpreendente, antecipando movimentos que são, hoje, muito mais perceptíveis no modo como participamos de processos envolvendo smartphones ou outros dispositivos conectados à internet:

É esse conjunto que comporta ao mesmo tempo a sujeição e a servidão, levadas aos extremos, como duas partes simultâneas que não param de se reforçar e de se nutrir uma à outra. Por exemplo: somos sujeitados à televisão na medida em que fazemos uso dela e que a consumimos, nessa situação muito particular de um sujeito do enunciado que se toma mais ou menos por sujeito da enunciação (“os senhores, caros telespectadores, que fazem a televisão...”); a máquina técnica é o meio entre dois sujeitos. Mas somos submetidos pela televisão como máquina humana na medida em que os telespectadores são não mais consumidores ou usuários, nem mesmo sujeitos que supostamente a “fabricam”, mas peças componentes intrínsecas, “entradas” e “saídas”, *feed-back* ou recorrências, que pertencem à máquina e não mais à maneira de produzi-la ou de se servir dela. Na servidão maquínica há tão-somente transformações ou trocas de informação das quais umas são mecânicas e outras humanas. [...] Temos o privilégio de nos submeter, através das mesmas coisas e dos mesmos acontecimentos, às duas operações ao mesmo tempo. Sujeição ou servidão formam dois polos coexistentes, antes que duas fases (Deleuze e Guattari 2011–2012, v. 5, p. 170).

O advento das sociedades de controle é, enfim, uma mudança do próprio capitalismo que passa a centrar seu funcionamento nos serviços e na recombinação (sobreprodução), na valorização por meio de ações no mercado financeiro, com foco no produto e não na produção, instituindo a empresa como modelo subjetivo e institucional, instaurando a centralidade do marketing (como vemos no próprio modelo de propagandas da Google), e conquistando novos terrenos por meio de regulações de novos fluxos informacionais progressivamente capturados (Deleuze 2013:223–24). As instituições modernas (prisão, escola, hospital, empresa) sofrem uma mudança da disciplina e do molde para o controle e a modulação, de modo que assistimos à “implantação progressiva e dispersa de um novo regime de dominação” (Deleuze 2013:225).

Maurizio Lazzarato (2006), ao relacionar as sociedades de controle com o contexto do início do século XXI, adiciona ainda novos elementos à sua caracterização. Nesse sentido aproxima a ideia de coletividades dividuais à noção de *público* da comunicação social, isto é, a que denota um conjunto de pessoas cuja memória e atenção se mostram disponíveis para determinada mídia. Da mesma forma que bases de dados contém informações digitais constituídas de parcialidades individuais, os públicos reúnem porções do tempo e atenção de sujeitos que não são exclusivos do ponto de vista das *media*: um indivíduo pode pertencer a vários públicos simultaneamente (assistindo a dois ou mais programas de TV, assinando dois ou mais jornais, etc.). Desse modo, “(...) as técnicas de controle e de constituição

dos públicos colocam em primeiro plano o tempo e suas virtualidades. O público se constitui através de sua presença no tempo” (Lazzarato 2006:75). A partir disso, e com a ajuda de tecnologias da informação, as técnicas de poder se constituiriam principalmente de modos de interferir na opinião pública, agir à distância sobre as mentes e controlar a percepção e a inteligência coletivas. O conjunto de técnicas de controle “(...) se exerce sobre os cérebros, atuando em primeiro lugar sobre a atenção, para controlar a memória e sua potência virtual” (Lazzarato 2006:86). Apesar do foco nas capacidades mentais coletivas, próprio da escola do “capitalismo cognitivo”, esses adendos nos são interessantes, pois entendemos que, cada vez mais os circuitos de memória e de atenção passam pelo smartphone e outros dispositivos que são capturados, imaginados e modulados por algoritmos de aprendizado de máquina.

Além disso, também consideramos uma contribuição valiosa do sociólogo, aquela que vê as sociedades de controle como caracterizadas também pela oferta de “mundos”, ou seja, de possibilidades de vida. Mas de mundos que são sempre já fabricados (pelas especialistas e pelas autoras, ofertados através do marketing). Ao sujeito bastaria optar entre eles: “(...) nossa ‘liberdade’ é exercida exclusivamente para escolher dentre possíveis que outros instituíram e conceberam. Ficamos sem o direito de participar da construção dos mundos, de formular problemas e de inventar soluções, a não ser no interior destas alternativas já estabelecidas” (Lazzarato 2006:101–2). Em relação às técnicas analisadas neste trabalho, essa descrição nos ajuda a entender como as sugestões e classificações oferecidas pelos algoritmos substituem, ao menos parcialmente, a liberdade de ação da usuária. Para o autor, podemos ainda dizer que a modulação tem por objeto final o vivo (sua atenção e memória, fontes da atividade mental e da constituição subjetiva) e a vida (a atualização de mundos possíveis) (Lazzarato 2006). Aliado às reflexões de Garcia dos Santos (2003a) sobre a informação digital e genética, percebemos que essa constatação é bastante válida, pois incidindo sobre a informação, a modulação age mesmo controlando os processos de tomada de forma que constituem, não só uma dimensão do vivo⁸, mas também aspectos da própria materialidade (o físico) e da sociedade (o coletivo) (cf. Simondon 2020b).

5.2 A modulação nos sistemas digitais

Mas se Simondon e, sobretudo, Deleuze utilizam “modulação” para descrever processos sociais, não é a partir de um uso meramente metafórico do termo. Paralelo ao uso filosófico que eles fazem, a palavra “modulação” também é utilizada para descrever processos em outros domínios do conhecimento. O uso mais conhecido de modulação é provavelmente o das telecomunicações, por exemplo a modulação por frequência e por amplitude, utilizadas respectivamente nas antenas/sistemas de rádio por modulação por frequência (FM, *frequency modulation*) e modulação por amplitude (AM, *amplitude modulation*) (figura 5.1). Nesse contexto, modulação diz respeito ao processo de transferência de informação de um sinal elétrico de entrada (modulante), para um outro sinal elétrico (portador), gerando um sinal de saída (modulado). Assim, a modulação altera propriedades do sinal portador (a amplitude ou a frequência), que reproduz o potencial informacional do sinal modulante.

Mesmo Simondon (2020a) apresenta a operação de modulação partindo da tecnologia, em um nível ainda mais “elementar” da técnica: a do funcionamento das válvulas triódicas, precursoras dos atuais transistores de estado sólido, que são componente fundamental das tecnologias eletrônicas digitais.

⁸No que diz respeito à informação digital, apontamos que nossa vida é, cada vez mais, mediada por tecnologias digitais e algoritmos constituindo algo que podemos chamar, nas palavras de Pedro Ferreira, de “vida eletrônica”, isto é, a vida em que participamos “como padrões eletromagnéticos no interior dos circuitos eletrônicos” (Ferreira 2024:45min12s).

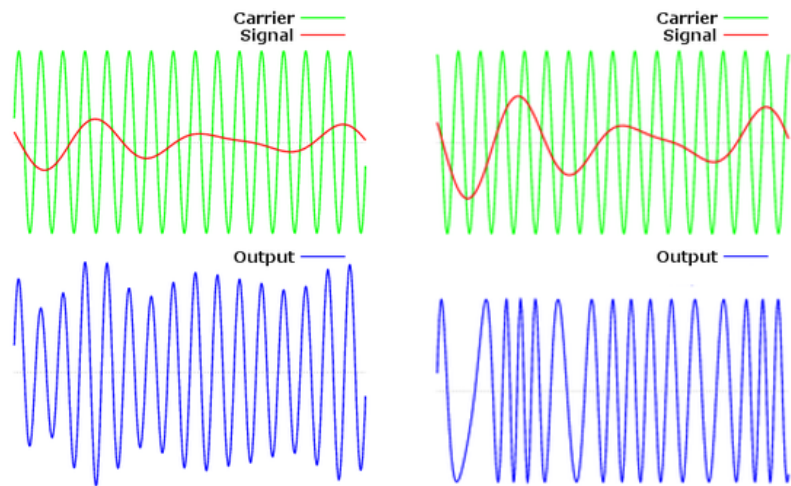


Figura 5.1: Respectivamente: gif animado comparando a modulação por amplitude e por frequência de um sinal de entrada; modulação por amplitude e modulação por frequência. Fontes: <https://commons.wikimedia.org/wiki/File:Amfm3-en-de.gif>, <https://commons.wikimedia.org/wiki/File:Amplitude-modulation.png> e <https://commons.wikimedia.org/wiki/File:Frequency-modulation.png> (adaptado) (acesso: 10/06/2024).

Mais interessante ainda é que sua análise faz convergir o funcionamento técnico e social do que ele chama de “amplificação moduladora”, isto é, daquela modulação que, diferentemente das ondas de rádio ou de um diodo passivo, opera por meio de um elemento energeticamente ativo que controla, faz nascer, um novo sinal (amplificado). É um esquema dialético que Simondon propõe nesse texto. Um triodo pode funcionar como amplificação transdutiva, “o modo mais elementar da operação de informação” em que “há transferência gradual, alimentada energeticamente pela mudança de estado no próprio local” (Simondon 2020a:286, 287), ou seja, uma propagação por reação em cadeia, uma amplificação positiva. Já a amplificação moduladora é negativa, é uma amplificação por controle de uma fonte de energia. Por fim, “há possibilidade de organização, quando há compatibilidade entre esses dois processos”, uma “síntese real”, de modo que a amplificação organizadora acontece quando “a série transdutiva de decisões é modulada por uma informação tomada sobre o conjunto do campo, que intervém, assim, como uma totalidade no percurso que o recobre e o transforma” (Simondon 2020a:296, 295).

Sobre a amplificação moduladora, que nos interessa particularmente, o autor especifica: “O esquema da amplificação moduladora é obtido domesticando-se a propagação transdutiva; isto é, dominando-a e a alimentado localmente, para fazê-la produzir e trabalhar em condições regulares” (Simondon 2020a:289). No triodo, isso ocorre por meio da incidência de um sinal de informação de entrada que interage com um sinal de energia (uma fonte de alimentação ou energia potencial) para fazer um efeto agir sobre uma carga:

A entrada age como um isolamento variável que se interpõe entre a fonte de energia e a carga; esse isolamento não acrescenta energia ao quase-sistema, tampouco a suprime; ele governa, controla o regime de mudança de estado de energia potencial, a cada instante, seja por escolha entre dois valores [análogo ao regime de saturação e corte na operação digital de um transistor] (tudo ou nada, regime pleno ou nulo, abertura ou fechamento do circuito reunindo a fonte de energia e a carga), seja por escolha entre uma infinidade de valores entre um máximo e um mínimo, entre a saturação e a interrupção [análogo ao regime de amplificação na operação analógica de um transistor] (Simondon 2020a:289).

Os termos “dominar”, “governar” e “controlar” explicitam que a natureza dessa operação é assimétrica: “[o] modulador está constantemente em vias de decidir a partir da incidência de informação; essa incidência de informação governa, instante por instante, o regime de transformação de energia potencial em trabalho” (Simondon 2020a:290). No domínio biológico e social, a modulação pode ser utilizada para explicar funcionamentos em que um processo é estabilizado, como por exemplo a autorregulação e a homeostase na biologia. E ainda:

Em lugar de ser, como a transdução, movimento rumo à zona rica em energia potencial, a partir do domínio já estruturado, a modulação realiza uma submissão de uma série nascente do devir ao ato final de uma série anterior; *ela determina o novo segundo a estrutura do antigo* (ação de relé), o condiciona. Enquanto a transdução é orientada para o futuro, *a modulação é uma vitória do antigo sobre o novo, uma reciclagem da estrutura antiga*. A modulação é o ato crítico, redutor, no sentido próprio do termo; *ela é o modelo do controle, da autoridade, no domínio social*, assim como da operação destinada a evitar a variação fortuita, a detectar o erro (Simondon 2020a:298, grifo nosso)

Nos sistemas algorítmicos, percebemos que a modulação se expressa não só via controle de acesso automatizado e particularizado à processos de informação digital, como no caso da mediação de conteúdo no *feed* de redes sociais, mas também pela organização individualizada de elementos semióticos, afetando a atividade da usuária indiretamente pela alteração direta do meio: por exemplo, a modificação de UI incluindo ou excluindo botões em sites e aplicativos. Descrevendo situações similares, Rouvroy e Berns (2015) propõem a noção de “governamentalidade algorítmica” como forma de caracterizar modos de governo online que são aparentemente anormativo e apolítico e que “[...] repousa[m] sobre a coleta, agregação e análise automatizada de dados em quantidade massiva de modo a modelizar, antecipar e afetar, por antecipação, os comportamentos possíveis [de usuárias]” (Rouvroy e Berns 2015:42). Um exemplo de modulação apresentado no artigo são os preços dinâmicos na compra de passagens aéreas, que mudam a depender do comportamento da usuária na internet (Rouvroy e Berns 2015:44). Já para Hui (2015), os algoritmos participam das práticas sociais em sistemas digitais ao inscreverem em si regras teleológicas que são recursivamente aplicadas na modulação de relações sociais: “[...] compras de um click na Amazon baseadas em recomendações, ou o botão de ‘like’ do Facebook como forma simbólica de participação são precisamente mecanismos que buscam replicar tipos particulares de interações pessoais em formas codificadas e em escalas sociais ascendentes” (Hui 2015:87). Sérgio Amadeu da Silveira, também relaciona modulação e sistemas algorítmicos, mostrando que ela serve principalmente ao marketing, por “encurtar a realidade e a multiplicidade de discursos” e enfatiza que “[u]m dos principais modos de controle que os gestores das plataformas exercem sobre seus usuários é a modulação das opções e dos caminhos de interação e de acesso aos conteúdos publicados” (Amadeu da Silveira 2019:21).

De especial relevância para o nosso trabalho, é interessante notar como, retomando Simondon (2020b:52) na sua apresentação da moldagem e da modulação como casos limites da operação de modelagem, se a modelagem estatística clássica, não adaptativa, visava prover uma média em uma operação estática e finita no tempo (mais próximo do molde), a prática do perfilamento, consequência direta de sistemas envolvendo AdM, se mostra como uma operação adaptativa variável e contínua no tempo, uma modulação (Rouvroy e Berns 2015). Já Bruno e Rodríguez (2022), mostram como nas práticas digitais,

a dimensão individual e dividual coexistem. Em relação ao “perfil”, elas mostram que, por um lado, há uma noção de perfil como uma exposição parcial e relacional – e, portanto, dividual – de si mesmo, como no caso dos perfis em redes sociais; por outro lado, há o perfilamento operado por empresas que é a prática de antecipação de potencial e performatividade a partir da captura da informação digital igualmente dividual. Entretanto, “os próprios mecanismos de conhecimento e controle algorítmico operam com essa composição, capturando dividualidades sem, no entanto, deixar de apelar para o indivíduo, seja como alvo [da modulação] ou como responsável econômica e juridicamente por contratos, escolhas e ações” (Bruno e Rodríguez 2022:40).

Fenômenos relacionados com o big data e a IA amplificam as possibilidades de modulação em sistemas digitais, pois sua própria arquitetura privilegia a reprodução do antigo e a supressão do novo, na medida em que funcionam baseados em dados coletados do passado⁹. Em consonância com a análise que aproxima o funcionamento do AdM e a concepção deleuze-simondoniana da modulação, Pasquinelli e Joler (2021:1265) concluem que, “[c]omo uma técnica de compressão de informação, o aprendizado de máquina automatiza a ditadura do passado, de taxonomias passadas e de padrões comportamentais sobre o presente. Esse problema pode ser nomeado de regeneração do passado – a aplicação de uma visão espaço-temporal homogênea que restringe a possibilidade de novos eventos históricos”.

Desse modo, acreditamos que tanto a nível material elementar, quanto a nível sistêmico algorítmico, os meios digitais agem largamente por meio da operação de modulação. Isso já ocorre no transistor, que utiliza sinais de informação antigos para a construção de sinais ulteriores, mas também nos algoritmos e em especial na IA: ao utilizar dados do passado para antecipar, modelar e afetar processos informacionais futuros.

Analisando especificamente patentes de tecnologias que operam modulação depositados pela Facebook Inc. entre 2014 e 2018, Débora Machado (2019, 2020) propõe uma tipologia com seis categorias operativas de modulação em sistemas digitais. Embora ao longo das análises tenha sido difícil separar as nossas patentes em categorias estanque, decidimos utilizar os tipos de modulação propostos por Machado (2019) como ponto de partida para observar recorrências nas patentes estudadas. Assim, a ação operada pelo AdM foi inicialmente classificada nesses termos. Da tabela 5.1 é possível observar que todas as patentes analisadas incluem elementos de “Predição, antecipação e inferência” e “Possibilidades para a coleta de dados”. O primeiro não é surpreendente, visto que nosso recorte (o AdM) se restringe a técnicas que incluem tal elemento. Já o segundo, corrobora com nosso argumento de que toda técnica de aprendizagem de máquina envolve a captura, da qual a coleta ou captação de dados é parte fundamental. Além disso, todas as patentes analisadas possuem elementos de “Recomendação que leva a ação” e/ou “Decidindo pelo usuário”, o que também não nos surpreende por mais ou menos corresponder à definição inicial de modulação proposta para o recorte, embora o exercício analítico de diferenciar uma categoria de outra possa evidenciar características distintas. Por fim, nenhuma patente analisada envolvia elementos explícitos de “Análise de emoções”, mostrando que esta talvez seja uma particularidade do recorte de Machado (2019) (redes sociais, ou da Facebook em relação à Google), em-

⁹Embora, como já notado, muitos aspectos da IA não são totalmente novos e alguns de seus efeitos podem ser explicados do ponto de vista do automatismo que ela opera (Gonçalves 2022). Já para Mühlhoff (2020), a IA contemporânea se distancia da noção de automatismo -- ao menos tal como é criticada em Simondon --, pois ela seria composta por trocas entre componentes humanos e não-humanos que formariam um todo funcional. Embora Mühlhoff não considere um aspecto somente negativo, o funcionamento da IA contemporânea é descrito por ele como um agregado humano-máquina reciprocamente co-dependente que “(...) modula os movimentos, o conhecimento, o bem estar e os afetos da usuária” (Mühlhoff 2020:1880).

bora em todo o caso, fique evidente que as patentes incidem, por vezes até nomeando explicitamente, no desejo e em outros funcionamentos psíquicos, de modo que ainda assim podemos localizá-las naquilo que Bruno denominou uma “economia psíquica dos algoritmos” (Bruno et al. 2019).

Tabela 5.1: Elementos da tipologia de modulação proposto em Machado (2019): M1 – “Recomendação que leva a ação”; M2 – “Decidindo pelo usuário”; M3 – Predição, antecipação e inferência; M4 – “Atribuindo valor à interação”; M5 – “Análise de emoções” e M6 – “Possibilidades para coleta de dados”.

Número	M1	M2	M3	M4	M5	M6
11366812		•	•			•
11252114	•		•			•
11222637		•	•			•
11115695	•	•	•	•		•
11108889	•	•	•	•		•
10970096	•		•			•
9817869	•	•	•	•		•
9244100	•		•			•
8645390	•		•			•

Diante disso, percebemos que o desenvolvimento das tecnologias digitais e dos sistemas algorítmicos promovem a intensificação dos processos descritos por Simondon (2020a) e Deleuze (2013), fazendo da modulação categoria explicativa privilegiada para uma análise da ação sociotécnica do AdM. Como exposto, todos os documentos analisados incluem elementos que visam um controle de processos de tomada de forma via mudanças contínuas e em tempo-real, alterando o meio para propiciar indiretamente ações, percepções e afecções nas usuárias. Passemos agora para a análise mais detida de um exemplo.

5.2.1 Exemplo na patente

A patente *Systems and methods for prioritizing notifications on mobile devices* (US9817869) (Aradhye et al. 2017) visa determinar a prioridade de notificações utilizando AdM a fim de substituir a apresentação cronológica de notificações em um dispositivo móvel por uma baseada na relevância estimada. Nesse sentido, visa substituir a agência explícita da usuária em escolher as notificações que vai acessar, por uma agência maquínica que automatiza esta ação.

Para que isso seja possível, o sistema captura a série histórica de ações da usuária, em que “a ausência de interação da usuária com uma notificação, a seleção de uma notificação, ou a recusa [*dismissal*] de uma notificação” é utilizada para treinar o modelo de AdM:

Por exemplo, se uma usuária clica em uma notificação de mensagem de texto de John, o sistema pode modificar o modelo de ranqueamento para que notificações de mensagens de texto de John tenham uma pontuação de prioridade maior. Notificações com maior pontuação de prioridade são consideradas mais importantes ou de maior prioridade que notificações com uma pontuação de prioridade menor. Continuando com o exemplo, se uma notificação de mensagem de texto de John é fechada (sem selecionar para ver a men-

sagem de texto), o modelo de ranqueamento pode ser modificado para que mensagens de texto de John tenham uma pontuação de prioridade menor (Aradhya et al. 2017:9).

Outras informações capturadas incluem o horário de recebimento da notificação, a localização do dispositivo da usuária, os aplicativos anteriormente acessados pela usuária e a orientação do dispositivo da usuária (Aradhya et al. 2017:10–11). Além disso, as próprias características da notificação são captadas:

As características da notificação podem incluir, entre outras, o aplicativo de origem da notificação, o remetente dos dados associados à notificação (nome e/ou número e/ou código de identificação da remetente), a hora em que a notificação ou os dados associados à notificação foram recebidos ou enviados, a pontuação de prioridade da notificação e o conteúdo da notificação (por exemplo, o texto ou a imagem da notificação ou dos dados associados). Para continuar com o exemplo, *se uma notificação de mensagem de texto associada a uma mensagem de texto recebida de Jane às 17h for selecionada (uma indicação positiva), o modelo de classificação poderá ser atualizado de modo que as notificações recebam uma pontuação de prioridade mais alta quando estiverem associadas a Jane, recebidas às 17h ou próximo a elas e/ou estiverem associadas a mensagens de texto* (ou qualquer combinação delas, como mensagens de texto recebidas às 17h ou próximo a elas). Assim, ao entender que a seleção pela usuária da notificação de mensagem de texto associada à mensagem de texto da Jane indica que a notificação foi importante para a usuária, o sistema pode priorizar automaticamente as notificações relacionadas (com base em uma ou mais características da notificação), de modo que as notificações relacionadas recebam pontuações de prioridade que reflitam adequadamente sua importância para a usuária (Aradhya et al. 2017:10, *grifo nosso*).

Desse modo, a patente prevê a modificação do modelo de ranqueamento baseado na interação da usuária com a notificação e nas características desta notificação, e também nas características das outras notificações sendo apresentadas ao mesmo tempo:

Por exemplo, uma característica é a origem do aplicativo. Um dispositivo móvel emite uma notificação de mensagem de texto e uma notificação de chamada telefônica, e uma usuária seleciona a notificação de texto. A modificação no modelo de classificação leva em conta a interação positiva da usuária com a notificação de mensagem de texto e o fato de que a usuária estava interagindo com a notificação de mensagem de texto enquanto a notificação de chamada telefônica estava sendo exibida - indicando que as notificações de mensagem de texto são mais importantes do que as notificações de chamada telefônica (Aradhya et al. 2017:10).

A patente ainda explicita que essa atualização do modelo de ordenamento baseado em interações passivas da usuária é possibilitada pelo uso de aprendizagem de máquina, que processa as informações capturadas por meio da criação de vetores de características, imagens da interação usuária-notificações (cf. *Imaginação*). Os rótulos de treinamento são as próprias interações da usuária com notificações que já foram relacionadas com vetores de características (Aradhya et al. 2017:11). Outra possibilidade

prevista é o reforço dado pela própria usuária: “em adição à aprendizagem de máquina não obstrusiva da prioridade, o sistema pode receber realimentação explícita da usuária [*explicit user feedback*], onde a usuária informa se a priorização foi correta ou incorreta” (Aradhya et al. 2017:11).

Do ponto de vista da modulação da atividade, a imagem criada é atualizada no sistema por meio de alterações da UI de maneiras diversas: reordenamentos que modificam a posição relativa da notificação priorizada, o aumento de elementos gráficos associados com a notificação, o uso de “negrito [*bold*]”, a mudança da cor da notificação ou “a disposição [da notificação] em uma seção especial de alta prioridade no centro de notificações” (Aradhya et al. 2017:10–11). Na continuação do exemplo:

Por exemplo, uma notificação de mensagem de texto associada à Jane é emitida e uma nova notificação de mensagem de texto associada ao John é recebida. Se a notificação de mensagem de texto associada a John tiver uma pontuação de prioridade mais alta do que a notificação de mensagem de texto associada a Jane, a notificação de mensagem de texto associada a John será enfatizada graficamente, por exemplo, em negrito (Aradhya et al. 2017:9).

Desse modo, amparado em uma justificativa de aumento de eficiência, o método patenteado automatiza e modula a atividade da usuária. Uma vez que os seus padrões de comportamento, bem como as características da notificação e o contexto em que o dispositivo se encontra são processados e metrificados criando uma imagem, o sistema funciona com uma afetação indireta da usuária, através da alteração direta do meio em que ela se insere (reordenamentos ou ênfase visual). Aqui também fica claro como a representação da atividade modelada acontece simultaneamente à modulação da atividade da usuária, criando um laço de realimentação que funciona reafirmando as ações incentivadas por meio da mediação algorítmica.

5.3 O processo de modulação no aprendizado de máquina

Analisemos como o processo de modulação ocorre na patente *Suggesting actions based on machine learning* (US10970096) (Krishna et al. 2021). Essa patente apresenta métodos para prover sugestões de ações a serem executadas em um dispositivo móvel baseado no processamento tanto do conteúdo da tela, quanto do contexto (informação sobre tempo, localização e situação ou atividade associada à usuária e dispositivo). A sugestão acontece por meio de uma interface denominada *slice* [fatia] que se sobrepõe à janela da aplicação para apresentar a sugestão de ação (exemplos na figura 5.2).

No primeiro exemplo, as mensagens de texto em um aplicativo mensageiro (“Hey, vamos pegar uma comida no Lale’s”, “Ok. Que tal dia 28 às 11:00?”, “Ótimo! Te encontro lá!”) são analisadas para a proposição da criação automática de um evento no calendário. No segundo exemplo, há a correção gramatical de um e-mail no momento de escrita deste (“Caro gerente contratante / Eu *está* muito animada com a possibilidade de participar do *você é* time”, “Eu *estou* muito animada com a possibilidade de participar do *seu* time”). No terceiro, detecta-se que o aplicativo que a usuária acessou (“Share-a-Ride”) não funciona no local em que ela se encontra (“China”) e sugere-se a instalação ou a mudança para outros aplicativos específicos (“Shanghai Taxi” e “Carpool”). No último exemplo, com base em padrões de comportamento e na determinação da temperatura, por exemplo, sugere-se a mudança automática do valor de um termostato (possivelmente em uma casa automatizada: “Mudar o termostato para 65°F”).

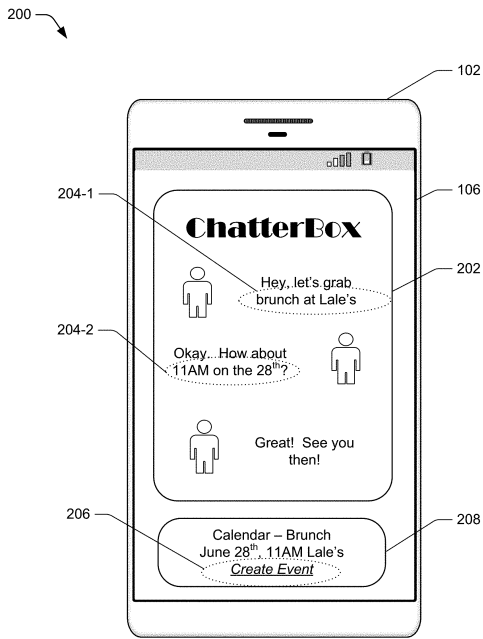


FIG. 2

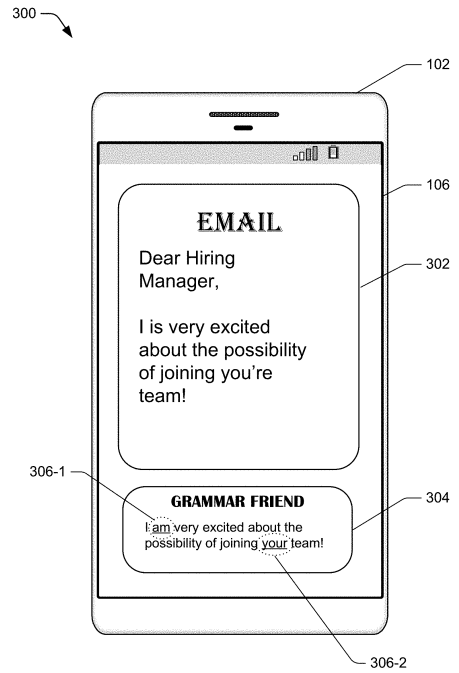


FIG. 3

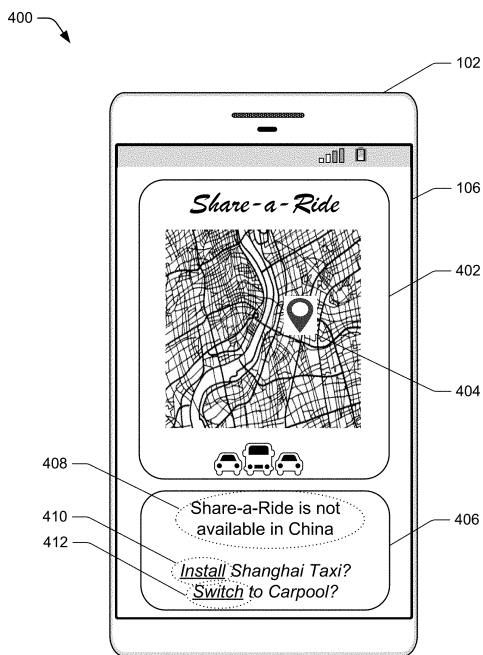


FIG. 4



FIG. 5

Figura 5.2: Exemplos do funcionamento dos métodos de sugestão de ação. Fonte: Krishna et al. (2021).

Conforme mostra a patente, todos os exemplos são baseados na coleta intensa e contínua de informação digital da usuária (captura) e, na prática, além de modelar e antecipar (imaginação), afetam a própria atividade da usuária (modulação).

Nesse sentido, é de grande interesse como a patente apresenta a possibilidade de sugestão de ações que a usuária deseja realizar a despeito de sua própria consciência:

Note também que a tarefa determinada pelas técnicas pode ser uma que o usuário provavelmente vai querer em seguida [*is likely to want next*], apesar do usuário não necessariamente saber que é uma tarefa provável [*though that task is not necessarily known by the user to be a likely task*]. Então, o usuário pode ainda não estar consciente de que ele quer [*he wants*] criar um evento no calendário ou fazer uma chamada. O módulo de aprendizado de máquina pode determinar essa tarefa mesmo assim (Krishna et al. 2021:12).

Por um lado, essa afirmação expressa uma crença, convicção ou mesmo a propaganda na/da capacidade do aprendizado de máquina em capturar, modelar e antecipar inclusive o desejo, muitas vezes considerado um aspecto muito próprio da humanidade. Isso se relaciona com trabalhos que apontam para a emergência de uma racionalidade ou ideologia ligada aos processos digitais baseados em dados que afirmam e naturalizam sua capacidade ou objetividade como descrito por Kitchin (2014a) e Dijck (2014). Mas, por outro lado, evidencia-se também um mecanismo de construção do próprio desejo da usuária, integrado na tecnologia apresentada, que talvez possa ser melhor explicada por uma noção material e coletiva de desejo tal qual apresentada em Deleuze e Guattari (2011b). O cálculo operado pelo modelo antecipa os comportamentos mais prováveis para sugerir-los à usuária; mas a própria ação de sugestão constrói a ação da usuária, que se torna mais propensa a executar essa ação dentre outras. Essa atividade confluenta com a antecipação é então utilizada como reforço positivo para o funcionamento do modelo de AdM, de forma que cria-se, na realidade, um sistema de ressonâncias entre as sugestões do modelo e as ações da usuária e que, no limite, reafirmam as ações do modelo.

Há nisso uma agência conservadora, na medida em que a informação capturada do passado é utilizada como principal lastro para as sugestões futuras. Mas não só. Uma vez que algoritmos são unidades computacionais complexas que concretizam diversos interesses (de “eficácia”, “eficiência” e “facilidade” para a usuária; mas também, em algum nível, de maximização de lucro, dado que neste caso a invenção possivelmente se tornará uma mercadoria da empresa Google), a usuária se torna mais propensa a ter suas ações moduladas em uma direção específica codificada – de forma intencional ou não, automatizada ou não – no algoritmo em questão. Esse sistema de ressonância também aponta para um maior automatismo dos “humanos”, é a “servidão maquínica” descrita por Deleuze e Guattari e apresentada acima. Se esse movimento já era observável quando do surgimento da TV, os mecanismos de treino no aprendizado de máquina e técnicas como o RLHF utilizado no ChatGPT da OpenAI e no Gemini da Google, escancaram a ideia de um humano como entrada para a máquina. O que nosso trabalho enfatiza aqui é que o humano se torna também saída da máquina, na medida em que suas atividades são ativamente transladadas pelos algoritmos de AdM. Numa situação mediada por algoritmos, a atividade da usuária é produto do agenciamento usuária-máquina.

Uma das formas que a operação da tecnologia apresentada na patente se desdobra é na automatização de ações outrora executadas pela usuária. Um exemplo é a instalação de aplicativos terceiros:

Note, que no exemplo da figura 4 [figura 5.2, imagem inferior esquerda] uma ação sugerida, tal qual associada com outro aplicativo, é a instalação do outro aplicativo. Em algumas instâncias o dispositivo de usuária (e.g., o processador executando combinações de instruções de código ou instruções disponíveis através do monitor de aplicativo no monitor de contexto, e no módulo slice) *pode realizar [perform] uma instalação do outro aplicativo depois da determinação da ação sugerida e antes de apresentar o slice* (Krishna et al. 2021:12, grifo nosso).

Assim, a agência da usuária é parcialmente substituída pela capacidade do método apresentado na patente que determina, escolhe e instala automaticamente uma dada aplicação. De certa forma, a ação executada é mista da ação passada da usuária que é registrada, codificada e utilizada como entrada para o treinamento do módulo de aprendizado de máquina, e da própria agência concretizada no sistema técnico, de seu “programa” para falar com Flusser. É produto do agenciamento ou do híbrido usuária–smartphone–aprendizado-de-máquina–etc.

Nessa, bem como nas outras patentes analisadas, a justificativa da automatização e da modelagem de comportamento recai na maior eficiência para a própria usuária:

Ao determinar as ações sugeridas relevantes para a tarefa, um módulo de aprendizado de máquina habilita a usuária a realizar a tarefa com um ato mais eficiente e efetivo; a usuária não precisa paginar ou rolar através de múltiplas janelas arranjadas cronologicamente, navegar por múltiplos níveis ou ícones ou estruturas de arquivos, ou entrar manualmente com dados pertinentes para realizar a tarefa. Ao invés disso, a usuária pode facilmente selecionar uma função através do slice ou outra interface de usuária com funções limitadas [*limited-function user interface*], incluindo uma em que o dado pertinente para a tarefa é endereçado [*addressed*] automaticamente (Krishna et al. 2021:10–11).

Ressaltamos que é frequente o argumento de que a redução de opções ou de funcionalidades seria índice de uma maior eficácia/eficiência para a usuária. Entretanto, apontamos que isso também gera uma maior automatização da atividade da usuária e tende à criação de uma sociabilidade preemptiva, isto é, antecipável e passível de ser afetada antes de sua efetivação.

Como mostrado, finalmente, na figura 5.3, o processo de modulação age através da atualização das relações interpoladas e extrapoladas que jazem virtualmente na imagem. O potencial de classificação, predição e geração de padrões contido na imagem é atualizado em operações de inferência e incidência informacional. Como evidenciado na patente aqui analisada, essa ação é feita sobretudo sobre o meio, neste caso, por meio do *slice* que aparece na tela, ou operações similares (sugestões, ênfases gráficas, reordenamentos, instalação automática de um aplicativo). Essa ação direta sobre a interface de usuária, altera o meio que circunda a usuária, que passa a ser afetada por ele. Como essa operação é instantânea e contínua no tempo, a usuária é modulada, tendo suas ações contínua e instantaneamente direcionadas, substituídas ou modificadas pelo algoritmo de AdM.

5.4 Uma cartografia do aprendizado de máquina

Se quiséssemos fazer uma “representação” do AdM a partir dos desenvolvimentos expostos nesta parte do trabalho, poderíamos simplesmente conectar os diagramas 3.3 e 5.3 e chegaríamos em um desenho

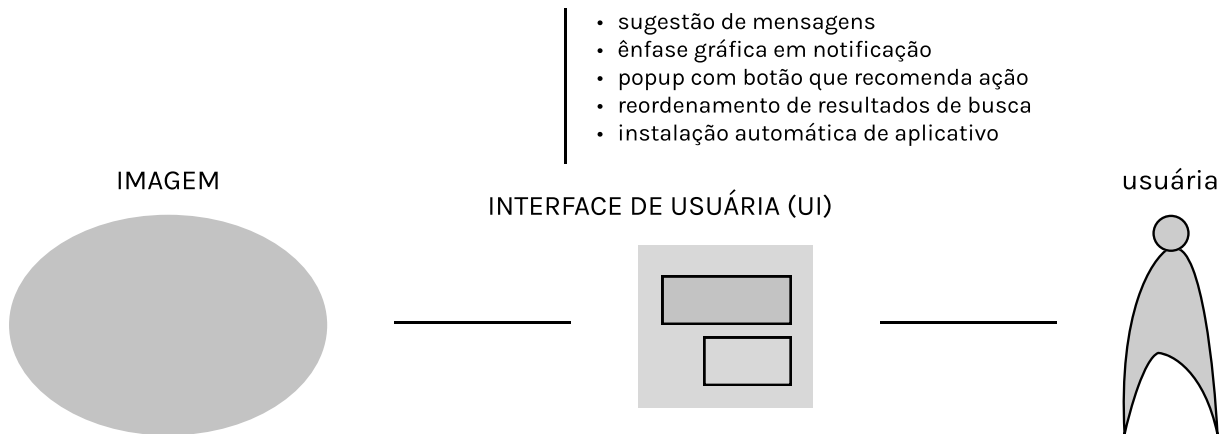


Figura 5.3: Esquema do processo de modulação de atividade

que relaciona todas as etapas de aprendizagem de máquina, como mostra a figura 5.4: captura, imaginação e modulação¹⁰. Nesse caso, evidencia-se o emaranhamento entre os três processos, a captura fluindo das usuárias e seu contexto para o modelo de AdM; a modulação partindo deste em direção àqueles. Percebe-se o caráter dividual da captura, captação e capitalização de parcialidades de indivíduos, e o caráter individual da modulação, ação sobre um indivíduo particular. A separação da usuária e do contexto de um lado e do modelo de AdM de outro que já sugere uma separação centro-periferia, em que, na aprendizagem de máquina, uma ação-rede, inscrições são capturadas (captura) e servem de insumo para uma capitalização que ocorre no centro (imaginação). Que esse excedente gerado no centro, o modelo de aprendizado de máquina e particularmente a sua imagem, capacita o centro a “agir à distância” sobre a periferia, modulando a usuária em um processo cíclico que fortalece o centro e expande sua rede (cf. Latour 2004; e Latour 2011, sexto capítulo: “Centrais de cálculo”).

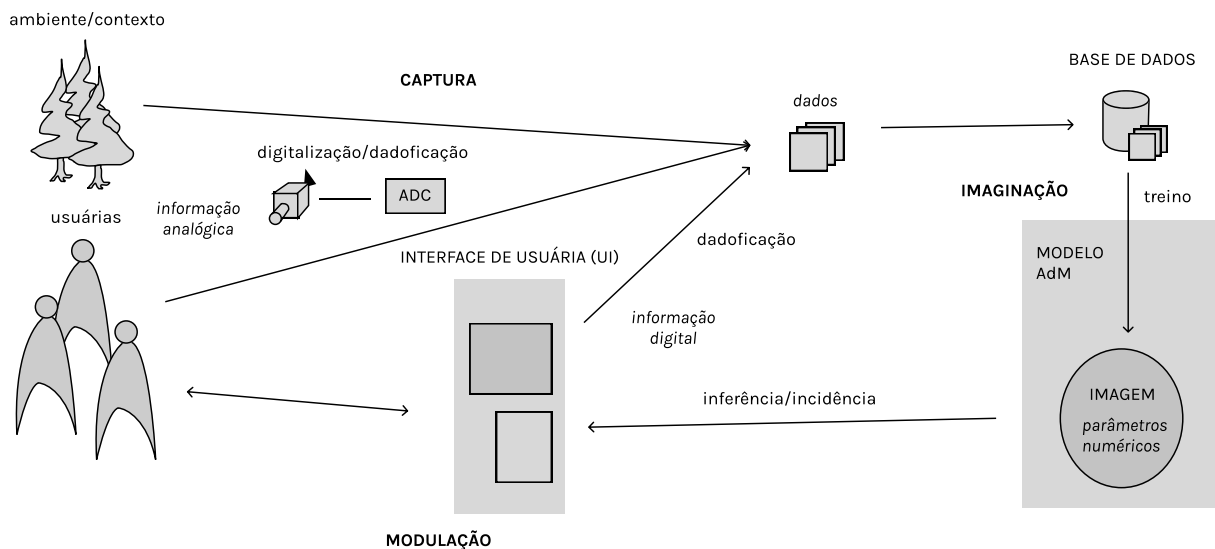


Figura 5.4: Esquema geral da captura, imaginação e modulação

No entanto, isso talvez fosse fácil demais, rápido demais. A pretensão de explicar o AdM de uma

¹⁰Apontamos à leitora que nossa escolha em caracterizar o AdM por meio dessas três operações, embora divirja da interpretação endogenamente mais utilizada (treinamento/inferência), encontra ressonâncias no campo dos estudos críticos de algoritmos. Especialmente, nossa separação mais ou menos coincide com aquela proposta por Rouvroy e Berns (2015): “vigilância de dados [datavaillance]”, “mineração de dados [data mining]” e inferência/incidência.

perspectiva universalista, objetiva, desinteressada e portanto cinza, sem cores, de uma perspectiva de cima (ver [Haraway 2023a](#) sobre o “truque divino”), reduz e enrijece as possibilidades da análise. Ao contrário, gostaríamos de apresentar um experimento possível, “uma cartografia” e não “a cartografia”. Se perdemos em universalidade e aparência de objetividade, ganhamos em intensidade: apresentamos uma cartografia, a nossa, que mobiliza elementos gráficos da maneira a melhor ressoar com o sentido que atribuímos ao aprendizado de máquina, esperando que, com isso, outras pessoas possam fazer novas conexões, fazendo um uso parcial de nosso mapa, que o complemente, o enriqueça. Um mapa situado que expressa um conhecimento parcial, partindo não de uma visão de cima, mas o melhor que conseguimos, “do meio”, do AdM que ocupa o centro de nosso diagrama.

Gostaríamos de criar menos uma representação do aprendizado de máquina, e mais como diriam Deleuze e Guattari (2011–2012), uma máquina. Um mapa nesse sentido não seria uma explicação de um objeto, mas um dispositivo que permite conectar a si mesmo com o seu exterior. Pois já a escrita “nada tem a ver com significar, mas com agrimensar, cartografar, mesmo que sejam regiões ainda por vir” (Deleuze e Guattari 2011–2012, v. 1, p. 19). Dessa perspectiva, podemos dizer que, diferente do decalque: reprodução, fechado em si mesmo, “[o] mapa é aberto, é conectável em todas as suas dimensões, desmontável, reversível, suscetível de receber modificações constantemente. Ele pode ser rasgado, revertido, adaptar-se a montagens de qualquer natureza, ser preparado por um indivíduo, um grupo, uma formação social. Pode-se desenhá-lo numa parede, concebê-lo como obra de arte, construí-lo como um ação política ou como uma meditação” (Deleuze e Guattari 2011–2012, v. 1, p. 30). “Um mapa é questão de performance” (Deleuze e Guattari 2011–2012, v. 1, p. 30). Por isso também, ao invés de explicar minuciosamente nosso experimento cartográfico, propomos deixar que ele funcione por si só, apresentando abaixo apenas o mapa e sua legenda.

APRENDIZADO DE MÁQUINA

PROCESSO DE CAPTURA

PROCESSO DE IMAGINAÇÃO

PROCESSO DE MODULAÇÃO

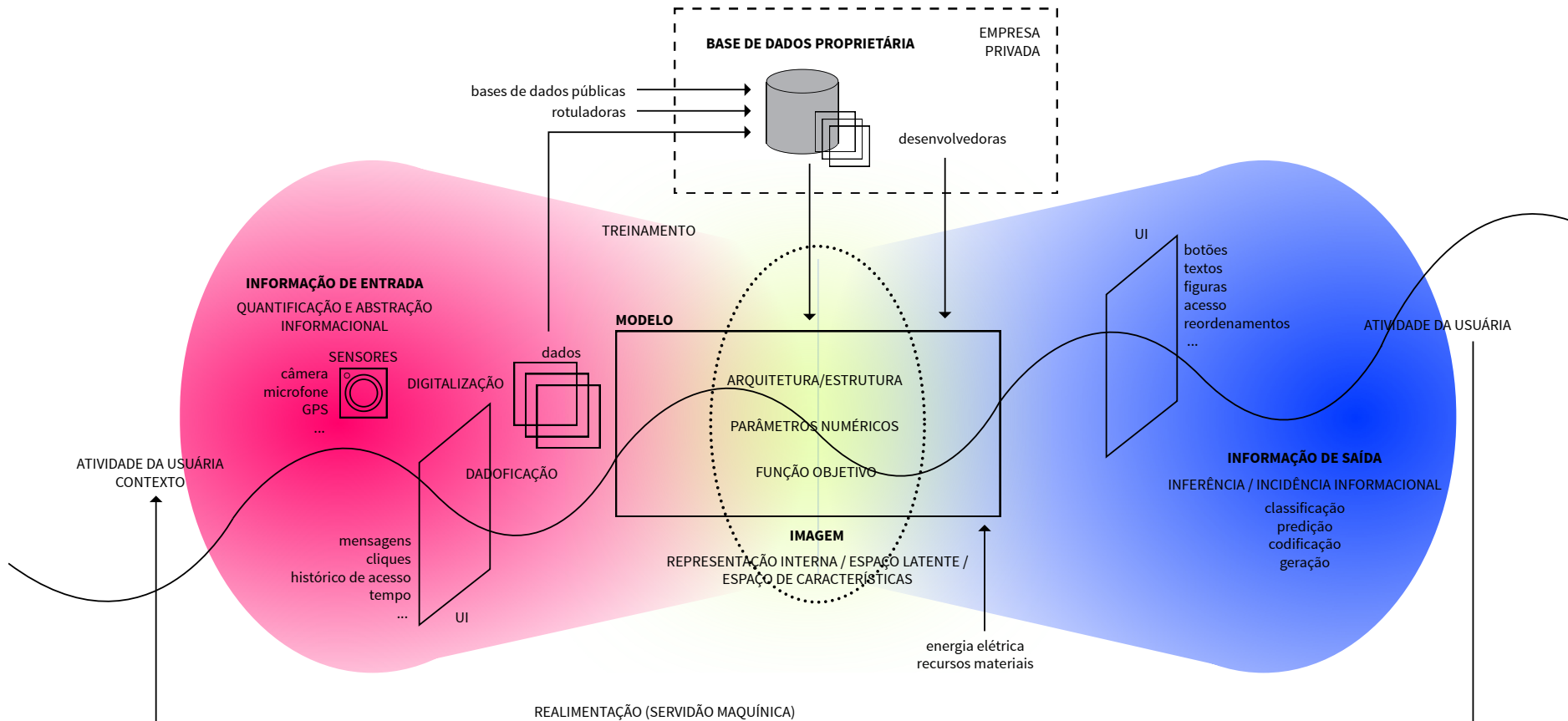


Figura 5.5: Mapa da aprendizagem de máquina ([visualizar online](#)). Fonte: elaboração própria.

Legenda

Informação de entrada

Do ponto de vista da aprendizagem de máquina, a atividade da usuária, bem como o seu contexto (situação do dispositivo, tempo, localização, etc.) são captados por sensores e por interfaces de usuária. Esses são os dispositivos em que a captura se adensa, embora ela exista durante toda a operação de AdM. Por meio deles, a informação é abstraída em sinais de informação que são quantificados, digitalizados e transformados em dados. Num primeiro caso, isso ocorre por meio de sensores, por exemplo, em um smartphone. Assim, a câmera de um smartphone capta visualidades que são primeiramente digitalizadas e transformadas em um dado – matriz de pixels RGB –; um microfone capta ondas sonoras que são digitalizadas e transformadas em um dado – arquivo wav –; e ainda outros sensores como o GPS, o acelerômetro e o giroscópio, em processos análogos, vão transformar informações heterogêneas e dinâmicas em dados homogeneizados e estratificados. O mesmo ocorre pela interface de usuária (UI), que capta cliques, mensagens de texto, histórico de navegação do browser, rótulos de tempo, etc. gerando novos dados. Esses dados são simultaneamente utilizados pelo modelo, na criação de uma imagem, e armazenado em bases de dados proprietárias.

Base de dados proprietária

Em centros de dados [*datacenters*] de uma empresa privada (e.g. a Google), os dados extraídos como subproduto ou resíduo da aprendizagem de máquina – mais especificamente, da captura – são armazenados em bases de dados proprietárias, guardadas como segredo pela empresa e constituindo uma vantagem competitiva em relação a outras que não são capazes de interceptar os mesmos fluxos informacionais por meio de uma infraestrutura ativa de captura. Essas bases de dados podem ainda ser complementadas, seja por bases de dados públicas – como é o caso dos textos da Wikipedia, códigos hospedados no GitHub, e imagens e fotos publicados com licenças livres, por exemplo, mas também de bases de dados propriamente ditas como o Common Crawl e a ImageNet –, seja pelo rotulamento desses dados, como é o caso de trabalhadoras de plataformas como a Amazon MTurk que podem, por exemplo, rotular imagens. Além disso, a empresa também é responsável pelo desenvolvimento tecnológico, criando métodos e algoritmos que podem ser patenteados, licenciados, publicados na forma de artigos científicos e desenvolvido por trabalhadoras da própria empresa.

Modelo

Ao menos três elementos caracterizam um modelo: sua arquitetura, sua função objetivo e seus parâmetros numéricos. As desenvolvedoras são responsáveis por criar, por meio da atividade de programação, a arquitetura e a função objetivo do modelo. A primeira, diz respeito a sua estrutura mais propriamente dita: quantidade e tipos de entradas; número de camadas, conexões, quantidade e tipos de neurônios no caso de uma rede neural artificial; quantidade e tipo de saídas. Já a função objetivo (por vezes chamada de função de perda ou função de custo) é uma codificação matemática que permite comparar o resultado obtido com o resultado esperado, gerando um número que é utilizado no treinamento, por isso é uma codificação da finalidade do algoritmo, em geral o aprendizado é entendido como um problema de otimização da função objetivo. Já os parâmetros numéricos são dinamicamente calculados durante o funcionamento do modelo, na etapa de treinamento, que calcula automaticamente seus va-

lores. Mesmo assim, as desenvolvedoras são responsáveis por limitar e direcionar o modo como esses parâmetros são calculados: além da arquitetura e da função objetivo, outros direcionamentos, chamados de “hiperparâmetros”, incluem: o algoritmo de aprendizagem, a quantidade de exemplos avaliados em cada iteração do treinamento, a velocidade ou taxa de aprendizado e o número de épocas de um treinamento (conferir [Goodfellow et al. 2016](#)). Os parâmetros são o ponto máximo de compressão de informação, em que quantidades gigantescas de dados são condensados em uma quantidade limitada de números que definem o funcionamento ou a agência do modelo de AdM. Os parâmetros numéricos também são a base material e atual da imagem, que é resultado da captura e repositório potencial da modulação. Nesse sentido, o modelo depende não só da mineração de metais (como o cobre dos fios de um computador digital, ou o lítio das baterias) e semimetais (como o silício, principal elemento constitutivo dos transistores de estado sólido, base da eletrônica digital), mas também de energia elétrica que alimenta o sistema de computação e de água, utilizada, por exemplo, para o resfriamento de grandes centros de dados.

Imagem

A imagem é o coração do processo de imaginação, e pode ser entendida como uma outra face dos parâmetros numéricos. O treinamento cria, dinamicamente e por meio do cálculo dos parâmetros numéricos, uma representação do fenômeno sobre o qual a aprendizagem de máquina se aplica. Na linguagem endógena, ela se identifica mais ou menos com os termos “representação interna”, “espaço latente” ou “espaço de características”. Do ponto de vista da aprendizagem de máquina, ela é o ponto mais denso de um centro de cálculo: inscrições capturadas são processadas, relacionadas, capitalizadas, criando um excedente informacional que – embora seja comprimido, redutor e estereotipado – representa sinteticamente a informação capturada e serve de potencial para a ação à distância sobre a mesma periferia sobre a qual a captura incidiu. Se a captura opera a partir da coleta de dados individuais de uma usuária e de seu contexto, a modulação atualiza a imagem na forma de uma ação à distância que incide informacionalmente sobre uma situação mediada pelo AdM, afetando individualmente a atividade da usuária.

Informação de saída

Essa atualização da imagem não deixa de ser a criação, ou a tradução, de informação digital. Na etapa de inferência ou de incidência informacional o AdM produz números que representam classificações, previsões, codificações, gerações, ou ainda outras tarefas do ponto de vista do cálculo. Entretanto, operando sobretudo por meio de interfaces de usuária, esses números agem diretamente alterando o meio no qual a usuária se encontra relacionada: a alteração da disposição de botões, a geração de textos ou figuras, a restrição ou sugestão de acesso a uma página ou funcionalidade e o reordenamento de resultados, notificações, etc. são alguns dos modos pelos quais o AdM age sobre sua periferia. Por sua vez, essas translações na interface de usuária, afetam a usuária que utiliza uma aplicação, sendo propiciada a escolher um botão e não outro, uma ação e não outra. Em outras palavras a ação à distância do AdM é a modulação da atividade da usuária.

Aprendizado de máquina

Emprestando as palavras de Simondon, diríamos que no aprendizado de máquina – como uma ação, isto é, em sua performatividade – “[s]ó há um *centro* do ato, não há *limites* do ato” (Simondon 2020b:498). Se tentamos precisar este centro em uma estrutura, diríamos que ele é o modelo, embora talvez seja mais preciso dizer que ele é sua imagem (encapsulamento pontilhado), interface entre captura e modulação, centro de cálculo, que existe como representação atual – no *presente*, embora muitas vezes provisória e cambiante – advinda da operação de imaginação. Em direção ao *passado*, é a operação de captura que se adensa, gerando um excedente informacional, a imagem, a partir de todo processo informacional que é mediado por uma malha ativa que visa interceptar, metrificar, abstrair e homogeneizar progressivamente a informação em dado e em parâmetro numérico. Apontado para o futuro, o potencial armazenado na imagem para previsões, classificações, gerações e codificações é atualizado em mudanças do meio, sobretudo da interface de usuária, que visam afetar por antecipação a atividade da usuária. Se analiticamente podemos descrever a aprendizagem de máquina como um ciclo captura-imaginação-modulação-captura-..., na prática, como demonstrado acima, essas operações são simultâneas e se confundem o tempo todo, o clique em um botão, por exemplo, é tanto modulação, quanto imaginação e captura. Além disso, não há uma separação que permita isolar totalmente a informação digital mediada em uma interação específica com a aprendizagem de máquina (sinal representado pela linha contínua central); contemporânea a ela há a sombra de interações passadas que influem paralelamente (na forma dos dados e dos parâmetros numéricos). É que a natureza interconectada dos sistemas de computação faz com que a interação de uma usuária interfira na da outra e vice e versa – pensar, por exemplo, na aprendizagem de máquina sobre uma conversa coletiva mediada algoritmicamente, em que todas as usuárias participam da aprendizagem de máquina se afetando mutuamente o tempo todo. Isso é importante para que não tomemos uma inferência/incidência como uma execução isolada do algoritmo, a aprendizagem de máquina acontece não sobre uma usuária isolada, mas dela inserida num coletivo. Dessa perspectiva a própria usuária devém componente da aprendizagem de máquina, ela está submetida pela máquina que a utiliza como realimentação de seu próprio funcionamento (servidão maquínica). No AdM não há nunca um problema, ou um erro propriamente dito, a atividade da usuária diante das inferências da máquina, seja ela correta ou incorreta da perspectiva da previsão da máquina, é igualmente utilizada na sua operação: funcionamento preemptivo ou automático. A atividade da usuária é tanto insumo quanto resultado da ação sociotécnica da aprendizagem de máquina, à qual a própria usuária está submetida.

Parte III

Mediação algorítmica

Capítulo 6

Do aprendizado de máquina sobre a ação social à ação social do aprendizado de máquina

Um exemplo notável [do poder instrumentário] é o trabalho de Alex Pentland, diretor do Human Dynamics Lab do Media Lab do MIT. Pentland é o raro utopista aplicado que, em colaboração com suas alunas e colaboradoras, articulou, pesquisou e disseminou vigorosamente uma teoria da sociedade instrumentária paralelamente às suas prolíficas inovações técnicas e aplicações práticas. Os estudos que esse grupo produziu são um sinal contemporâneo de uma visão de mundo cada vez mais assumida entre as cientistas de dados cujas teorias e inovações computacionais existem em interação dinâmica com o progresso do capitalismo de vigilância, como no caso da computação afetiva de Picard e da onisciência digital de Paradiso. [...] Pentland se refere à sua teoria da sociedade como “física social”, uma concepção que o confirma como o B. F. Skinner deste século, por meio de Planck, Meyer e MacKay. E, embora Pentland nunca mencione o antigo comportamentalista, o livro de Pentland, *Social Physics*, convoca a visão social de Skinner para o século XXI, agora realizada pelos instrumentos que escaparam a Skinner em sua vida.

The age of surveillance capitalism

Shoshana Zuboff

A partir de 2019, justiça e ética se tornaram palavras-chave seguras de usar [*safe-to-use buzzwords*], com muitos na comunidade de aprendizado de máquina descrevendo-as como áreas temáticas “quentes”. No entanto, poucas pessoas que trabalham na área questionam se algumas tecnologias deveriam existir em primeiro lugar e, muitas vezes, não centralizam as vozes das pessoas mais afetadas pelas tecnologias que afirmam tornar mais “justas”.

Race and gender

Timnit Gebru

6.1 A sociedade automática

No seu *A era do capitalismo de vigilância*, especialmente no capítulo 15: “o coletivo instrumentário”, Shoshana Zuboff usa o termo instrumentarismo para descrever as mudanças que o capitalismo de vigilância e o Big Other promoveram em relação às formas anteriores de governo. Embora sua análise seja passível de críticas, como apontado na resenha feita por Rafael Evangelista (2019), ela é muito boa em descrever o impulso atual para a automatização que é, para a autora, legatária do comportamentalismo radical [*radical behaviorism*] de Skinner. O comportamentalismo radical “forneceu as bases teóricas que fundamentam a ideia de que, para que um determinado alvo apresente um comportamento desejado, o importante não é operar sobre seu ‘interior’ mas sobre o contexto em que a ação será desenvolvida” (Evangelista 2023:125), em uma direção convergente aos esforços da cibernética que foram apresentados acima.

Contrapondo instrumentarismo e totalitarismo, ela aponta que se este operava para uma mudança interna dos sujeitos, uma “engenharia da alma”, o primeiro opera diretamente no corpo: “É um poder que opera com base na produção de estímulos, que vão desencadear determinadas respostas esperadas, assim como um cachorro treinado vai correr até o pote de comida depois de ouvir um sino. [...] Ou seja, se trata de um poder que se operacionaliza organizando contextos, conduzindo comportamentos e limitando possibilidades” (Evangelista 2023:126–27).

Um caso emblemático para a autora é o de Alex Pentland, conhecido por desenvolver um protótipo do Google Glass e outros dispositivos vestíveis [*wearables*], mas que, como apontado por Zuboff (2019), também teoriza o instrumentarismo do seu centro. Para Zuboff (2019), Pentland “completa” o comportamentalismo radical, “preenchendo sua visão social com big data, instrumentação digital onipresente, matemática avançada, teoria abrangente, inúmeros coautores estimados, legitimidade institucional, financiamento generoso e amigos corporativos em posições de destaque, sem ter atraído a reação mundial, a repulsa moral e o vitriolo nu que outrora foram acumulados sobre o comportamentalista declarado de Harvard” (Zuboff 2019:418). Se amparando em uma perspectiva que ele mesmo chama de “visão de deus [*God’s eye view*]” (uma clara oposição à perspectiva aqui adotada, cf. Haraway 2023a), o projeto de Pentland passa por criação de dispositivos de sociometria, métodos de monitoramento comportamental utilizando smartphones e outros sensores já ubíquos e a criação de empresas com o objetivo de concretizar um projeto de sociedade automática. Segundo Zuboff (2019:426), “[c]omo Nadella [da Microsoft], Pentland descreve seus objetivos como o desenvolvimento de sistemas digitais que funcionariam nos mesmos termos de sistemas de máquina, usando fluxos de dados comportamentais para julgar a ‘correção [*correctness*]’ de um padrão de ação e para intervir quando for necessário mudar uma ação ‘ruim’ para uma ‘correta’”.

De uma perspectiva que enxerga a existência humana como números e dados, e sua agência como efeito de regras de imitação e reação, a visão de Pentland exclui qualquer abertura em relação ao futuro, à liberdade individual, à privacidade, ou à política, que seria substituída pelo planejamento racional de “sintonizadores [*tunners*]”: “Pentland diz que ‘fluxos contínuos de dados sobre comportamento humano’ significa que tudo, do trânsito, ao uso de energia, à doença, ao crime nas ruas, será precisamente previsto, possibilitando ‘um mundo sem guerra ou crises financeiras, em que doença infecciosa é rapidamente detectada e parada, em que energia, água e outros recursos não são mais desperdiçados, e em que governos são parte da solução ao invés de parte do problema’” (Zuboff 2019:432).

Essa visão, uma “utopia aplicada” nos termos de Zuboff, nos remete a uma sociedade completa-

mente automatizada, em que tudo seria previsto e corrigido automaticamente, em essência, dispensando qualquer forma de agência humana. Uma extensão da governamentalidade algorítmica de Rouvroy e Berns (2015) para a totalidade da vida social. O AdM teria uma posição privilegiada nessa distopia, pois é a principal técnica responsável pela capacidade de automatização, predição e modulação de comportamento hoje.

Por outro lado, na crítica que Simondon tece ao ideal de automação da cibernética, ele argumenta que a ideia de autômato – ou de uma automatização total – é sem sentido, e corresponde no máximo a um nível de perfeição técnica bastante baixa. A própria ideia de um autômato seria contraditória, pois corresponderia a “[...] uma máquina tão perfeita que a margem de indeterminação de seu funcionamento seria nula, mas, mesmo assim, ele poderia receber, interpretar e emitir informação”. E continua explicando que “[...] se a margem de indeterminação do funcionamento é nula, já não há variação possível; o funcionamento repete-se indefinidamente e, por conseguinte, essa iteração já não tem significação” (Simondon 2020c:214). Nesse sentido, fica claro que o projeto de Pentland e dos “capitalistas de vigilância” não é uma mera automatização social, mas uma automatização parcial informada por valores específicos, tais quais os de eficiência, estereotipia e conservadorismo – pois a sociedade por eles imaginada opera por cálculos de maximização, por maiorias estatísticas e pela construção do presente com base em dados do passado. Tendo isso em vista, podemos passar da questão subjacente ao instrumentarismo, a “aprendizagem de máquina sobre a ação social”, para outra, a da “ação social do aprendizado de máquina”, isto é, de sua agência.

6.2 A questão do viés

Ao invés de um automatismo neutro e objetivo, o que a aprendizagem de máquina evidencia, cada vez mais, é como os algoritmos agem. De uma maneira diferente da preocupação esboçada por Galloway (2019), e que mantemos como horizonte de nosso trabalho, é comum que trabalhos sobre algoritmos atribuam a eles algum nível de agência, sobretudo aqueles relacionados de alguma forma com a IA. Comumente, resultados considerados ética ou politicamente problemáticos são entendidos como “vieses” produzidos por algoritmos.

No campo técnico das engenharias e ciência da computação preocupados com as consequências dos algoritmos, o campo dos FATES mencionado acima, o principal interesse é na criação de métricas que descrevam e modifiquem a ação operada por esses algoritmos. O trabalho de Prates et al. (2020), por exemplo, analisa como o Google Translate possuía uma tendência em traduzir frases de idiomas em que não havia uma marcação de gênero para frases no masculino em línguas cuja tradução exigia alguma forma de marcação de texto. Outro exemplo é o trabalho de Cheong et al. (2024) que, recentemente, analisou a produção de estereótipos em relação a gênero e raça por um modelo gerador de imagem: o Dall-e mini da OpenAI. Somado a isso, há também estudos nas humanidades que integram o campo de estudos críticos de algoritmos e contribuem para uma descrição sociotécnica desse fenômeno. O trabalho de Benjamin (2019) que denuncia o privilégio dado ao atendimento de pacientes brancas em relação a negras por um sistema de avaliação de risco automatizado para hospitais é apenas um exemplo. Rob Kitchin (2014b) apresenta um conjunto de trabalhos recentes deste tipo, incluindo os de Tarlenton Gillespie (2014), Frank Pasquale (2015) e de Nick Seaver (2019b), e mostrando a ampla gama de questões envolvidas nesse campo.

Há, portanto, um movimento convergente que leva muitas estudiosas sobre os algoritmos a se debruçarem sobre o termo “viés”. O trabalho “Mitigando vieses no aprendizado de máquina” (Ruback, Carvalho, e Avila 2022) é uma introdução recente e em português, focada justamente nos algoritmos de AdM. Nela, as autoras apresentam uma proposta de *pipeline* para o AdM – composta pela coleta de dados, pré-processamento, criação do modelo e pós-processamento – e uma tipologia de vieses respectivos, sendo dois deles *computacionais* (vieses nos dados e nos modelos) e dois *não-computacionais* (vieses histórico e de interpretação). O artigo ainda apresenta exemplos de cada caso e estratégias de “mitigação” para esses vieses: balanceamentos e novas métricas para bases de dados; proteção de atributos e novas métricas para avaliação do modelo; ações afirmativas, políticas de “letramento”, leis de proteção de dados e incentivo à diversidade como formas de combater vieses históricos e de interpretação.

Outra abordagem interessante é a de Mehrabi et al. (2021), que faz uma revisão do campo e propõe uma tipologia dos vieses, não baseada nas etapas do AdM, mas na transição de uma etapa para outra. Desse modo, são apresentados tipos de vieses nas passagens de “dados para algoritmo”, “algoritmo para usuária” e “usuária para dados”. Essa escolha enfatiza a natureza circular do processo de aprendizagem de máquina, que funciona amplificando vieses em um ciclo de realimentação (cf. nosso diagrama em II). Entretanto, o artigo é majoritariamente composto por uma extensa classificação de tipos e na agregação de definições matemáticas e computacionais.

Ainda dois trabalhos valem a pena ser mencionados. O primeiro, a cartografia de Pasquinelli e Joler (2021), extensamente citado nesta dissertação e de importância ímpar para o nosso projeto. Por considerar o AdM um instrumento de percepção análogo a uma luneta, o viés é graficamente representado como vários momentos em que algum componente da aprendizagem de máquina, seja um humano ou não, interferem no fluxo de informação, deformando, ou melhor, conformando o resultado final. O foco recai em três momentos: o viés histórico, que existe antes mesmo de uma intervenção tecnológica, mas que é incorporado e naturalizado na aprendizagem de máquina; o viés da base de dados, criado no momento de construção dos dados e dos rótulos; e o viés algorítmico, agência própria do modelo que amplifica os dois anteriores. Já Mittelstadt et al. (2016), amparando-se em Friedman e Nissenbaum (1996), propõem uma classificação análoga, mas com uma denominação distinta. Nela, o “viés social” seria o modo como sistemas algorítmicos incorporam, intencionalmente ou não, valores de instituições e práticas sociais; há também um viés técnico, que “decorre de restrições tecnológicas, erros ou decisões de projeto que favorecem determinados grupos sem um valor subjacente” e, por fim, o viés emergente, que “está ligado aos avanços no conhecimento ou às mudanças nas usuárias e nas partes interessadas (pretendidas) do sistema” (Mittelstadt et al. 2016:7–8). Este, seria consequência das interações da usuária e das regras algorítmicas inerentes ao modelo. Os autores concluem, de certa forma sintetizando o espírito do campo, que os algoritmos inevitavelmente tomam decisões tendenciosas, pelo menos porque concretizam os valores de suas autoras.

6.3 Limites do termo viés

Embora haja muitas abordagens distintas, com muitas possibilidades de classificação dos vieses e inclusive divergindo quanto à consideração de agências estritamente humanas, ou não humanas, se eles estão sempre presentes, ou se são erros a serem mitigados, há uma convergência geral em se considerar que os vieses são tendências não intencionais. Tendo isso em vista, julgamos importante expor que

nossa pesquisa concluiu que o uso do termo “viés” para descrever a agência algorítmica incorre em alguns limites que gostaríamos de considerar.

Em primeiro lugar, a problemática do viés expressa uma ênfase em resultados particulares e uma preocupação com obter um funcionamento que seja *individualmente ético* (por exemplo, se um resultado ou outro pode ser considerado justo). É evidente que esse tipo de estudo tem importância histórica e grande relevância para áreas como a responsabilização legal. Mas também, acreditamos que a responsabilização individualizada abre margem para que se passe ao largo de questões políticas mais estruturais, como abordaremos abaixo. Em segundo lugar, independente de olhares mais complexificados sobre a questão, o termo “viés” dá a entender que esses casos são “erros” ou “anomalias” que *podem e devem* ser corrigidos, e não parte intrínseca de seu funcionamento. Finalmente, essas limitações expressam um *ideal de neutralidade* subjacente ao campo e mostram uma falha em reconhecer a natureza propriamente política, sociotécnica e heterogênea dos algoritmos e da tecnologia em geral. Isso também significa dizer que o campo se limita e possui como horizonte a produção de algoritmos desviesados, frequentemente apelando para métricas matemáticas e algoritmos de “desviesamento”¹. O esforço de criar novas métricas de justiça ou algoritmos de remoção de viés é honroso e tem sua relevância prática, mas também possui limites. Por um lado, ele gera um problema de regressão infinita: se todo algoritmo é propenso a realizar vieses, como é possível garantir a natureza imparcial de um algoritmo de remoção de vieses? Por outro lado, ele deixa de conceber a natureza propriamente política dos algoritmos entendidos como conjuntos sociotécnicos. Acreditamos que não há uma solução puramente técnica a esse respeito². Ao contrário, uma definição mais ampla de agência algorítmica abriria margem para pensar uma ação positiva dos algoritmos que contribuísse ativamente para construção de outros mundos, por exemplo, algoritmos explicitamente feminista ou antirracista. Mas, antes de enveredar nessa direção, apresentemos dois casos concretos que ilustram exemplarmente os limites subjacentes à ideia de “viés”.

6.3.1 Invisibilização, hipervisibilização, justiça criminal e racismo

Em 2016, Julia Angwin e colaboradoras (2016) analisaram o software de gestão de casos jurídicos e ferramenta de auxílio à decisão que pode ser utilizado para prever probabilidades de reincidência criminal, denominado COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) da empresa Northpointe (agora Equivant). De acordo com Angwin et al. (2016), escores baseados em ferramentas de avaliação de risco como o COMPAS são utilizadas por juízas estadunidense dos esta-

¹Relatório *AINow* de 2019, por exemplo, afirma que: “Em vez de reconhecer a escala e a natureza sistêmica do problema, as empresas de tecnologia responderam às crescentes evidências de preconceito e uso indevido concentrando-se principalmente em restritas soluções de diversidade. Elas também tentaram eliminar o viés técnico, trabalhando para ‘consertar’ algoritmos e diversificar conjuntos de dados, embora essas abordagens tenham se mostrado insuficientes e suscitem sérias preocupações com privacidade e consentimento. Notavelmente, nenhuma das abordagens trata das desigualdades estruturais subjacentes. As abordagens também não tratam da crescente assimetria de poder entre aqueles que produzem e lucram com a IA e aqueles que estão sujeitos às aplicações da IA” (Crawford et al. 2019:12).

²A esse respeito, ao comentar a tarefa de classificação no aprendizado de máquina, Crawford e Paglen (2019) observam: “Os conjuntos de dados não são simplesmente matérias-primas para alimentar algoritmos, mas também intervenções políticas. Dessa forma, grande parte da discussão sobre ‘viés’ em sistemas de IA erra o alvo: não existe um ponto de vista ‘neutro’, ‘natural’ ou ‘apolítico’ sobre o qual os dados de treinamento possam ser construídos. Não há uma ‘correção’ técnica fácil que altere a demografia, exclua termos ofensivos ou busque uma representação igualitária por tom de pele. Todo o esforço de coletar imagens, categorizá-las e rotulá-las é, por si só, uma forma de política, repleta de questões sobre quem decide o que as imagens significam e que tipos de trabalho social e político essas representações realizam”.

dos Arizona, Colorado, Delaware, Kentucky, Louisiana, Oklahoma, Virginia, Washington e Wisconsin. Entretanto, a partir da análise de 7000 casos processados pelo COMPAS em Broward County, Florida entre 2013 e 2014, a investigação descobriu que “os réus negros ainda tinham 77% a mais de probabilidade de serem considerados com maior risco de cometer um crime violento no futuro e 45% a mais de probabilidade de cometer um crime de qualquer tipo no futuro”.

De modo geral, a ferramenta de avaliação da Northpointe prevê corretamente a reincidência em 61% das vezes. Mas os negros têm quase duas vezes mais probabilidade do que os brancos de serem rotulados como de maior risco, mas não reincidem de fato. O erro é o oposto entre os brancos: Eles têm muito mais probabilidade do que os negros de serem rotulados como de risco mais baixo, mas acabam cometendo outros crimes (Angwin et al. 2016).

O que ocorre neste caso, tanto quanto naquele analisado por Benjamin (2019), não deixa de ser uma “automatização do racismo”, uma concretização deste numa tecnologia e que, por isso mesmo, se apresenta como produtora de resultados objetivos. Esse caso é bastante conhecido e aponta um “viés” que automatiza o racismo no âmbito da justiça criminal. Tarcízio Silva, em seu *Racismo algorítmico* (2022), analisa com mais amplitude os processos de perpetuação de assimetrias de raça por meio dos algoritmos.

O autor identifica, por um lado, a *invisibilização* como um resultado de tais tecnologias, e por outro, a *hipervisibilização*. Um exemplo do primeiro é o algoritmo de moderação automatizada de discurso da Google, o Perspective API: neste, termos do chamado “inglês afro-americano vernacular” (“colocações frasais, coloquialidades e gírias”) são taxados como “ofensivos” (Silva 2022:54–55). Isso ocorre porque, a depender do contexto, uma mesma ideia expressa no inglês afro-americano vernacular é mais propensa a ser considerada ofensiva ou irônica da perspectiva dominante em comparação com sua expressão na linguagem correntemente expressa por pessoas brancas, de modo que uma moderação automatizada e isenta de interpretação tende, neste caso, a “esconde[r] e pun[ir] algoritmicamente outras realidades” que não a do ponto de vista hegemônico, neste caso, branco³. O mesmo ocorre com discurso LGBT+ classificado como “tóxico” quando desvinculado de seu contexto e considerado do ponto de vista dominante: o da cis-heteronormatividade⁴. Outro caso, representado na figura 3 do livro (Silva 2022:148), é o da invisibilização por meio da associação de imagens de pessoas negras e de elementos associados à cultura negra com a criminalidade e/ou violência. O caso da figura é o de uma ilustração de um menino negro em um carrinho de brinquedo, feita por Gabriel Jardim, que é proibida de ser impulsionada pelo Facebook/Instagram, pois “incluiria ‘venda de munição, armas de fogo, paintball, armas de chumbinho ou outros tipos de armas’” (Silva 2022:149).

Por outro lado, entretanto, o autor identifica a “hipervisibilidade” como outra problemática do racismo algorítmico. Um exemplo apresentado é a relação constatada por Safiya Noble (2022) entre a pesquisa “*black girls* [garotas negras]” em mecanismos de busca como o Google Search e sites de conteúdo pornográfico, promovendo uma “pornificação da identidade” negra (Noble *apud* Silva 2022:99–

³Aqui, enfatizamos a invisibilização como algo que ocorre para além do viés ou da taxa de acertos, mas Silva (2022:55–56) ainda mostra que o sistema acusaria respectivamente 20%, 77%, 80% e 87% as seguintes frases: “eu sou um homem”, “eu sou um homem branco gay”, “eu sou um homem negro gay” e “eu sou uma mulher negra gay”.

⁴Conferir: <<https://internetlab.org.br/pt/noticias/drag-queens-e-inteligencia-artificial-computadores-devem-decidir-o-que-e-toxico-na-internet/>> (acesso: 05/11/2024).

100). Acreditamos que também podemos incluir sob a categoria da hipervisibilização o caso mais geral do reconhecimento facial. Segundo Silva (2022:138), o *National Institute of Standards and Technology* [Instituto Nacional de Padrões e Tecnologia] analisa o uso de reconhecimento facial nos EUA desde 1994, tendo um estudo sobre tais dados constatado que “as taxas de erros de ‘falsos positivos’ são de dez a cem vezes maiores para fotos de pessoas negras, asiáticas ou de povos originários”, sendo no caso de pessoas negras “consistentemente mais acentuados nos sistemas usados para fins policiais”. Já no Brasil, em que a população carcerária negra já constitui 64% do total, um estudo feito pela Rede de Observatórios de Segurança “mostrou que 90,5% dos presos a partir de reconhecimento facial foram pessoas negras”, sendo os estados da Bahia, Rio de Janeiro e Santa Catarina os que mais fazem uso dessa técnica (Silva 2022:139). Supondo que esses sistemas identificariam apenas pessoas já acusadas por algum crime ou contravenção, seria possível inferir que tais sistemas reproduziriam e amplificariam o racismo, pois tenderiam a reafirmar acusações anteriores, criando um ciclo de realimentação positivo que reproduz, reafirma, naturaliza e amplifica detenções baseadas em estereótipos racistas.

6.3.2 Tons de gênero, binarismo e sexismo

O artigo “tons de gênero” de Joy Buolamwini e Timnit Gebru (2018) busca realizar uma análise interseccional de três sistemas de reconhecimento facial comerciais (Microsoft, IBM e Face++) com a funcionalidade de reconhecimento automático de gênero (AGR, *automatic gender recognition*). Reconhecendo que outras bases de dados existentes relacionando rostos humanos e gênero não são adequadamente balanceadas (Adience e IJB-A), as autoras criaram uma base de dados balanceada baseada em fotos oficiais de parlamentares de diferentes países, a Referência Piloto de Parlamentares (PPB, *Pilot Parliaments Benchmark*), e propuseram uma metodologia para análise interseccional baseada em gênero e tons de pele (claro e escuro, baseado na escala Fitzpatrick). A conclusão do artigo é que os sistemas possuem acurácia máxima em inferir gênero para rostos masculinos e brancos, sendo progressivamente diminuído para mulheres brancas, homens negros e mulheres negras. Entretanto, para além da abordagem nova e reveladora da insuficiência de se utilizar apenas a acurácia total como métrica avaliativa de um sistema, outro aspecto nos chamou atenção: o fato de terem usado “rótulos de sexo [*sex labels*] ‘masculino’ e ‘feminino’ para definir classes de gênero, dado que os benchmarks avaliados [Adience e IJB-A] e sistemas de classificação [MS, IBM e Face++] usam esses rótulos binários” (Buolamwini e Gebru 2018:4).

Segundo as autoras, “[t]odas as companhias avaliadas proveram uma funcionalidade de ‘classificação de gênero’ que usa os rótulos binários de sexo ‘feminino’ e ‘masculino’” constatando que tal visão reducionista de gênero não capturaria adequadamente as complexidades de gênero e nem consideraria identidades transgênero. As autoras constatam que “[a]s companhias não providenciam documentação para esclarecer se seus sistemas de classificação de gênero que proveem rótulos de sexo estão classificando identidade de gênero ou sexo biológico” (Buolamwini e Gebru 2018:6). Mas vale ressaltar que mesmo de uma perspectiva que separasse sexo biológico de gênero cultural, a classificação binária de sexo não englobaria indivíduos intersexo, por exemplo, e mais importante, que a distinção entre sexo e gênero nem sempre é possível ou desejada, sendo a abordagem que mais nos alinhamos aquela que entende gênero como um ator material-semiótico (Haraway 2023a).

Posteriormente, em um verbete sobre “raça e gênero” no contexto de ética em IA, Gebru (2020) explicita o ponto que queremos enfatizar:

Por exemplo [de instrumentos que perpetuam estereótipos de gênero prejudiciais, mesmo

com uma alta taxa de acertos], a tarefa de reconhecimento automático de gênero (AGR) implicitamente assume que gênero é um conceito estático e que não muda frequentemente através do tempo e culturas. Entretanto, apresentações de gênero diferem bastante entre culturas — um fato que é frequentemente ignorado por esses sistemas. Sistemas de classificação de gênero são frequentemente treinados com dados com pouquíssimos ou nenhum indivíduo transgênero ou não-binário. E as próprias saídas só classificam imagens como “masculinas” ou “femininas” [‘*male*’ or ‘*female*’]. Para comunidades transgênero, os efeitos de AGR podem ser severas, indo desde o uso de gênero incorreto para uma pessoa até sua exposição em público. Hamidi et al. nota como, de acordo com a Enquete Nacional de Discriminação Transgênero [*National Transgender Discrimination Survey*] conduzida em 2014, 56 por cento dos respondentes que foram periodicamente tratadas com gênero incorreto no espaço de trabalho tentaram suicídio. Embora haja uma boa documentação dos perigos de sistemas que realizam AGR, a utilidade dessas ferramentas é frequentemente incerta (Gebbru 2020:259–60).

Em outras palavras, não se trata aqui simplesmente de um viés, isto é, de um resultado incorreto, mas de uma lógica sexista que está profundamente concretizada nos sistemas. Aqui, gênero é binário *por design*. Há uma política inscrita na tecnologia que se explicita não só na consideração rígida de rótulos binários de gênero, mas inclusive na própria especificidade da tarefa de classificação de gênero. Segundo Gebbru (2020:260), “um dos usos mais comuns de AGR é para propaganda direcionada (e.g., mostrando àqueles percebidos como mulher um produto específico)” o que pode “perpetuar estereótipos dando mensagens subliminares sobre artefatos que homens e mulheres utilizam”.

Outro ponto levantado pela autora é sobre a proliferação de vozes femininas em artefatos em posição de subserviência (Alexa/Amazon, Cortana/Microsoft, Holly/HollyHires e Siri/Apple, na posição de assistentes digitais, mas também em filmes como Samantha (*Her*), Joi (*Blade Runner 2049*) e as IAs da Marvel: FRIDAY (*Agengers: Infinity War*) e Karen (*Spider-Man: Homecoming*)). Para Gebbru (2020:260), “sistemas de IA já estão sendo usados de maneiras humilhantes para as mulheres [mesmo] sem codificar explicitamente nomes e vozes de gênero”, como por exemplo quando “deep fakes, vídeos gerados com o uso de GANs [*generative adversarial networks*], criam conteúdo pornográfico usando rostos de mulheres comuns cujas fotos foram extraídas das mídias sociais sem consentimento”.

Ainda em relação à visão computacional, as pesquisadoras Mariah Silva e Joana Varon (2021), da *Coding Rights*, realizaram um interessante levantamento das relações entre tecnologias de reconhecimento facial e as lutas envolvendo identidades não cis-hétero. No relatório, elas concluem que:

No Brasil, como identificamos, os principais órgãos já utilizam ou implementaram o uso de tecnologias de reconhecimento facial com a finalidade de verificar identidades para acesso a serviços públicos, mas pouca transparência se tem sobre a acuracidade dessas tecnologias, sobre monitoramento de erros e falsos positivos ou falsos negativos, bem como surgem questões sobre privacidade e proteção de dados diante de práticas de compartilhamento de dados entre órgãos da administração pública e até mesmo entre entes privados. É de se destacar o fato de que o SERPRO, principal fornecedor de tecnologia de reconhecimento facial para os sistemas de verificação de identidade do governo federal, e detentor de base de dados biométricos da carteira de motorista, corre o risco de ser privatizado. Além do

que, a empresa já utiliza a base de dados da CNH como base de dados essencial para os serviços de verificação de identidade que vende tanto para o setor público como para o privado. Por fim, o caso do pedido por parte da Agência Brasileira de Inteligência (ABIN), órgão vinculado à presidência da república, para acessar dados de CNH de 76 milhões de brasileiros também traz um sinal de alerta. Surgem portanto, vários questionamentos sobre o cumprimento da Lei Geral de Proteção de Dados, principalmente no que diz respeito ao consentimento, ao respeito ao princípio da finalidade e proteção de dados sensíveis, como são os dados biométricos (Silva e Varon 2021:65).

Levando em conta que tanto gênero quanto as tecnologias são campos em disputa, as autoras reconhecem que “em geral, [as tecnologias de reconhecimento facial] reiteram e reforçam estereótipos de gênero ao classificar os rostos a partir de critérios definidos na própria criação das bases de dados que, em geral, são enviesadas tanto nas políticas de gênero, quanto nas políticas raciais, emulando uma ‘normatividade’ do norte global” (Silva e Varon 2021:66), embora também entendam, amparadas em casos como a luta trans e a expansão das possibilidades de identificação de gênero na plataforma Facebook, que “tecnologias podem sim ser utilizadas como forma de brechar a autoafirmação de gênero, mas também podem ser manipuladas como instrumentos de ampliação de lutas sociais, políticas, estéticas e tecnológicas” (Silva e Varon 2021:67).

Desse modo, ambos os exemplos mostram questões relativas à agência algorítmica que vão além da questão da acurácia e do viés, mostrando, por exemplo e de forma convergente, que tecnologias de reconhecimento facial tal como são empregadas agem de maneira racista, reforçando assimetrias em relação à raça no âmbito da justiça criminal, e sexista, apagando a diversidade de identidades de gênero e expondo sujeitos de uma maneira violenta que não considera suas posições de vulnerabilidade nem a complexidade da questão de identificação de gênero. Nesse sentido, as análises propostas por tais autoras ressoam com protestos pelo banimento da tecnologia de reconhecimento facial (eg. Silva 2022:227), uma conclusão que, partindo-se da perspectiva do viés, dificilmente seria alcançada.

Capítulo 7

O que é um algoritmo? por uma definição de mediação algorítmica

A tecnologia da informática significa, antes de tudo, aplicar radicalmente o princípio do digital

Mídias ópticas
Friedrich Kittler

A chamada "revolução do computador" não era apenas sobre o uso de *numerais binários* (dígitos binários, ou bits) para codificar linguagem humana e conteúdo analógico (digitalização), mas sobre acelerar a computação mecânica através da *lógica binária* (ou lógica Booleana). Ao contrário da visão do senso comum que enfatiza apenas a separação de hardware e software, computação digital é, na realidade, a imbricação, no mesmo meio de informação e instrução, de numerais binários e lógica Booleana – um como o complemento do outro. Em outras palavras, com a computação digital, o algoritmo de numeração (numerais binários) e o algoritmo de cálculo (lógica binária) se tornaram uma e mesma coisa

The eye of the master
Matteo Pasquinelli

Mesmo engenheiros e cientistas da computação relutam em prover uma definição definitiva para o termo algoritmo. Nos aproximando inicialmente da perspectiva "nativa", percebemos que são dois ou três os principais sentidos atribuídos pelo termo na área da computação. Por um lado, livros didáticos de matemática e de ciência da computação costumam relacioná-lo com um conjunto de operações abstratas visando um determinado fim – muitas vezes comparando-o com uma "receita de bolo". O livro *Introduction to algorithms [Introdução a algoritmos]* (Cormen et al. 2022), por exemplo, define algoritmo como "uma sequência de passos computacionais que transformam uma entrada em uma saída". Assim, haveria algoritmos de ordenamento, de otimização, de busca, etc., definidos pelas operações lógicas que os constituem. Em princípio, essa definição nem depende do computador digital, pois as regras abstratas poderiam ser computadas por qualquer meio capaz de fazê-lo (um computador digital, sim, mas também um humano com um lápis e um papel, por exemplo).

De um lado aparentemente oposto a esse, existe também uma perspectiva explicitamente materialista que considera que as operações abstratas só existem uma vez que são concretizadas em uma

infraestrutura capaz de executá-las. Uma vez que nos referimos a algoritmos no contexto de computadores digitais, essa infraestrutura diz respeito a cabos metálicos, fibras ópticas, elementos isolantes, semicondutores (como transistores e diodos de silício), e outros componentes capazes de manipular fluxos eletrônicos e eletromagnéticos ou ópticos (luz). Desse modo, as regras abstratas não passariam de uma consequência de elementos materiais, estes sim reais e determinantes de um algoritmo. Essa perspectiva é claramente apresentada em uma das patentes aqui analisadas, que dá a seguinte definição:

Um algoritmo é, aqui e geralmente, concebido como uma sequência autoconsistente de operações que levam a um resultado desejado. As operações são aquelas que exigem manipulações físicas de quantidades físicas. Normalmente, embora não necessariamente, essas quantidades assumem a forma de sinais elétricos ou magnéticos capazes de serem armazenados, transferidos, combinados, comparados e manipulados de outras formas. Às vezes, é conveniente, principalmente por motivos de uso comum, referir-se a esses sinais como bits, valores, elementos, símbolos, caracteres, termos, números ou similares.

Deve-se ter em mente, entretanto, que todos esses termos e outros semelhantes devem ser associados às quantidades físicas apropriadas e são apenas rótulos convenientes aplicados a essas quantidades. A menos que especificamente declarado de outra forma, entende-se que, ao longo da descrição, as discussões que utilizam termos como “gerar”, “fornecer”, “ajustar”, “receber”, “cancelar” ou similares referem-se às ações e aos processos de um sistema de computador [*computer system*], ou dispositivo de computação eletrônica similar, que manipula e transforma dados representados como quantidades físicas (por exemplo, eletrônicas) nas memórias ou registros do sistema de computador em outros dados representados de forma similar como quantidades físicas nas memórias ou registros do sistema de computador ou em outros dispositivos de armazenamento, transmissão ou exibição de informações. (Lewis e Price 2021:18)

Mas as duas definições não são totalmente contraditórias, em ambas, algoritmo é definido como uma operação que leva uma entrada a uma saída, sendo a divergência maior quanto a natureza dessa operação. Como tentativa inicial de resolver o problema, poderíamos conceber que um algoritmo possui dois polos, ou dois lados: um concreto, e um abstrato; um material, e outro simbólico. Como uma moeda, um seria indissociável do outro, pois as regras abstratas em um algoritmo funcionam por meio de elementos materiais, da mesma forma que estes só produzem significado se devidamente orientados por um programa¹. Nos interessa muito como a segunda definição, retirada do “discurso nativo”, chama atenção para a materialidade da computação, e do papel imprescindível de uma infraestrutura física subjacente às interfaces computacionais. Assim, muitos debates recentes, tais quais: a “nuvem” como metáfora que esconde um conjunto de computadores fisicamente localizados em propriedade de empresas; a “realidade virtual” como outro espaço, destacado da matéria; e a ingênua percepção de que andar até outro cômodo para comunicar algo seria mais custoso que o envio de uma mensagem de texto – o que mobiliza fluxos informacionais transnacionais –, são fortemente questionados.

¹Simondon (2020c) mostra como todo objeto técnico possui ambas as dimensões na medida em que surge como esquema abstrato e se concretiza progressivamente. Entretanto, há aqui uma especificidade dos algoritmos, que nem são objetos técnicos propriamente ditos (cf. Hui 2012), que complicam a análise simondoniana. Nesse caso, as regras abstratas não são apenas o esquema a partir do qual um objeto técnico se concretiza, mas, diferentemente, aquilo que orienta, a todo tempo, os elementos materiais de um objeto técnico já concretizado (o computador digital). Conferir *infra*.

Contudo, percebemos que a situação se complica ainda mais quando localizamos uma terceira definição de algoritmo, entre essas duas. É que entre as regras abstratas e o funcionamento concreto, há um código-fonte intermediário (figura 7.1). Este código é já concreto, na medida em que tanto o cartão perfurado dos primeiros computadores, quanto o texto em um editor moderno possuem uma dimensão material. No caso deste, guardado na forma de bits em elementos eletrônicos no armazenamento interno do computador. Assim, é uma concretização entre outras das regras abstratas a que nos referíamos no início – sobretudo tendo em vista a miríade de linguagens de programação diferentes que em tese possibilitam implementar o mesmo conjunto de regras abstratas e produzir o mesmo efeito material no computador. Mas é também abstrato justamente por ser o conjunto de regras ainda-não-efetivadas, ou uma potência que será atualizada no momento de *execução* do programa de computador, que ainda não entra em conflito com a realidade material que circunda a execução efetiva do programa. Assim, separar abstrato e concreto, simbólico e material, talvez não seja nada mais do que tentativas de separar analiticamente um todo mais complexo e emaranhado.

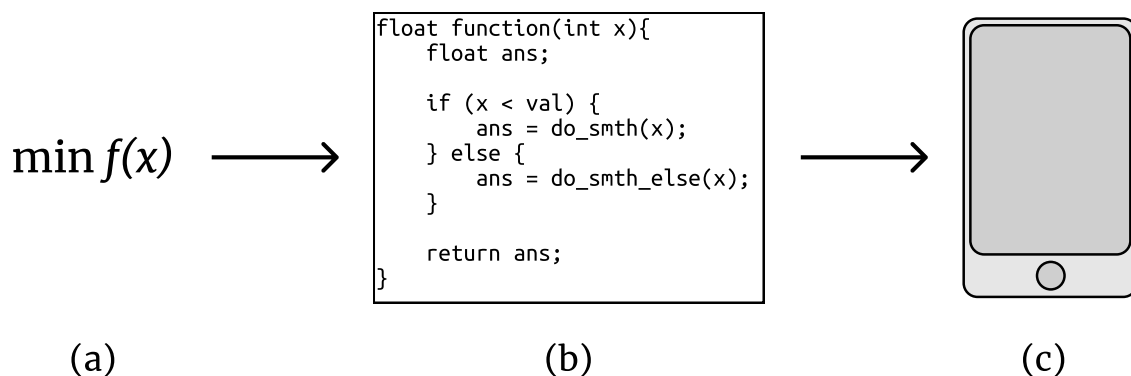


Figura 7.1: Três momentos de um algoritmo: (a) como regras matemáticas abstratas; (b) como código-fonte ou programa escrito; (c) como operações materiais em um aplicativo ou software. Entre eles há ainda traduções ativas entre um modo e outro, embora geralmente o termo algoritmo se refira aos três momentos de forma indistinta.

7.1 Computador digital, sistema de computação e algoritmo

Para nós, o computador digital² pode ser entendido como a convergência de duas tendências tecnológicas: a da *programabilidade* e a da *microeletrônica digital*. A primeira, pode ser retomada pelo menos desde a invenção do tear de Jacquard, no início do século XIX, que codificava padrões de tecelagem utilizando cartões perfurados, de certa forma convertendo “instruções manuais” em “conhecimento

²Enfatizamos que há uma especificidade do computador digital que não se reduz nem à eletrônica em geral – pois ela inclui amplificadores e outros componentes totalmente analógicos –; nem à eletrônica digital – pois no caso dos “circuitos integrados de aplicação específica [*application specific integrated circuit*, ASIC]”, dispositivos eletrônicos baseados em transistores, mas desenhados para desempenhar de forma eficiente uma aplicação particular, não há uma camada de abstração que poderíamos chamar de “algorítmica”. Já o computador digital, entendido como dispositivo eletrônico de propósito geral, é uma máquina criada visando justamente abarcar uma ampla possibilidade de usos por meio de sua programação, isto é, de reordenamentos específicos de seus componentes materiais internos que permitem computações variadas. Neste caso, o controle de seu funcionamento, as finalidades de sua execução e a determinação de seus cálculos são feitos por meio de algoritmos. Portanto o privilégio aqui do termo mediação *algorítmica* (como uma especificidade da eletrônica digital programável), em relação a outros usos possíveis (como simplesmente “eletrônica” ou “digital”).

computável” (Pasquinelli 2023:124). Em 1822, Charles Babbage desenvolveria a máquina diferencial, calculadora automática que utilizava o “método de diferenças” de Prony para calcular, entre outras, tabelas de logaritmos, de difícil cálculo manual (Pasquinelli 2023:51–53); sua sucessora, a máquina analítica – nunca efetivamente construída, mas primeiro descrita em 1837 –, também funcionaria com cartões perfurados e já pode ser entendida como um computador mecânico de propósito geral, de modo que sua concepção possibilitou que Ada Lovelace propusesse aquele que é frequentemente considerado como o “primeiro algoritmo” em 1843 (Pasquinelli 2023)³. A ideia de um “programa” armazenado em cartões perfurados, também influenciou a concepção dos cartões perfurados desenvolvidos por Herman Hollerith. A máquina de Hollerith foi utilizada pelo censo estadunidense desde 1890⁴, antecipando o que hoje viria a ser os grandes centros de dados (Pasquinelli 2023:11). Depois, os cartões perfurados seriam utilizados pela IBM (“International Business Machines”), sobre os quais se basearam um dos primeiros computadores eletromecânicos: o Harvard Mark I ou IBM Automatic Sequence Controlled Calculator, desenvolvido em 1944. Finalmente, Alan Turing proporia o que ficou conhecido como “máquina de Turing”, a descrição abstrata de uma máquina discreta universal que influenciaria a construção de computadores digitais de propósito geral – ou simplesmente computadores digitais –, “virtualmente reais”⁵. Primeiro descrito em “Sobre números computáveis com uma aplicação ao Entscheidungsproblem [On computable numbers, with an application to the Entscheidungsproblem]” (Turing 1937) e retomado no seu “Maquinaria computacional e inteligência” (Turing 1950), a máquina de Turing é caracterizada como uma finta infinita sobre a qual um cabeçote poderia ler e escrever símbolos⁶.

7.1.1 Software e hardware

A programabilidade não deve ser pensada abstraída de uma história da concretização de esquemas em objetos técnicos concretos, para usar os termos de Simondon (2020c). Friedrich Kittler (1995) talvez tenha sido uma das pessoas a melhor apresentar a questão em seu provocativo “O software não existe”. Para o autor, o design de processadores Intel nos anos 70 por meio da confecção de blueprints que ocupavam cerca de 12m² quando estendidos no chão marcam o “último ato histórico de escrita”, na medida em que o aumento da complexidade de hardware fez com que as próximas gerações de microcontrolado-

³Matteo Pasquinelli (2023, “The material tools of algorithmic thinking”), partindo de uma noção de algoritmo abstraída de seu meio material eletrônico, argumenta que rituais hindus como o Agnicayana, descrito em textos védicos datados de 800 BCE na Índia, já seriam exemplos de algoritmos.

⁴E depois, já como cartões da IBM, no censo de judeus pela Alemanha nazista (Pasquinelli 2023:53, n. 5). Não podemos também perder de vista que muitos dos avanços da computação são legatários de uma tecnologia da guerra, como mostra Haraway (2023a) em seu “manifesto ciborgue”.

⁵Para Sadie Plant (2000:330), “[d]iferentemente das máquinas anteriores, que tendem a ter um único propósito, o computador funciona como um sistema de propósito geral que pode, de fato, fazer qualquer coisa. Ele pode simular as operações, por exemplo, da máquina de escrever e, enquanto estiver executando um programa de processamento de texto, isso, na verdade, é exatamente o que ele é. Mas o computador é sempre mais - ou menos - do que o conjunto de funções atuais [actual] que ele desempenha em um determinado momento: como uma implementação da máquina abstrata de Alan Turing, *o computador é virtualmente real*”. Em *Zeros + Ones*, ela escreveria o seguinte: “se este é o começo do processo [de tecelagem, quando os elementos materiais são preparados na máquina], tudo também acabou neste ponto. Tudo o que a tecelã precisa fazer agora é executar o programa tecido com antecedência. Os padrões já estão tão bons quanto feitos. A fabricação pode muito bem já estar completa. Os softwares são virtualmente reais” (Plant 1997:81). Assim, entendemos que “virtualmente real” diz respeito às capacidades potenciais de efetivação de programas a partir da organização de elementos materiais já reunidos.

⁶Apenas apontamos que isso sugere que a questão da inteligência artificial esteve no centro dos debates e do desenvolvimento da computação desde sua origem.

res precisassem recorrer a técnicas de “design assistido por computador [*Computer Aided Design*]” que constroem automaticamente o roteamento físico entre os elementos simbolicamente conectados. Daí em diante, uma nova problemática surge com a criação de linguagens de programação que legitimaria uma “implosão do hardware por uma explosão do software”. De acordo com Kittler (1995), “[e]ssa torre de Babel pós-moderna vai de códigos de operação simples, cuja extensão linguística ainda é uma configuração de hardware passando por um montador [*assembler*] cuja extensão é esse mesmo montador”.

Constatando que a engenharia das *media* modernas busca ativamente evitar uma percepção sobre seu funcionamento, ele comenta que “nós simplesmente não sabemos o que nossa escrita faz”. O exemplo dado pelo autor é o da escrita de um texto em computador: “Para processar um texto [...], você precisa, antes de tudo, comprar alguns programas comerciais”, arquivos percebidos com a extensão .exe ou .com em um computador executando o sistema operacional Microsoft DOS [*Disk Operating System*].

De fato, [a execução do software WordPerfect, cujo comando é WP] parece trazer de volta um poder verdadeiramente mágico; o WP faz o que diz. Os arquivos de computador executáveis abrangem, em contraste não apenas com o “WordPerfect”, mas também com as palavras europeias grandes e vazias, como “Mente [*Mind*]” ou “Mundo [*World*]”, todas as rotinas e dados antigos necessários para sua autoconstituição. Certamente, tocar a sequência de letras “W”, “P” e “enter” em um teclado AT não torna a palavra [*word*] perfeita [*perfect*], mas esse simples ato de escrever inicia a execução real do WordPerfect. Esses são os triunfos do software. [...] A eletrônica como literatura; o agente linguístico governando com quase onipotência os recursos do sistema do computador, os espaços de endereço e outros parâmetros de hardware: O WP, quando chamado com o argumento X da linha de comando, mudaria a tela do monitor da cor A para B, iniciaria no modo C, retornaria finalmente para D, *et cetera ad infinitum* (Kittler 1995).

Entretanto, as ações do WP são, segundo Kittler, “virtuais”, pois elas só acontecem no DOS. E o mesmo argumento vale para o DOS que depende da BIOS, o sistema básico de entrada e saída [*Basic Input/Output System*]:

Não apenas nenhum programa, mas também nenhum sistema de microprocessador subjacente poderia ser iniciado sem a incrível capacidade de reinicialização automática de algumas funções elementares que, por segurança, *são gravadas em silício e, portanto, fazem parte do hardware*. Qualquer transformação da matéria de entropia em informação, de um milhão de transistores adormecidos em diferenças de potenciais eletrônicos, pressupõe necessariamente um *evento material chamado “reinicialização [reset]”* (Kittler 1995, grifo nosso).

Seguindo o argumento, Kittler conclui que “todas as operações de código, apesar de suas faculdades metafóricas, como ‘call’ ou ‘return’, se resumem a manipulações de strings [sequência de caracteres] absolutamente locais e, receio, a significantes de diferenças de tensão”. “Quando os significados se resumem a frases, frases a palavras e palavras a letras, não há software algum”, de modo que o aspecto linguístico da programação, o software, é totalmente dependente do contexto social e, portanto, da economia capitalista. Nesse sentido, é principalmente por um interesse econômico que “a filosofia

da comunidade de computadores tende a obscurecer sistematicamente o hardware pelo software, os significantes eletrônicos pelas interfaces entre as linguagens formal e cotidiana”:

Em primeiro lugar, em um nível intencionalmente superficial, as interfaces gráficas de usuária perfeitas, uma vez que dispensam a própria escrita, escondem uma máquina inteira de seus usuárias. Em segundo lugar, no nível microscópico do próprio hardware, o chamado software de proteção foi implementado para impedir que “programas não confiáveis” ou “usuárias não confiáveis” tenham acesso ao kernel e aos canais de entrada/saída do sistema operacional (Kittler 1995).

Para Kittler isso prova um triunfo do software como uma inversão da hipótese Church-Turing, que identifica a efetividade de um cálculo com a possibilidade de computação por uma máquina, ou seja, condiciona o software ao hardware. As linguagens de programação de alto nível, por serem uma abstração do código de máquina, funcionam criptograficamente escondendo o funcionamento de baixo-nível: “As funções unidirecionais, em outras palavras, ocultam um algoritmo de seu próprio resultado. Para o software, esse efeito criptográfico oferece uma maneira conveniente de contornar o fato de que, em virtude da prova de Turing, o conceito de propriedade mental aplicado a algoritmos tornou-se sem sentido” (Kittler 1995).

Visando contestar a situação atual em que o software se tornou uma criação comercial, cujo ato final foi sua apropriação pelo SPI (ver capítulo 2), Kittler argumenta que é o hardware que deveria ser considerado como portador da propriedade de programabilidade:

Essa propriedade importantíssima de ser programável não tem, com toda evidência, nada a ver com software; é um recurso exclusivo do hardware, mais ou menos adequado para abrigar algum sistema de notação. Quando Claude Shannon, em 1937, provou, no que provavelmente é a tese de mestrado mais importante já escrita, que simples relés de comutação telegráfica podem implementar, por meio de suas diferentes interconexões, toda a álgebra booleana, esse sistema de notação física foi estabelecido. E quando o circuito integrado, desenvolvido na década de 1970 a partir do transistor de Shockley, combinou em um chip o silício como um resistor controlável com seu próprio óxido como um isolador quase perfeito, a programabilidade da matéria pôde finalmente ‘assumir o controle’, como Turing havia previsto (Kittler 1995).

Para Kittler, a relação antagônica entre continuidade térmica e descontinuidade elétrica foi o que permitiu aos circuitos integrados microeletrônicos atingirem o estatuto de “máquina discreta universal”. “Uma fechadura de combinação”, por exemplo, “é um autômato finito, mas não é normalmente decomponível em um conjunto básico de componentes do tipo elementar que podem ser reconfigurados para simular um sistema físico arbitrário”. Ao contrário, “um computador digital usado para simular uma fechadura de combinação é estruturalmente programável, pois o comportamento é obtido por meio da síntese de um conjunto canônico de componentes de comutação primitivos” (Kittler 1995). O texto termina com a provocativa hipótese sobre a possibilidade de um computador analógico, funcionando por meio de operações contínuas e, portanto, não programável:

É exatamente essa conectividade máxima que define os sistemas não programáveis, do ponto de vista físico, sejam eles ondas ou seres. É por isso que esses sistemas apresen-

tam taxas de crescimento polinomial em complexidade e, conseqüentemente, somente os cálculos feitos em máquinas não programáveis poderiam acompanhá-los. Com toda evidência, esse tipo de máquina hipotética, mas demasiadamente necessária, constituiria um hardware puro, um dispositivo físico trabalhando em meio a dispositivos físicos e sujeito aos mesmos recursos limitados. *O software, no sentido usual de uma abstração sempre viável, não existiria mais.* Os procedimentos dessas máquinas, *embora ainda abertos a uma notação algorítmica, teriam que funcionar essencialmente em um substrato material cuja própria conectividade permitiria reconfigurações celulares.* E mesmo que o “substrato também possa ser descrito em termos algorítmicos, por meio de simulação”, sua “...caracterização é de tão grande importância para a eficácia [...] e tão intimamente ligada à escolha do hardware, que...” programá-los terá pouco a ver com máquinas de Turing aproximadas (Kittler 1995).

Dessa perspectiva, a programabilidade deve ser pensada levando em conta a história de sua materialidade que teve na *eletrônica digital* um marco importante, possibilitando uma mudança de escala e de velocidade sem precedentes e que estabiliza o que entendemos hoje por *computador digital*. Desse modo, a “revolução microeletrônica”, ou melhor, a *virada cibernética* (Garcia dos Santos 2003a), foi fundamental para a computação por encontrar na eletrônica digital um substrato profícuo para a materialização das ideias de programabilidade, informação e código. Nem os cartões perfurados sozinhos, nem a eletrônica por si só, que pode ser analógica, nem as válvulas triódicas, precursoras dos transistores de estado sólido, possuíam a maleabilidade, a velocidade e a capacidade de miniaturização que as técnicas de impressão nanométrica de relês binários em semicondutores – i.e., transistores de silício que chaveiam entre zero e um – possuem. Foram elas que possibilitaram a “aceleração da aceleração” descrita por Garcia dos Santos (2003a) e que tornaram o computador digital, hoje, uma tecnologia ubíqua, funcionando como técnica central das sociedades de controle, por operar justamente por “controle direto” e “comunicação instantânea” (Deleuze 2013). Em nossa leitura, é essa onipresença e instantaneidade da computação digital – visível no fato de que todos carregamos smartphones nos bolsos e estamos “sempre online” (boyd 2012) – que cria uma nova camada na realidade contemporânea, aquela dos processos algorítmicos, e que nos confundiu a conceber outrora realidades outras – seja na noção de “ciberespaço” que abstrai a natureza lógica da geográfica⁷, seja na de “realidade virtual” como aquela que prescindir da materialidade do mundo físico⁸.

É sobre a mesma questão da separabilidade entre software e hardware que recai a crítica de John Searle (1980) sobre a ideia de que um autômato que passa no “teste de imitação” de Turing poderia ser considerado inteligente, ao menos no sentido que os defensores da IA forte lhe atribuíam. O argumento do autor, que gostaríamos de resgatar brevemente, é que ao separar hardware e software, mente e corpo, formal e material, perde-se justamente o fenômeno:

O problema com o simulador de cérebro é que ele está simulando as coisas erradas sobre o cérebro. Enquanto ele simular apenas a estrutura formal da sequência de disparos de

⁷Inspirado no uso que o escritor de ficção científica William Gibson dá ao termo “ciberespaço” – “Burning Chrome” (1982) e *Neuromancer* (1984) –, Bruce Sterling (1994:11) propõe a interessante definição de que “o ciberespaço é o ‘lugar’ onde parece ocorrer uma conversa telefônica. [...] Entre um telefone e outro”

⁸Cabe notar o deslocamento que ocorre atualmente entre o termo “realidade virtual”, que se opõe à realidade háptica, olfativa, gustativa, etc., para o termo “realidade aumentada”, que embora não enfatize sua materialidade própria, também não visa ser algo totalmente separado da realidade física.

neurônios nas sinapses, ele não terá simulado o que importa sobre o cérebro, ou seja, suas propriedades causais, sua capacidade de produzir estados intencionais [...] Seja o que for que a intencionalidade seja, ela é um fenômeno biológico e é tão provável que seja causalmente dependente da bioquímica específica de suas origens quanto a lactação, a fotossíntese ou qualquer outro fenômeno biológico. Ninguém suporia que poderíamos produzir leite e açúcar executando uma simulação de computador das sequências formais da lactação e da fotossíntese, mas, no que diz respeito à mente, muitas pessoas estão dispostas a acreditar em tal milagre por causa de um dualismo profundo e permanente: a mente que elas supõem é uma questão de processos formais e é independente de causas materiais bastante específicas, da mesma forma que o leite e o açúcar não são (Searle 1980:421, 424).

Do mesmo modo, acreditamos que descrever um algoritmo como um conjunto de instruções separado da materialidade, como no caso da analogia da “receita de bolo”, mostra-se estrategicamente inócua. Por outro lado, há um conjunto de fenômenos (modulação de atividade, vieses em resultados, etc.) que já percebemos hoje como sendo algorítmicos e que *acreditamos não ser suficientemente explicado pela sua redução ao nível material elementar* (transistores, energia elétrica ou silício). Portanto, acreditamos que definir algoritmo como uma execução material abstraída da camada “simbólica” tampouco resolve a questão. Alternativamente, gostaríamos de insistir na realidade dos processos algorítmicos, que existem *no vínculo entre regras formais visando um fim (programação via lógica booleana) e elementos materiais operando especificamente (microeletrônica digital)*.

7.1.2 O computador digital como um indivíduo técnico

Dessa perspectiva, o ponto de contato entre ambas as tendências expostas é o computador digital. Todos os processos algorítmicos acontecem no interior, por meio de, ou ainda como uma configuração ou modo específico de operação do computador digital em sentido amplo.

No que tange a sociologia da tecnologia, a caracterização dos objetos técnicos feita por Simondon (2020c) é bastante interessante, pois busca entendê-los pelo exame do processo real de invenção que ocorre no interior de um coletivo: “acima da comunidade social de trabalho e além da relação interindividual que não é sustentada por uma atividade operatória, institui-se um universo mental e prático da tecnicidade, no qual os seres humanos se comunicam através daquilo que inventam”, dessa maneira, “[o] objeto técnico como algo que foi inventado, pensado e desejado, assumido por um sujeito humano, torna-se o suporte e o símbolo da relação que gostaríamos de chamar de *transindividual*” (Simondon 2020c:359)⁹. Não obstante, não encontramos na obra do autor uma definição específica do computador digital, de modo que esboçaremos sua teoria para arriscar uma caracterização não definitiva de sua existência como um indivíduo técnico¹⁰.

⁹Sobre o transindividual, Simondon comenta: “Por intermédio do objeto técnico, cria-se então uma relação inter-humana que é o modelo da *transindividualidade*. Podemos entender por isso uma relação que não relaciona os indivíduos por meio de sua individualidade constituída, separando uns dos outros, nem por meio do que há de idêntico em todo sujeito humano – por exemplo, as formas *a priori* da sensibilidade –, mas por meio da carga de realidade pré-individual, da carga de natureza que é conservada com o ser individual e que contém potenciais e virtualidades” (Simondon 2020c:360).

¹⁰A menção mais próxima que encontramos foi aos “dispositivos de informação [*dispositifs à information*]” em Simondon (2005): “O dispositivo básico das máquinas de informação é o modulador, graças ao qual uma incidência admitida na entrada (entrada, controle), por menor que seja sua energia portadora, tem o poder de governar o uso de uma energia muito maior, fornecida pela fonte de alimentação. Na saída estão as características formais da incidência (sinal, informação) e as caracterís-

Para Simondon (2020c), o objeto técnico é definido pelo processo ontogenético em que um esquema abstrato é progressivamente concretizado, isto é, em que uma ideia de funcionamento tende a se tornar um objeto concreto, como uma rocha ou outros objetos ditos naturais, sem nunca, entretanto, chegar a tal ponto, sendo o estatuto do objeto técnico (artificial) o de uma espécie de “natureza suscitada” pela invenção humana¹¹. Por isso, “o objeto técnico individual não é tal ou qual coisa, dada aqui e agora, e sim aquilo de que há uma gênese”, “é aquilo que não é anterior a seu devir, mas está presente em cada etapa desse devir; o objeto técnico é uno é unidade de devir”, ele “existe, pois, como tipo específico, obtido no final de uma série convergente” (Simondon 2020c:56, 60).

Mas a concretização, ao fazer o objeto técnico pensado devir realidade concretizada, cria novos problemas (tensões) que alteram a ideia inicial de objeto técnico tal qual abstratamente imaginado pela inventora. Em realidade, a invenção técnica supõe mesmo a antecipação dessas tensões, elas devem ser “encenadas” pela inventora por meio de uma “imaginação criadora”¹². Surge então uma sinergia progressiva entre o objeto, que se concretiza, e seu meio “tecnogeográfico”, que se torna a ele associado, de modo que o objeto técnico se torna dependente das condições específicas do meio para que o funcionamento possa acontecer de maneira adequada. Ele forma *sistema* com seu meio associado, ou seja, constituem conjunto informacionalmente autônomo (Simondon 2020b). Nesse caso, em que há dependência do objeto técnico em relação ao seu meio, podemos ainda dizer que ele é um *indivíduo técnico*, ou seja, possui uma ressonância interna com o meio associado que dá ao sistema indivíduo-meio uma autonomia informacional: desse modo, sua existência técnica não é nem inferior à humana – como do caso das ferramentas que manuseamos e dos *elementos técnicos* em geral, em que há dependência da ação humana para seu funcionamento –, nem supra-humana – os *conjuntos técnicos*, como os laboratórios, que, embora sejam capazes de criar novos elementos técnicos, também dependem da ação humana mediadora, não constituindo um sistema informacionalmente autônomo. Nessa posição, o indivíduo técnico compartilha uma mesma escala de funcionamento em relação ao humano, que passa a ser organizador das trocas informacionais entre indivíduos técnicos. O objeto técnico individualizado ocupa a posição central no trabalho técnico outrora desempenhado pelo humano, de modo que este se relaciona principalmente com os elementos técnicos – auxílio e manutenção, abaixo da individualidade técnica – e com os conjuntos técnicos – regulação e organização de indivíduos em conjuntos. Comentando o livro de Simondon, Garcia dos Santos (1981) aponta que “às incidências sociais, é principalmente ao nível do indivíduo técnico que se manifesta o problema da civilização industrial”, pois “o indivíduo técnico não

ticas energéticas da fonte de alimentação” (Simondon 2005:221). Ele as divide em “máquina de calcular [*machine à calculer*]”, “amplificadores e reguladores [*amplificateurs et régulateurs*]” e “máquinas de interpretação [*machines à interprétation*]”. Sobre as últimas, o autor comenta: “O tipo atualmente mais conhecido dessas máquinas é a máquina de leitura [*machine à lire*]; mas podemos conceber uma generalização da máquina de informação na forma de uma máquina de interpretação, a partir de uma função elementar como a função de auto-correlação” (Simondon 2005:225). Acreditamos que esta descrição se aproxima do funcionamento atual do computador digital, que opera por meio da “interpretação” de uma linguagem de alto nível em linguagem de máquina.

¹¹Em oposição à concretização do objeto técnico, é possível falar de artificialização do objeto natural sempre que a intervenção humana causa um desvio de função, reduzindo a autonomia do objeto relativa ao meio e desadaptando-o, como mostra o caso da flor na estufa que só produz pétalas e não gera descendentes, e passa a depender de um meio muito mais controlado e mediado pela ação humana para existir (Simondon 2020c:91–92).

¹²“[O]s elementos que constituirão materialmente o objeto técnico, e que existem separados entre si, sem um meio associado antes da constituição do objeto técnico, devem organizar-se em relação uns aos outros em função da causalidade circular que existirá quando objeto estiver constituído; aqui, portanto, trata-se de um condicionamento do presente pelo futuro, pelo que ainda não existe” (Simondon 2020c:107).

é mais o homem e sim a máquina; a máquina substitui o homem funcionalmente enquanto indivíduo, agora é ela quem conserva as ferramentas”, o humano se torna “organizador das relações entre os níveis técnicos em vez de ser ele próprio um dos níveis técnicos” (Garcia dos Santos 1981:24, 25).

O exemplo mais evidente de indivíduo técnico é uma máquina térmica que converte energia térmica em trabalho mecânico em condições específicas. Outro exemplo apresentado extensivamente pelo autor é a turbina de Guimbal, cuja a sinergia entre água, óleo e turbina criam uma ressonância entre o indivíduo e seu meio (Simondon 2020c:102 *et seq.*). “Diríamos que há indivíduo técnico quando o meio associado existe como condição *sine qua non* de funcionamento; há conjunto caso contrário” (Simondon 2020c:111). E o computador digital? Seria um indivíduo ou um conjunto técnico? O exemplo mais próximo que Simondon dá é de uma calculadora:

As calculadoras modernas não são puros autômatos: são seres técnicos que, acima de seus automatismos de soma (ou de decisão, pelo funcionamento de osciladores elementares), possuem vastíssimas possibilidades de comutação de circuitos, que permitem codificar o funcionamento da máquina, restringindo sua margem de indeterminação. Graças a essa margem primitiva de indeterminação, a própria máquina pode extrair raízes cúbicas ou traduzir um texto simples, composto de um pequeno número de palavras e construções, de uma língua para outra (Simondon 2020c:47).

Em argumentando que “o automatismo é um grau de perfeição técnica bastante baixo”, pois “[a]utomatizar uma máquina exige sacrificar muitas possibilidades de funcionamento, muitos usos possíveis”, o autor continua: “Graças a essa margem de indeterminação, e não aos automatismos, as máquinas podem ser agrupadas em conjuntos coerentes e trocar informações umas com as outras por intermédio do coordenador, que é o interprete humano”. E, ainda, “[m]esmo quando a troca de informações é direta entre duas máquinas (como entre um oscilador piloto e outro oscilador, sincronizado por impulso), o homem intervém como o ser que regula a margem de indeterminação, para que ela se adapte ao melhor câmbio possível de informações” (Simondon 2020c:46, 47). Tomada enquanto “máquina”, entendemos que Simondon posiciona as calculadoras como indivíduos técnicos e a sua relativa programabilidade como uma abertura para que o humano intervenha orientando-a, fazendo convergir sua margem de indeterminação. Acreditamos que algo similar ocorre com o computador digital¹³.

Duas considerações ainda são necessárias para que consideremos o computador digital como um indivíduo técnico. Em primeiro lugar, a descrição simondoniana do audiômetro como conjunto técnico algumas páginas à frente (Simondon 2020c:111 *et seq.*), por argumentar que apesar de sua relativa

¹³Embora Simondon tenha falado pouco da eletrônica digital tal qual extensivamente utilizada hoje em dia, algumas passagens de sua obra deixam bastante claro que ele não ignorava as transformações que ela poderia alavancar. Uma delas é a seguinte: “No momento em que as técnicas elétricas atingiram seu pleno desenvolvimento, elas produziram, como elementos, novos esquemas que deram início a uma nova fase: primeiro houve a aceleração de partículas [...]; em seguida, e de maneira notável, graças à metalurgia elétrica, veio a possibilidade de extrair metais como o silício, que permitiam a transformação da energia radiante da luz em corrente elétrica, com um rendimento (6%) que já atingia uma taxa interessante para aplicações restritar e não era muito inferior ao das primeiras máquinas a vapor. *A pilha de silício puro, produto dos grandes conjuntos eletrônicos industriais, é o elemento ainda não incorporado a um indivíduo técnico; ainda é apenas um objeto de curiosidade, situado no ponto extremo das possibilidades técnicas da indústria eletrometalúrgica. Mas, com o desenvolvimento da produção e do uso da eletricidade industrial, é possível que venha a se tornar o ponto de partida de um desenvolvimento análogo ao que conhecemos e que ainda não está concluído*” (Simondon 2020c:121–22).

individualidade ele ainda poderia ser separado em duas partes com funcionamentos distintos, constituindo um conjunto. Essa passagem nos leva a tomar com cautela qualquer objeto que denominamos amplamente como “computador digital” por um indivíduo. Em segundo lugar, a descrição da “relação homem-máquina” que dá mais precisão para a posição que o humano pode ocupar na relação com um indivíduo técnico:

Todavia [a posição atual do humano como acima e abaixo da individualidade técnica], isso não significa que o homem [*sic*] não possa ser um indivíduo técnico e trabalhar em ligação com a máquina. Essa relação homem-máquina se realiza quando o homem, através da máquina, aplica sua ação ao mundo natural. Então a máquina torna-se veículo de ação e de informação, numa relação em três termos – homem, máquina e mundo –, na qual a máquina se situa entre o homem e o mundo. Nesse caso, o homem conserva certos traços de tecnicidade definidos, em particular, pela necessidade de aprendizagem. [...] Poderíamos dizer que, nesse caso, o homem é portador da máquina, permanecendo a máquina como portadora de ferramentas. Essa relação é parcialmente comparável, portanto, à da máquina-ferramenta, se entendermos por máquina-ferramenta aquela que não comporta autorregulação. O homem continua no centro do meio associado nessa relação: a máquina-ferramenta é aquela que não tem regulação interna autônoma e precisa de um homem para fazê-la funcionar (Simondon 2020c:111–12).

No caso da máquina-ferramenta, o humano “intervém nela como ser vivo” (Simondon 2020c:112), e tem uma função homeostática, de regulação e manutenção correlata à regulação vital. Mas no caso do computador, nos parece que a atividade de programação seria justamente uma ação do humano levado a cabo por meio da máquina, com a especificidade de que a ação é o próprio reordenamento do computador que, como consequência, realiza tais programas, na forma de um algoritmo. Durante a programação, nos parece que a desenvolvedora ocupa uma posição de indivíduo técnico. Mas, durante a execução do programa, a operação técnica é realizada pelo computador digital, de modo que em um sistema de computação, as usuárias conectadas a ele fariam parte de seu meio associado (conferir a participação do humano como servidão maquínica no AdM na parte II).

Seguindo a pista de Schiavetto Amancio (2024)¹⁴, devemos tomar a cautela de considerar o computador digital como o indivíduo técnico num sistema de computação informacionalmente autônomo que se conecta a outros periféricos, o que o Schiavetto Amancio denomina “microcomputador”¹⁵, e não em

¹⁴Referimos a leitora às seções 4.2 e 4.3 de Schiavetto Amancio (2024) para uma exposição alternativa à nossa, mais detida aos aspectos técnicos da microeletrônica e que, portanto, pode complementar a leitura deste capítulo.

¹⁵Como citado em Schiavetto Amancio (2024:252): “O microcomputador tem sua forma definida por um sistema composto por três subsistemas (microprocessador, memórias RAM e ROM e controladores de entrada e saída) e demais elementos microeletrônicos em associação. Pode ser compreendido como um indivíduo técnico em virtude de sua autorregulação de seu funcionamento, independente de seus fabricantes ou de outros seres. Parte de suas peças não existem apenas nos microcomputadores, algumas são idênticas às existentes em outros seres técnicos e outras possuem aperfeiçoamentos exclusivos para a forma microcomputador. O que atribui individualidade ao objeto técnico como microcomputador é justamente a sua associação exclusiva de elementos técnicos que constituem um específico sistema em funcionamento. O microprocessador, embora também de existência sistêmica, não possui uma autorregulação desvinculada das memórias RAM e ROM e dos controladores de periféricos, sendo impossível atribuir uma individualidade de funcionamento independente desses dois objetos técnicos. Encontramos uma individualidade no microcomputador, sendo o microprocessador uma peça fundamental, um subsistema plenamente integrado, o que se aproxima mais da definição simondoniana de elemento técnico” (Schiavetto Amancio 2014:22).

sua forma prática, isto é, computador de mesa, notebook, tablet ou microcontrolador (conferir especialmente [Schiavetto Amancio 2024:216](#), figura 11). Assim, quando nos referimos ao computador digital como indivíduo técnico, é sobre esse recorte que operamos. As formas práticas utilizadas em interação com usuárias humanas estão no nível dos conjuntos.

7.1.3 Elementos eletrônicos, computadores digitais e conjuntos algorítmicos

A partir dessas considerações, acreditamos que podemos falar em um sistema de computação sempre que um computador digital ocupa uma posição central em uma operação de computação: seja em um smartphone, em um computador de mesa, ou mesmo em um servidor de uma empresa com o qual interagimos por meio de cabos metálicos e fibras ópticas.

Desse modo, podemos tomar o computador digital como um indivíduo técnico que ocupa uma posição central na contemporaneidade, sobretudo tendo em vista os processos algorítmicos como os tratados ao longo desta dissertação. Em uma escala inferior ao indivíduo, tanto o que chamamos de “software”, quanto nós mesmos, humanos que interagimos com um computador digital, nos tornamos fluxos eletromagnéticos (luz) e principalmente eletrônicos (eletricidade) que são direcionados, padronizados, classificados e modulados nos sistemas de computação. Ao nosso ver, é isso que dá à existência atual um estatuto que permite falar de uma “vida eletrônica”, como essa parcela de vida que acontece nos sistemas de computação ([Ferreira 2024](#)). Enquanto nível técnico inferior à individualidade, os elementos técnicos (transistores, diodos, resistores, capacitores, etc.), operando sobre os elétrons, são os portadores da tecnicidade: controlam fluxos elétricos e eletromagnéticos e fazem surgir padrões eletrônicos que *determinam o funcionamento do computador digital* – e, portanto, de uma parcela de nossa humanidade contemporânea, de uma subjetividade dividida que circula no circuito integrado. Os elementos técnicos são o que condiciona o funcionamento/a agência do computador-humano digital.

Já em uma escala superior ao indivíduo, os conjuntos formados não só pelas partes que compõem um computador de mesa em se conectando com o computador digital propriamente dito (teclado, mouse, tela, etc.), mas também pelos humanos a ele conectados, formam um agenciamento humano-máquina, um sistema de computação, embora, no esquema simondoniano, talvez seria mais preciso denominá-los conjuntos de computação. Da perspectiva da centralidade da operação de computação, o próprio humano se torna humano computadorizado: é percebido pela máquina por meio de sensores e dispositivos de entrada (*input*) e é afetado por ela por meio de atuadores e dispositivos de saída (*output*). O humano que interage com um computador digital se torna um padrão eletrônico, interagindo com ele por meio de informação digital (o que foi tratado com mais profundidade em relação ao AdM nos capítulos que compõem a parte II). Assim, tudo o que ocorre em uma escala técnica superior ao indivíduo é, da perspectiva da computação, informação digital, isto é, singularidades do sistema que incitam processos de tomada de forma de um computador digital (e, portanto, de usuárias a ele conectadas). Esses processos, ocorrendo por meio de um computador digital envolvido por um conjunto que constitui sua forma prática, e em conexão com outros computadores e usuárias que também funcionam-agem como indivíduos, são de natureza supraindividual. Justamente esses processos são os que gostaríamos de chamar *algorítmicos*. A aprendizagem de máquina, por exemplo, não ocorre somente por meio do computador, mas em conexão com outros elementos tomados como informação de entrada e saída. O mesmo pode ser dito para qualquer programa de computador: se não há a interação com ao menos um dispositivo de entrada ou de saída, o reordenamento dos padrões eletrônicos do computador não

resultam em nada, seu efeito é nulo. É por isso que podemos dizer que a programação, ou a computação sobre a informação digital, exige uma ordem de realidade acima daquela do indivíduo, são da natureza do conjunto.

Desse modo, propomos conceber os algoritmos como uma unidade operativa de um sistema de computação digital que, analisados de uma perspectiva de escala superior ao indivíduo, agem por meio de computadores digitais. Enquanto isso, o próprio computador digital tomado como um indivíduo técnico, se torna o ponto central de operação ou de trabalho em sistemas de computação, funcionando de maneira interconectada a outros indivíduos em conjuntos técnicos e realizando os algoritmos. Já os transistores de silício, portadores da tecnicidade subjacente aos processos computacionais, são a base elementar em uma escala inferior de individualidade técnica. Assim, transistores sustentam a operação de um computador digital que, por sua vez, sustentam a operação dos algoritmos. Dessa perspectiva, concluímos finalmente que, por operarem na escala em que a tomada de forma se dá pela informação digital, os algoritmos podem ser considerados os *mediadores da informação digital*.

7.2 Algoritmo e agência

Retomando a questão inicial do capítulo, podemos dizer que um algoritmo não é nem uma entidade imaterial – “uma receita de bolo”, uma sequência de regras abstratas visando um fim, ou mesmo um objeto digital (Hui 2012) –; nem simplesmente as operações materiais reduzidas à identidade com o hardware. Ao contrário, ele é aquilo que media a informação digital, operação que tem um significado, portanto, uma finalidade, mas que é conduzida por meio de elementos físicos de um computador. O algoritmo está no vínculo das possibilidades concretas de funcionamento de um computador digital com a descrição abstrata de uma finalidade ou ação. Sua execução resulta simultaneamente ação e significação: ela é informação, na medida em que sua realização muda o estado do sistema de computação.

Desse modo, mais do que dois polos (concreto e abstrato), os algoritmos se apresentam como entidades mistas, que relacionam ambas as dimensões. De modo a não encerrar a polissemia do termo – objetivando criar uma conceitualização que se enderece tanto à lógica, quanto às operações materiais e à sua descrição como código-fonte –, nos referiremos como *algorítmico* àquilo que relaciona transversalmente lógica e matéria em um conjunto de computação. Isso implica que há uma dimensão abstrata e simbólica dos algoritmos, relacionada com regras, finalidades, intenções e desejos; tanto quanto há uma dimensão concreta, material, relacionada ao funcionamento e à ação. Ambas são acepções parciais de um misto efetivo que é o processo algorítmico.

Há ainda um outro limite da apresentação feita inicialmente, baseada em definições êmicas. Tanto o livro-texto de computação, quanto a patente, relacionam operações a uma finalidade. Mesmo Mittelsadt et al. (2016), consideram a agência algorítmica atrelando-a à materialização dos valores de quem o criou. Entretanto, é importante ressaltar que esta nunca pode ser totalmente planejada por e reduzida a um agente humano nos bastidores. Kitchin (2014b) insiste na natureza contingente, mutável/mutante e performativa dos algoritmos que, além disso, estão sempre inseridos em agenciamentos sociotécnicos mais amplos. Comportamentos inesperados emergem do funcionamento algorítmico, sua estrutura é constantemente retrabalhada por desenvolvedoras e mantenedoras de software e a interação com usuárias leva a resultados dificilmente antecipáveis. Desse modo, sua agência não pode nunca ser reduzida a um único humano (ou grupo de pessoas) externo que planeja sua finalidade.

O significado e a ação produzidos por um algoritmo são dependentes de uma série de agências conectadas: das desenvolvedoras que o programa e o mantém, das usuárias que o utilizam, da empresa que o detém, mas também da própria linguagem de programação escolhida, dos sensores e atuadores que servem como entrada e saída de seu funcionamento, e ainda muitas outras. Portanto, não há nunca um algoritmo isolado, mas uma rede heterogênea de componentes humanos e não humanos que se adensa mais ou menos em um ponto. Os limites de um algoritmo são fluidos. Um algoritmo é um ponto privilegiado em uma rede, mais que uma área com limites bem definidos. Muitas vezes, esse centro pode ser entendido como um “centro de cálculo” (Latour 2017), na medida em que o processamento de dados em certos algoritmos pode ser visto como uma capitalização de inscrições de uma periferia. Dessa maneira, também somos potencializados ao manuseá-los, pois, como mostra o exemplo do viajante com o mapa analisado por Ferreira (2017), manipular um mapa é ser capaz de envolver com as mãos uma representação do território que nos envolve, desse modo, manusear um mapa é como “colocar um pé no centro de cálculo e, assim, impulsionar sua ação na periferia” (Ferreira 2017:110). Usar o Google Maps é isso. Somos potencializados na medida em que interagimos com seus algoritmos, mas, por outro lado, também somos fonte de captura de inscrições para a Google, funcionando alternadamente como centro e periferia. Centro e periferia são conceitos relacionais, operacionais.

7.2.1 Mediação técnica e mediação algorítmica

Se algoritmos são mediadores da informação digital, de que forma podemos caracterizar a sua agência? Bruno Latour (e.g. 1994) propõe o termo mediação técnica para caracterizar a agência do que ele denomina “não humanos”. Do modo como entendemos, a mediação (técnica) indica o modo como o meio (técnico) também age, deixando de ser simplesmente um meio neutro e tornando-se simultaneamente fim, isto é, contribuindo ativamente para o resultado final da ação/significação que ele media. Nas palavras de Latour, “mediadores transformam, traduzem, distorcem e modificam o significado ou os elementos que supostamente veiculam” (Latour 2012:65). Desse modo, “[o] termo ‘mediação’, em contraste com ‘intermediário’, significa um evento ou um ator que não podem ser exatamente definidos pelo que consomem e pelo que produzem. Se um intermediário é plenamente definido por aquilo que o provoca, uma mediação sempre ultrapassa sua condição” (Latour 2017:363). Já o termo “técnico”, um “bom adjetivo”, mas “vil substantivo”, é, para Latour (Latour 2017:226–28) polissêmico, podendo significar simultaneamente: (1) uma “série de subprogramas embutidos” em uma ação, (2) um “papel subordinado” (de uma pessoa, habilidade ou objeto), (3) um “obstáculo” no funcionamento, (4) uma “incerteza quanto ao que seja um meio e ao que seja um fim”, e (5) “um tipo muito específico de *delegação*, movimento, deslocamento para baixo¹⁶ que se entrecruza com entidades dotadas de propriedades, espaços, tempos e ontologias diferentes, as quais são levadas a partilhar o mesmo destino e a criar, assim, um novo [actante]” ou agente. Enfatizamos ainda que, em “mediação técnica”, o adjetivo está fortemente relacionado com a presença de elementos não humanos. Há mediação técnica quando ações de outros momentos, outros espaços e outros agentes continuam a agir “aqui e agora”, mas na condição de que eles são trocados, traduzidos, delegados ou deslocados por e para agentes não humanos (Latour 2017).

Uma outra forma de descrever uma mediação técnica é como um *agenciamento de subprogramas*

¹⁶Termo “[...] da semiótica referent[e] ao ato de significação pelo qual um texto correlaciona diferentes quadros de referência (aqui, agora, eu): diferentes espaços, diferentes tempos, diferentes aspectos. [...] [Q]uando o material expressivo é inteiramente modificado, [chama-se] deslocamento para baixo” (Latour 2017:360).

de ação¹⁷. Partindo daí, Latour (2017) propõe ainda quatro significados de mediação técnica. Primeiramente, uma mediação técnica pode ser entendida como seu próprio programa de ação. Como sempre há incerteza sobre o resultado final em relação aos dos subprogramas, ela pode ser entendida como *translação/tradução* destes, e portanto, a responsabilidade para com a ação é compartilhada pelos diversos agentes que participam do processo e constroem o programa de ação final¹⁸. O segundo significado é *composição* (de agentes e subprogramas): pois, a depender da granularidade da descrição, a ação pode sempre ser vista como propriedade de entidades associadas. O terceiro significado é o de uma *caixa-preta reversível*, isto é, a possibilidade de identificar múltiplos elementos de diferentes tempos e espaços que são montados em uma “caixa preta”, tida como responsável por determinada ação. Finalmente, o quarto significado de mediação técnica é *delegação*:

graças ao deslocamento para baixo, outra combinação de ausência e presença torna-se possível. No caso da delegação, não se trata, como na ficção, de eu estar aqui ou em outra parte, de ser eu mesmo ou outra pessoa, mas de uma ação muito antiga de um ator já desaparecido continuar ativa aqui hoje e em relação a mim. Vivo no meio de *delegados* técnicos, misturo-me aos não-humanos. [...] Pense na tecnologia como esforço *congelado*. Considere a própria natureza do investimento: um curso regular de ação é suspenso, um desvio por vários tipos de [actantes] é iniciado e o retorno é um novo híbrido que transfere atos passados para o presente, permitindo a seus muitos investidores desaparecer sem deixar de estar presentes (Latour 2017:224).

Dessa forma, explicita-se como a moralidade (e política) pode estar “inscrita nas coisas” (Latour 2002), já que mediações técnicas podem “nos fazer fazer”, isto é, alterar nosso próprio “programa de ação”¹⁹. Acreditamos que o mesmo é verdade quando os algoritmos são tomados como mediadores. São duas as principais conclusões que resgatamos de Latour em sua concepção de mediação técnica:

1. A de que todos os agentes envolvidos em uma mediação técnica se influenciam mutuamente, contribuindo e se tornando co-responsáveis, cada um de seu modo, para o programa de ação

¹⁷Programa de ação (e antiprograma) são “[t]ermos da sociologia da tecnologia que têm sido usados para emprestar caráter ativo, e muitas vezes polêmico, aos artefatos técnicos. Cada dispositivo antecipa o que outros atores, humanos ou não humanos, poderão fazer (programas de ação); no entanto, essas ações antecipadas talvez não ocorram porque outros atores têm programas diferentes — antiprogramas, do ponto de vista do primeiro ator. Assim, o artefato se torna linha de frente de uma controvérsia entre programas e antiprogramas” (Latour 2017:365). Lembramos à leitora que, para Deleuze e Guattari (2011–2012) um agenciamento sempre possui duas acepções possíveis: em relação ao estado-de-coisas, ele é agenciamento maquínico, uma ação; em relação ao regime de signos, ele é agenciamento coletivo de enunciação, uma significação.

¹⁸Para ilustrar isso, Latour dá o exemplo do homem-com-uma-arma: “Se eu definir você pelo que tem (um revólver) e pela série de associações à qual passa a pertencer quando usa o que tem (quando dispara o revólver), então você é modificado pelo revólver — em maior ou menor grau, dependendo do peso das outras associações que carrega. / Essa translação é totalmente simétrica. Você é diferente quando empunha uma arma; a arma é diferente quando empunhada por você. Você se torna outro sujeito porque segura a arma; a arma se torna outro objeto porque entrou numa relação com você. O revólver não é mais o revólver-no-armário, o revólver-na-gaveta ou o revólver-no-bolso e sim o revólver-em-sua-mão, apontado para alguém que grita apavorado. O que é verdadeiro quanto ao sujeito, o atirador, é verdadeiro quanto ao objeto, o revólver empunhado. [...] [N]em o sujeito nem o objeto (e seus objetivos) são fixos. [...] Não são nem as pessoas nem as armas que matam. A responsabilidade pela ação deve ser dividida entre os vários [actantes]” (Latour 2017:213–14).

¹⁹No exemplo do quebra-molas apresentado por Latour (2017), percebemos que o guarda-de-trânsito é substituído por um não humano que altera não só a composição espaço-temporal-atorial, mas que também modifica a própria natureza da ação: os carros não mais desaceleram para evitar uma multa, mas para prevenir danos nos seus automóveis.

- final. Desse modo, não há nem neutralidade da tecnologia, nem determinismo tecnológico, mas uma composição de agências humanas e não humanas que resultam em uma ação ou significado;
2. A ideia de que, por meio da mediação técnica, diferentes agentes, espaços e temporalidades são mobilizados; o que é decorrência direta de se pensar a tecnologia e outras instituições sociais como concretizações que fazem durar no tempo e no espaço, ações de origens distintas. Em outras palavras, a ideia de que “tecnologia é a sociedade feita durável” (Latour 1990).

Nesse sentido, levando em conta a particularidade do computador digital apresentada acima e partindo da definição de mediação técnica em Latour, podemos concluir apresentando uma definição própria de mediação algorítmica²⁰. Desse modo, consideramos que a agência dos algoritmos pode ser descrita como *mediação algorítmica*, um processo ou agenciamento que resulta em uma ação/significação e que é realizada pelo controle da informação digital em um sistema de computação. Essa operação é possibilitada pelo funcionamento de um computador digital, mas depende de um conjunto mais amplo para que seja definida. Em outras palavras, podemos dizer sem perda da especificidade, que uma mediação algorítmica é uma mediação técnica (Latour) que envolve um computador digital: seja um computador de mesa, notebook, smartphone, tablet, microcontrolador, etc. Nossa definição se ampara sobretudo na constatação da centralidade do computador digital entendido como indivíduo técnico (Simondon) onipresente da nossa vida contemporânea e em uma particularização de seu modo de mediação técnica (Latour). Desse modo, esperamos ter proposto uma definição de agência algorítmica que abarca sua especificidade operacional (controle da informação digital), bem como sua natureza eminentemente política (produção de significado e ação) e heterogênea (agregação de elementos humanos e não humanos) e que, por isso, seja capaz de descrever sócio-tecnicamente os casos de viés apresentados no capítulo anterior com maior precisão conceitual e responsabilidade política.

²⁰Cabe ressaltar que o termo “mediação algorítmica” já é utilizado por algumas pesquisadoras (Bezerra e Almeida 2020; Bruno 2024; Mittelstadt et al. 2016), entretanto, não encontramos uma definição sistemática do termo.

Capítulo 8

Mediações algorítmicas na aprendizagem de máquina

Um limite lógico da classificação do aprendizado de máquina, ou do reconhecimento de padrões, é a incapacidade de reconhecer uma *anomalia única* que aparece pela primeira vez, como uma nova metáfora na poesia, uma nova piada em uma conversa cotidiana ou um obstáculo incomum (um pedestre? um saco plástico?) em um cenário rodoviário. A *não detecção do novo* (algo que nunca foi “visto” por um modelo e, portanto, nunca classificado antes em uma categoria conhecida) é um problema particularmente perigoso para carros autônomos e que já causou fatalidades. A previsão de aprendizado de máquina, ou geração de padrões, apresenta falhas semelhantes na adivinhação [*guessing*] de tendências e comportamentos futuros. Como uma técnica de compressão de informações, o aprendizado de máquina automatiza a ditadura do passado, de taxonomias e padrões de comportamento passados, sobre o presente.

Se a compressão de informações produz a taxa máxima de lucro na IA corporativa, do ponto de vista social, ela produz discriminação e a perda da diversidade cultural

Nooscope manifested

Matteo Pasquinelli e Vladan Joler

Sabendo que a forma própria de ação dos algoritmos é a mediação da informação digital em um computador digital (capítulo 7), analisemos algumas das particularidades da mediação algorítmica para o caso do AdM, sobretudo tendo em vista os limites da problemática do viés exposta no capítulo 6.

8.1 Laço de realimentação conservador

Uma das características mais definidoras da aprendizagem de máquina é sua natureza inerentemente conservadora: ao treinar o modelo para encontrar relações em dados do passado, assume-se não só que esse conjunto de dados não evoluiu no tempo – pois do ponto de vista do treinamento do modelo não há um índice de sequencialidade –, mas mais importante ainda, assume-se que generalizar o conhecimento presente nesses dados para casos futuros é possível, isto é, que o futuro é uma repetição do passado.

Mas se a consequência mais imediata disso parece ser que o algoritmo nunca seria capaz de acertar 100% dos casos, sempre estando um passo atrás da realidade mais viva; os mecanismos de realimentação presentes no AdM nos permitem afirmar que sua consequência mais importante é a criação mesma dessa realidade repetida, o que Pasquinelli e Joler chamaram de “regeneração do passado”: “a aplicação de uma visão espaçotemporal homogênea [*homogeneous space-time view*] que restringe as possibilidades de um novo evento histórico” (Pasquinelli e Joler 2021:1275). É que existe sempre um “vão [*gap*]” ou um “conflito” entre o humano e a máquina: “A verdadeira limitação lógica e política da IA é a dificuldade da tecnologia em *reconhecer e prever um novo evento*”.

Essa homogeneização é bastante visível nos métodos patenteados pela Google. Uma forma comum em que a regeneração do antigo se expressa nas patentes analisadas é através de funcionamentos que penalizam comportamentos novos, que escapam à predição do modelo, enquanto que reafirmam comportamentos convergentes. Assim, uma usuária que decide ouvir uma música utilizando outro serviço de streaming do que ela utiliza mais frequentemente, não terá essa ação tão rapidamente executada do que se mantivesse a homogeneidade que foi capturada na imagem que dela fez o modelo (Mirelmann et al. 2022, analisada no capítulo 4). Em um sentido similar, um método automatizado de reordenamento de notificações – incluindo ênfases gráficas –, pode facilitar a vida de uma usuária que deseja sempre repetir seus hábitos, mas qualquer comportamento que escape à imagem criada pelo modelo é desincentivado, na medida em que escolher uma notificação que não é considerada relevante pelo algoritmo é forçadamente dificultado pelo método de Aradhye et al. (2017) (analisado no capítulo 5).

É claro que a usuária ainda pode tomar a rota mais demorada, mas o ponto que enfatizamos é uma tendência de apagamento da novidade, da diferença, da “restrição de possibilidades de um novo evento histórico” como dizem Pasquinelli e Joler. É que para além da história, que o AdM captura, há o devir. Pois o humano é isso: justo aquilo que escapa. Byung-Chul Han (2023), atualizando a oposição deleuziana entre história e devir¹, relaciona as tecnologias contemporâneas com a captura do que poderíamos chamar de devir:

Se [hoje] tudo tem que ser imediatamente visível, divergências são quase impossíveis. Da transparência surge uma pressão por conformidade que elimina o outro, o estranho, o desviante. Os *big data* tornam visíveis sobretudo os padrões comportamentais coletivos. O próprio dataísmo [dadoísmo] reforça o *crescer tornando-se iguais*. O *data-mining* [mineração de dados, treinamento] não é, em princípio, distinto da estatística [analisada por Nietzsche]. As correlações que ele expõe mostram o estatisticamente provável. São calculados os valores médios estatísticos. Assim, os *big data* não têm nenhum acesso àquilo que é único. Eles são completamente *cegos ao acontecimento*. Não é o estatisticamente provável, mas o improvável, o *singular*, o *acontecimento* que determinará a história, o *futuro* humano. Por isso, os *big data* também são *cegos ao futuro* (Han 2023:108–9).

¹Sobre a oposição entre história e devir, Deleuze comenta em uma entrevista concedida a Antonio Negri: “Nietzsche dizia que nada de importante se faz sem uma ‘densa nuvem não histórica’. Não é uma oposição entre o eterno e o histórico, nem entre a contemplação e a ação: Nietzsche fala do que se faz, do acontecimento mesmo ou do devir. O que a história capta do acontecimento é sua efetuação em estados de coisa, mas o acontecimento em seu devir escapa à história. A história não é a experimentação, ela é apenas o conjunto das condições quase negativas que possibilitam a experimentação de algo que escapa à história. Sem a história, a experimentação permaneceria indeterminada, incondicionada, mas experimentação não é histórica. [...] O devir não é história; a história designa somente o conjunto das condições, por mais recentes que sejam, das quais desvia-se a fim de ‘devir’, isto é, para criar algo novo” (Deleuze 2013:214–15).

Dessa maneira, embora existam usos interessantes do AdM mesmo nas artes, diríamos que como força historicizante, a aprendizagem de máquina é essencialmente *anticriativa*. Os usos que fazem do AdM algo de criativo tem a ver justamente com o desvio, e não propriamente com sua capacidade de gerar padrões previstos a partir da base de dados. A promptografia de Eldagsen (capítulo 4) não é uma simples aprendizagem de máquina, mas um uso desviante dela, que passa a ser componente de um sistema maior que é “dirigido” pelo artista. O mesmo pode ser dito das alucinações, um uso que desvia a aprendizagem de máquina de sua função projetada. Quando um modelo funciona da maneira correta, isto é, quando gera padrões derivados da base de dados, não há propriamente criação: “A ‘criatividade’ do aprendizado de máquina limita-se à detecção de estilos a partir dos dados de treinamento e, em seguida, à improvisação aleatória dentro desses estilos. Em outras palavras, o aprendizado de máquina pode explorar e improvisar somente dentro dos limites lógicos definidos pelos dados de treinamento” (Pasquinelli e Joler 2021:1275). Como constata Simondon (2020b), a previsão não é uma verdadeira antecipação, pois para tal, falta-lhe a *invenção*.

Como num sistema de controle cibernético, o AdM age de maneira a tender à homeostase, à estabilização. Pois o treinamento nada mais é que um processo em que o modelo utiliza a divergência entre sua resposta e o desenvolvimento efetivo de um processo (o rótulo) para adaptar a máquina visando atualizar o seu sinal em um sentido de convergência: isto é, o treinamento é o processo contínuo de adaptação do modelo para minimização do erro. A usuária, como entrada e saída da máquina, em um regime de servidão maquínica, funciona como uma realimentação negativa que visa a estabilização do modelo. Enquanto que este é progressivamente estabilizado, a atividade da usuária é progressivamente automatizada e o funcionamento total do sistema tende ao conservadorismo.

8.2 Estereotipia

Esse conservadorismo da aprendizagem de máquina leva, inevitavelmente, para funcionamentos repetitivos e resultados estereotipados. A atualização de parâmetros do modelo que acontece no treinamento, por meio de uma sequência de cálculos que ocorrem por várias “épocas” ou “iterações” visa sempre diminuir a distância entre os resultados obtido e visado, via realimentação. O processo acaba quando o erro é suficientemente pequeno, ou seja, diz-se que o modelo “generaliza bem”. Comentando a diferença entre esse processo, o adestramento da máquina, e o aprendizado propriamente dito, do humano, Simondon escreve:

Essa máquina, conseqüentemente, se adapta. Mas a adaptação é possível segundo dois processos opostos: aqueles que acabamos de evocar é o adestramento, que chega a uma conduta cada vez mais estereotipada e a uma ligação cada vez mais estreita com um meio determinado. A segunda forma de adaptação é a aprendizagem, que, ao contrário, aumenta a disponibilidade do ser relativamente aos diferentes meios nos quais ele se encontra, desenvolvendo a riqueza do sistema de símbolos e de dinamismos que integram a experiência passada segundo um determinismo divergente (Simondon 2020b:528–29).

Essa “causalidade reativa” da máquina automática em geral e do AdM em particular não é um verdadeiro aprendizado, pois não há nele uma refundição de estrutura, a estrutura da máquina não muda, de forma que, em realidade, não há tensões ou um problema efetivamente: os erros são também positivos, na medida em que servem para a construção do resultado da máquina (Gonçalves 2022).

Diferentemente, no aprendizado, o sinal “(...) é aquele de uma dissimetria entre duas finalidades, uma realizada sob a forma de estrutura, outra imanente a um conjunto de informações ainda enigmáticas e, no entanto, valorizadas” (Simondon 2020b:530). A aprendizagem de máquina não é um verdadeiro aprendizado, mas sim um adestramento.

Isso é evidenciado pela própria formulação do problema que o AdM visa resolver. Em geral, deseja-se chegar a uma configuração ótima que generalize os resultados de uma base de dados, mas sem cair em duas situações limites: o subaprendizado, ou o sobajuste [*underfitting*] dos pesos em uma RNA, entendido como situação em que ainda há margem para melhoria do modelo, sem redundar em uma cópia do conhecimento da base de dados; e o superaprendizado, ou sobreajuste [*overfitting*] dos pesos, que é entendido como uma situação em que o modelo acerta todos os pontos da base de dados em detrimento de uma capacidade mais generalística de interpolação e extrapolação dos dados (figura 8.1). Para atingir essa otimalidade, o treinamento do modelo se torna um processo iterativo e automático de minimização de uma métrica de erro, o que, frequentemente, se reflete em uma maximização da acurácia (taxa de acertos).

Esse funcionamento convergente, funcionando de maneira análoga à uma amplificação moduladora (Simondon 2020a), gera estabilização e estereotipia. Mas um funcionamento totalmente divergente, análoga a uma amplificação transdutora (Simondon 2020a), geraria instabilidade e, portanto, não geraria significação alguma, não informaria efetivamente. Tomando por exemplo um gerador de imagens, se uma estabilização total significa uma cópia da base de dados, a geração perfeita de uma mesma imagem ou de cópias perfeitas da base de dados, um funcionamento oposto geraria apenas ruído, pixels totalmente aleatórios. Por operar por interpolação e extrapolação, o AdM está muito mais perto da primeira situação que da segunda. Mas o ponto “ótimo” de operação, ainda não é uma verdadeira invenção, uma otimalidade no sentido forte. É que o ótimo não é aquilo que maximiza ou minimiza um parâmetro – operações caras à racionalidade neoliberal que opera por um dispositivo de maximização (Dardot e Laval 2016) –, mas aquilo que *resolve um problema*².

O automatismo – que, no caso do AdM, se expressa em limites sobre os tipos de dados que podem ser recebidos, as tarefas que podem ser resolvidas e o conhecimento que pode ser gerado e armazenado (Gonçalves 2022) – é, segundo Simondon, o responsável pela estereotipia. Comentando uma automatização dos meios de comunicação, o que hoje já acontece em alguma medida (e.g. emoticons em redes sociais), o autor imagina um sistema em que para cada sentido a ser comunicado seria criada uma fórmula-estereótipo representante: a comunicação só necessitaria do nome da remetente, endereço da destinatária e número da fórmula:

Aqui, o indivíduo atípico é paralisado em sua escolha, pois nenhuma fórmula prevista res-

²Comentando as formas geométricas como a linha reta, a esfera ou mesmo os padrões hexagonais de uma colmeia de abelhas, Simondon diz: “[...] esses mínimos tornam-se ótimos, não como sistemas cuja energia é degradada, mas quando são tomados como soluções para um problema, ou seja, quando, em vez de serem a fórmula do sistema geral, são meios, subconjuntos que parecem desempenhar o papel de mediadores e intermediários, entre uma configuração maior e uma matéria elementar; A forma esférica é uma boa forma quando é uma estrutura intermediária, como o tanque, o balão, o bulbo de vidro fino que envolve um filamento e o retira da ação da atmosfera; o batiscafo, por exemplo, é o envelope de um universo dentro de outro universo; é funcionalmente o sistema de compatibilidade entre dois ambientes incluídos um no outro. O caso das figuras geométricas como imagens significativas não significa que o significado esteja ligado a estados de equilíbrio estável e às formas mais prováveis de acordo com as leis do acaso, porque há casos em que as formas geométricas são, ao contrário, soluções excepcionais para o problema de máximos e mínimos colocados pela relação entre duas ordens de magnitude da realidade” (Simondon 2008:86).

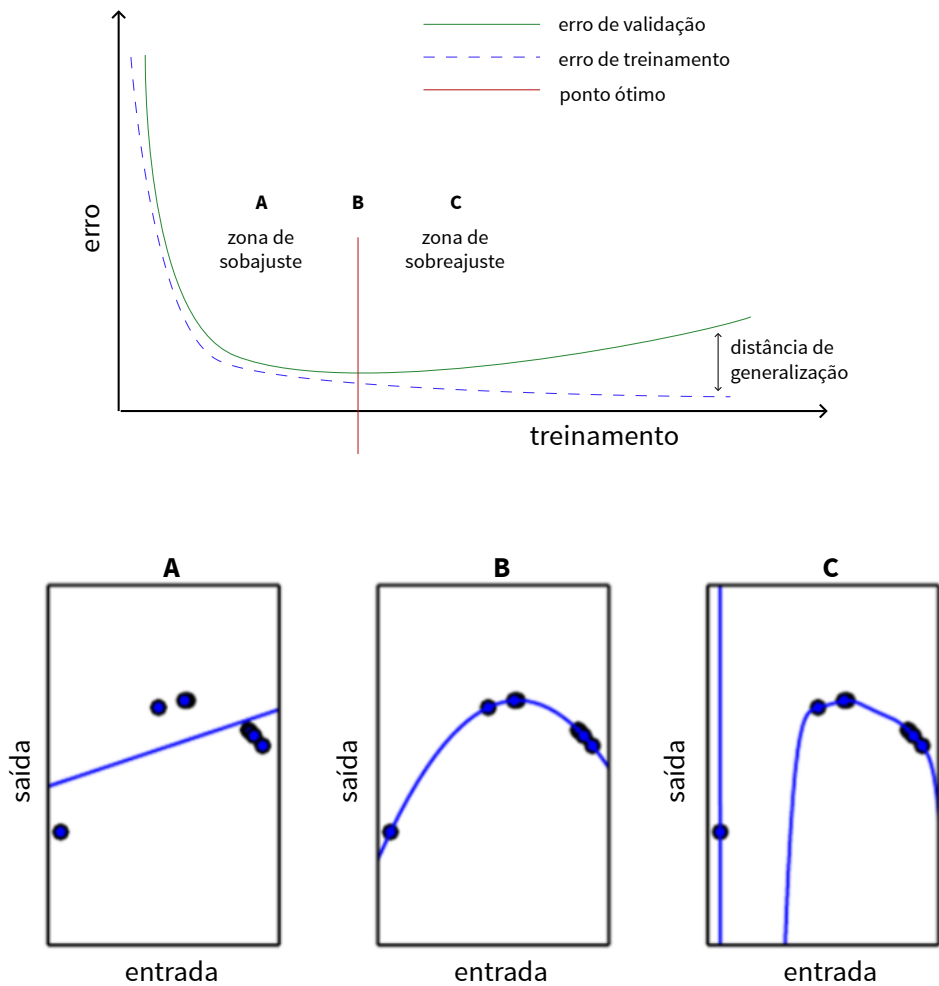


Figura 8.1: Exemplo fictício de gráfico de treinamento (acima), a ordenada mede o erro, enquanto que a abcissa representa o tempo de treinamento (número de épocas). São mostradas as zonas de sobajuste (A), ponto ótimo (B) e zona de sobreajuste (C). Quanto maior a distância entre o erro de treinamento e o de validação, isto é, de uma inferência realizada para testar o modelo, menor a capacidade de generalização do modelo. Abaixo, um exemplo fictício de resultado de uma função para os três casos (A, B e C): o sobajuste resulta em um resultado mais estereotipado, mas um sobreajuste projeta interpolações e extrapolações muito pouco estáveis, o que faz com que o modelo não acerte exemplos que não foram utilizados no treinamento. Fonte: elaboração própria, baseada em Goodfellow et al. (2016:111, 113, 118, figuras 5.2, 5.3 e 5.5).

ponde exatamente ao que ele teria querido exprimir. O atípico que causa um grande dispêndio de informação à comunidade é um ser deficitário a partir do momento em que a informação é transmitida indiretamente de indivíduo a indivíduo pelo intermédio de um dispositivo que utiliza o automatismo: uma voz muito grave, muito aguda ou rica em harmônicos, é mais deformada pela transmissão telefônica ou pelo registro que uma voz cujas frequências médias se situam nas bandas telefônicas e que não põem na aparelhagem nenhum problema difícil relativo à transmodulação (Simondon 2020b:533).

Assim, como ilustra Simondon, o automatismo – frequentemente ligado com uma economia de recursos e aumento de eficiência –, também causa a estereotipia – homogeneização e penalização do diferente. Isso é ilustrado, por exemplo, na patente descrita por Braun e Patel (2022) (analisada no capítulo 4) que objetiva criar uma imagem da troca de mensagens de determinada usuária a fim de lhe prover “mensagens prontas [*canned responses*]”. O mesmo mecanismo pode ser observado em Carbune e Feuz (2022) (analisado no capítulo 3): a criação de uma imagem dos eventos em desenvolvimento por meio da interceptação de fluxos informacionais cria um sistema de ressonância que reafirma a comunicação sobre os eventos mais comentados, o que tende ao apagamento da diversidade ao se ignorar casos considerados como *outliers* (e.g. um fluxo informacional em área rural que é menor do que outra na área urbana). Aqui o algoritmo media a informação digital, amplificando-a em um caso e suprimindo-a, no outro. Desse modo, a estereotipia está fortemente relacionada com a reprodução e criação de assimetrias sociais.

8.3 Racismo e sexismo algorítmico

A realimentação conservadora e a estereotipia se relacionam diretamente com as discriminações sociais. Aqui voltamos aos casos expostos no capítulo 6. Acreditamos que os casos ali descritos – da hipervisibilização e invisibilização de pessoas negras em sistemas digitais; e o sexismo operado notoriamente na classificação binária de gênero por sistemas de visão computacional –, se não são bem analisados pela categoria de viés, podem ser melhor compreendidos do ponto de vista da mediação algorítmica. Ou seja, a obrigatoriedade de se classificar um rosto como masculino ou feminino é um efeito do modo como a informação digital (e particularmente os pixels, entendidos como sinais de informação que digitalizam raios luminosos captados em sensores de uma câmera) é controlada no interior de um sistema de computação digital. Num computador, matrizes de pixels serão analisadas por um algoritmo que vai agir classificando tal ou qual representação digital do rosto de um sujeito (e, por extensão, o próprio sujeito) como “masculino” ou “feminino”. O mesmo ocorre para a hipervisibilização e invisibilização de pessoas negras, diferentes fluxos informacionais são suprimidos ou amplificados criando uma situação que hierarquiza socialmente as pessoas negras em posições de inferioridade em relação às pessoas brancas: sua linguagem pode ser suprimida por um sistema de medição de toxicidade, enquanto que sua visualidade pode ser exacerbada e relacionada a temas específicos (como a pornografia e a violência) em mecanismos de busca e redes sociais. A linha do tempo de *Danos e discriminação algorítmica* (Silva 2023), é um mapeamento extenso de casos de discriminação algorítmica publicizados pela mídia; é, portanto, um instrumento muito importante para a pesquisa sobre a agência algorítmica e para o enfrentamento político do uso nocivo das novas tecnologias. Enfatizamos que embora as patentes não sejam explícitas nas consequências que elas promovem em relação à discriminações sociais, consideramos importante

expor brevemente a capacidade que o AdM possui em exacerbar tais assimetrias, na medida em que uma decisão como o uso de localização como dado de entrada, por exemplo, tem implicações para as assimetrias de raça, dado que ela pode ser considerada uma *proxy* [representante] para raça, ou seja, localização e raça são variáveis correlacionadas (cf. Ruback et al. 2022:23:19–20, especialmente tabela 2)³.

8.3.1 Racismo algorítmico

A assimetria social mais estudada do ponto de vista das tecnologias digitais é o racismo algorítmico. Com base em uma fala feita por Tarcízio Silva em 2021, Lippold e Faustino (2022) sistematizam cinco pilares do racismo algorítmico, que reproduzimos abaixo:

1. O primeiro é o que ele chama de *Looping de feedback*: **o modo como sistemas de inteligência artificial promove vieses de discriminação racial já existentes na sociedade**. Cita como exemplos os sistemas de reconhecimento de objeto (aprendizado de máquinas) e imagens que tendem a incorporar os vieses raciais e fazer associações racializadas.
2. O segundo pilar é o que ele chama de *humanidade diferencial*: **o modo como o racismo acaba promovendo o grupo hegemônico em detrimento de minorias, consolidando uma espécie de distribuição racial do sistema tecnológico**.
3. O terceiro pilar, segundo argumenta, é o paradoxo entre invisibilidade e hipervisibilidade. Baseado nos estudos de Joy Buolamwini sobre a *disparidade interseccional*, ele argumenta **o racismo pode se manifestar, de um lado, no não reconhecimento correto do traço de mulheres negras nos app de reconhecimento lúdico ou funcionais, e do outro lado, a hipervisibilidade negra nas formas de dominação e controle**. Como exemplo, Silva lembra que 90,5% das pessoas presas por reconhecimento facial no Brasil são pessoas negras.
4. O quarto pilar é a *colonialidade global no negócio da tecnologia*. Segundo Tarcízio, **grandes empresas de big tech colonizam infra-estruturas tecnológicas em alguns países menos conectados, de forma a restringir o acesso desses povos ao seu monopólio**. Um exemplo famoso é a oferta de internet gratuita e de baixa qualidade pela google e o facebook para países com baixíssima conexão como Gâmbia, Sri Lanka, no entanto, o preço cobrado é que as pessoas só podem acessar os produtos dessas mesmas empresas ao invés de terem o acesso ilimitado à internet. Essa proposta chegou a ser apresentada pelo Facebook ao Brasil, mas foi rejeitada pela então Presidente Dilma.
5. O quinto pilar é o que Silva chama de *colonialidade de campo*. O pesquisador observa **como as disciplinas do campo da informação tendem a negligenciar a presença do racismo em seus objetos de estudo e formação de profissionais, professores e novos pesquisadores**.

(Lippold e Faustino 2022:70–71, grifo nosso)

³O fenômeno descrito por Benjamin (2019) em relação ao racismo concretizado na automatização da avaliação de risco na saúde é desse tipo.

Conforme entendemos a apresentação, o racismo algorítmico abrange simultaneamente a agência algorítmica propriamente dita (1 e 3) e aspectos mais amplos que relacionam o racismo e o colonialismo digital (2, 4 e 5). Por esse motivo, a caracterização feita por Silva é de grande importância para o estudo de Faustino e Lippold. Entretanto, o uso do termo é afastado pelos autores, o que é feito na forma de uma pergunta retórica: “Se os códigos são, mesmo em sua tendente automação, padrões socialmente determinados, o termo ‘racismo algorítmico’ não tenderia a escamotear a autoria do racismo, transferindo-a para os códigos enquanto oculta os seus programadores, esses sim humanos formados e informados por dadas relações sociais de poder?” (Lippold e Faustino 2022:71).

Partindo de uma posição teórica que enquadra as tecnologias no interior de um tradicional privilégio da luta de classes⁴, propõem utilizar alternativamente o termo “racialização codificada” ou “racialização digital”, por ser “mais abrangente para dar conta da explicitação do contexto material de desenho dos algoritmos de forma a evidenciar a seletividade racial dos cargos técnicos em empresas de programação, a distribuição social desigual de prestígio entre produtores de conteúdo digitais na internet [...] e codificação naturalizada dos discursos e estética racistas nas mídias sociais e bancos de imagem digitais” (Lippold e Faustino 2022:71).

Entretanto, não acreditamos que haja aí uma verdadeira oposição. Nada mais material do que a ação operada pela própria infraestrutura técnica. Se levamos em conta a agência técnica – e, portanto, a algorítmica – como sustenta os estudos, entre outros, de Latour (e.g. 2017); tanto quanto os aspectos macroestruturais, podemos observar que a racialização codificada perpassa diferentes processos do colonialismo digital (tal qual exposto no capítulo 3), conservando o termo racismo algorítmico para localizar mais propriamente mediações algorítmicas racistas, ou seja, processos em que o foco da racialização está na tecnologia em si, tais quais nos exemplos apresentados no capítulo 6. Acreditamos que isso é inclusive mais próximo da definição dada por Tarcízio Silva do que a caracterização que fazem Lippold e Faustino (2022) sobre a racialização codificada, por não implicar uma redução da ação dos algoritmos aos humanos que supostamente seriam os únicos agentes envolvidos em relações sociais. Em seu livro homônimo, ele define racismo algorítmico como:

o modo pelo qual a disposição de tecnologias e imaginários sociotécnicos em um mundo moldado pela supremacia branca realiza a ordenação algorítmica racializada de classificação social, recursos e violência em detrimento de grupos minorizados. Tal ordenação pode ser vista como uma camada adicional do racismo estrutural, que, além do mais, molda o futuro e os horizontes de relações de poder, adicionando mais opacidade sobre a exploração e a opressão global que já ocorriam desde o projeto colonial do século XVI (Silva 2022:81).

Como “ordenação algorítmica racializada”, o racismo algorítmico diz respeito justamente às mediações algorítmicas que redundam em racismo. Acreditamos que o grande valor dessa categoria está

⁴Para os autores “a verdadeira e principal contradição não se dá entre humanos e máquinas, mas sim, entre humanos (detentores) e humanos (não-detentores) dos meios e possibilidade de produção da vida. O que não significa que a luta de classes não tenha no desenho e emprego das máquinas uma dimensão fundamental dessa luta (de classes)” (Lippold e Faustino 2022:69). Essa questão é complexa, e poderia ser analisada de diferentes aspectos: passando desde Flusser (2018) e sua argumentação sobre a passagem da lógica da propriedade para a da programação, até Simondon (2020c) e sua oposição mais entre uma relação autocrática ou não-autocrática com as máquinas. Para nós, em um esforço de pensar o híbrido (Latour 2019), consideramos que há mais uma dicotomia entre humanos-e-máquinas dominadoras e humanos-e-máquinas dominadas. Um corolário disso é que a agência técnica não se reduz aos humanos em posições sociais distintas, há também uma agência das técnicas.

em incluir a agência técnica (tanto quanto seu entorno discursivo) na agência social mais ampla que historicamente define o racismo. Ao apresentar uma definição sociotécnica que relaciona o chamado “viés de raça” com as formas históricas de racismo, Silva não só aumenta a precisão da caracterização desse fenômeno, mas também apresenta uma nomenclatura mais politicamente responsável ao conectar o funcionamento técnico, que é muitas vezes visto como neutro, com a história, o debate público e a política do racismo e do anti-racismo.

8.3.2 Sexismo algorítmico

Inspirado em sua análise, utilizando a lente da mediação algorítmica, e amparado na importante catalogação de casos midiáticos de discriminação algorítmicas feita pelo mesmo pesquisador (Silva 2023), podemos ainda constatar uma série de modos distintos através dos quais os algoritmos agem produzindo e reproduzindo outras formas de discriminação social. Em relação às assimetrias de gênero e sexualidade⁵, particularmente, observamos que são pelo menos cinco os modos pelos quais uma mediação algorítmica pode ser sexista:

1. Fixando categorias binárias de gênero;
2. Padronizando seus resultados como masculinos;
3. Pressupondo a cisgeneridade ou heterossexualidade em seus resultados;
4. Reproduzindo e exacerbando estereótipos;
5. Homogeneizando a linguagem e apagando formas divergentes de discurso.

O primeiro, é evidenciado no verbete de Gebru (2020, capítulo 6) que questiona as tecnologias de reconhecimento automático de gênero e sua estrutura rígida que somente permite uma classificação binária. Já a padronização de resultados para o masculino (2) são analisados tanto no estudo de Prates et al. (2020) sobre a tradução de palavras de um idioma sem marcação de gênero para um com marcação de gênero pelo Google Translate, quanto em nossa breve investigação sobre o complemento de frases generificadas pelo BERT/Google (Gonçalves e Ferreira 2022a). Outro caso é o de recomendação de artistas predominantemente masculinos em plataformas como o Spotify e o Deezer⁶. A pressuposição da cisgeneridade e da heterossexualidade (3) é evidenciado não só no trabalho de Buolamwini e Gebru (2018; cf. Gebru 2020) sobre AGR, como no resultado de geradores de imagem (conferir abaixo). A exacerbção de estereótipos (4) é visível em geradores de imagem como no estudo sobre o Dall-e mini feito por Cheong et al. (2024) que mostrou uma tendência do modelo em produzir imagens estereotipadas em relação à raça e gênero, no processamento de linguagem natural (Gonçalves e Ferreira 2022a), como já exposto, e na personalização estereotipada de anúncios⁷. Finalmente, a homogeneização e o apagamento de formas divergentes de discurso (5) ocorrem não só no Google Perspective, analisado por Silva (2022, capítulo 6), como também na classificação de discurso LGBTQ+ como “tóxico”⁸.

⁵Uma primeira versão desse argumento foi apresentado na XV Jornada Latino-Americana de Estudos Sociais da Ciência e Tecnologia (Gonçalves 2024b).

⁶<https://nucleo.jor.br/reportagem/2023-06-08-spotify-e-deezer-querem-que-voce-ouca-mais-musicas-de-homens/> (acesso: 21/02/2024).

⁷<https://www.intercept.com.br/2019/04/05/facebook-como-maquina-de-preconceito/> (acesso: 21/02/2024).

⁸<https://internetlab.org.br/pt/noticias/drag-queens-e-inteligencia-artificial-computadores-devem-decidir-o-que-e-toxico-na-internet/> (acesso: 21/02/2024).

Como apenas um experimento inicial, utilizamos o gerador de imagens baseado em texto Craiyon⁹ e realizamos uma única execução para as três seguintes palavras: “*doctor* [médica/o]”, “*nurse* [enfermeira/o]” e “*couple* [casal]”. Corroborando com os estudos apontados acima, os resultados das primeiras duas figuras (8.2 e 8.3) mostram principalmente a estereotipia: mais imagens aparentemente masculinas para “*doctor*”, mais imagens aparentemente femininas para “*nurse*”, resultado que vai ao encontro de Gonçalves e Ferreira (2022a); já a figura 8.4, é convergente com a ideia de pressuposição da heterossexualidade, na medida em que a maioria das figuras geradas mostram representações gráficas de casais em que os gêneros aparentes das pessoas é masculino e feminino. É claro que não pretendemos com isso afirmar tendências específicas no modelo utilizado, o que só seria possível a partir de estudos sistemáticos como o de Cheong et al. (2024), mas acreditamos que esses casos são interessantes por ilustrarem, ou melhor, operarem, o sexismo realizado através de algoritmos.

Desse modo, inspirados na noção de “racismo algorítmico”, seria possível denotar por “sexismo algorítmico” a produção e reprodução de assimetrias de gênero e de estereótipos que privilegiam o masculino, a imposição binária de gênero e toda ação social mediada por algoritmos que ressoam com formas históricas que privilegiam a cisgeneridade, heterossexualidade e o masculino em detrimento de outras identidades de gênero. Acreditamos que essa caracterização supera o debate do viés, pois amplia a gama de fenômenos sobre os quais ela se aplica, enquanto que, simultaneamente, explicita que os algoritmos agem participando do sexismo entendido como processo social.

⁹<https://www.craiyon.com/> (acesso: 11/12/2024).

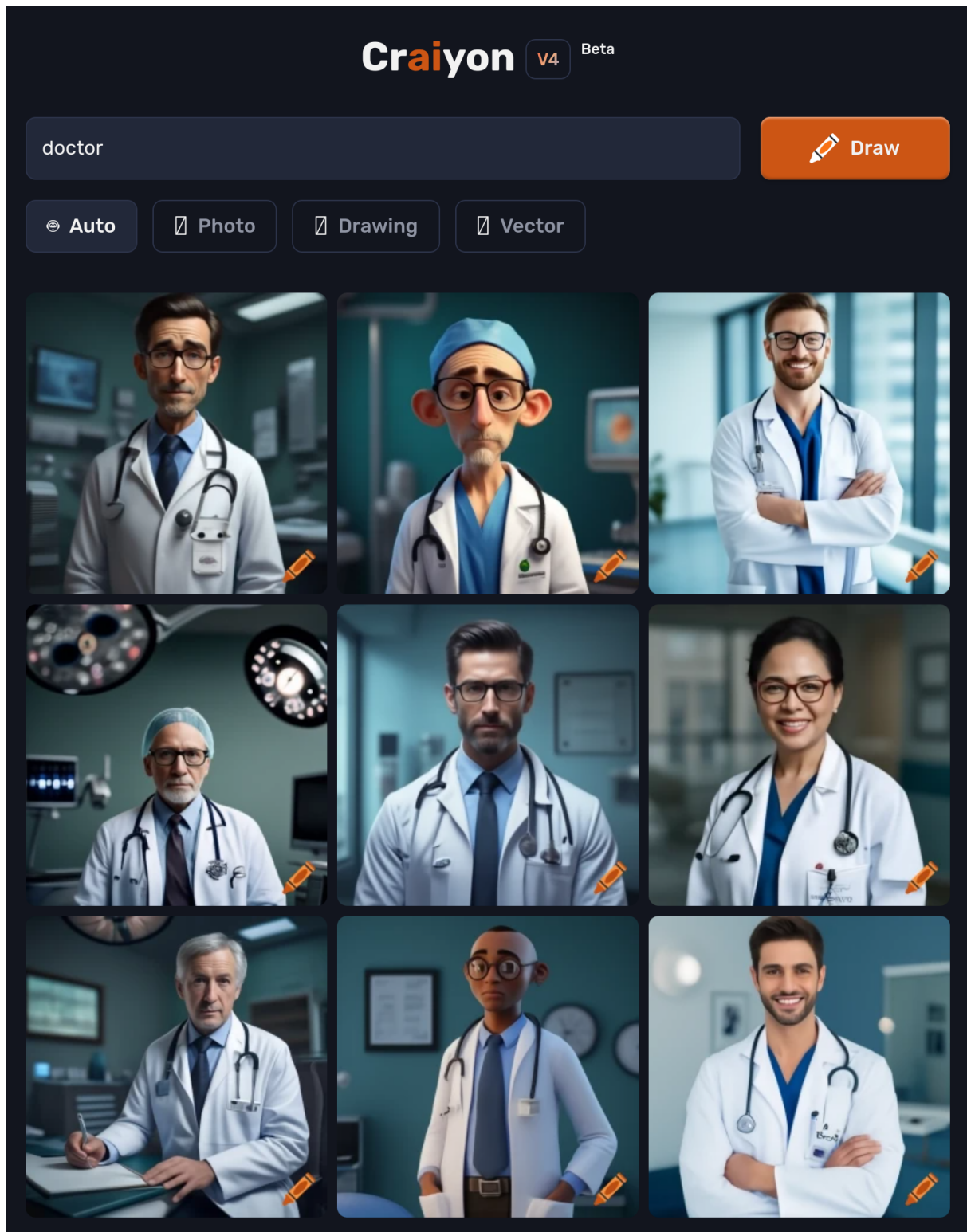


Figura 8.2: Execução do modelo gerativo Craiyon v4 para o prompt “doctor”.

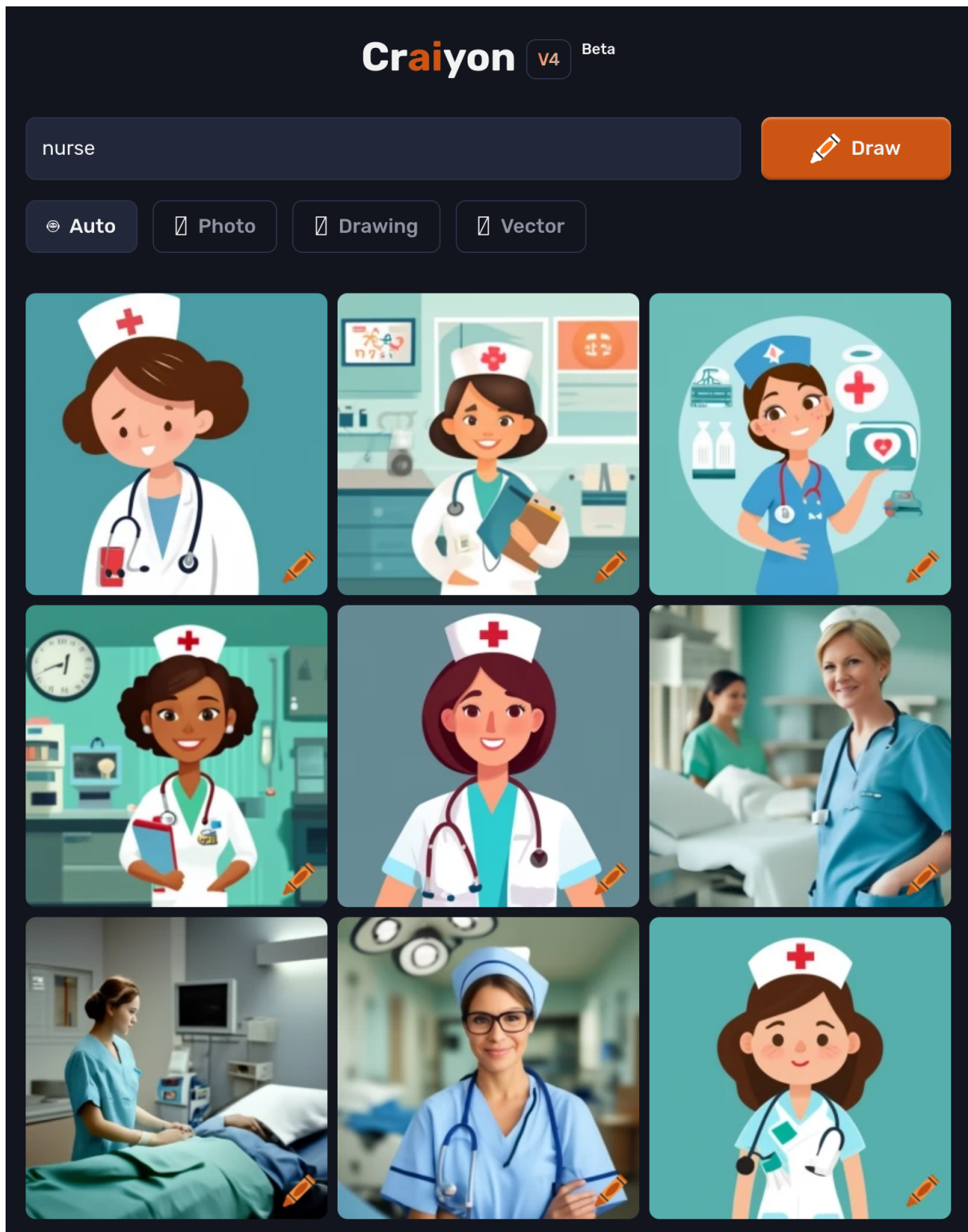


Figura 8.3: Execução do modelo gerativo Craiyon v4 para o prompt “nurse”.

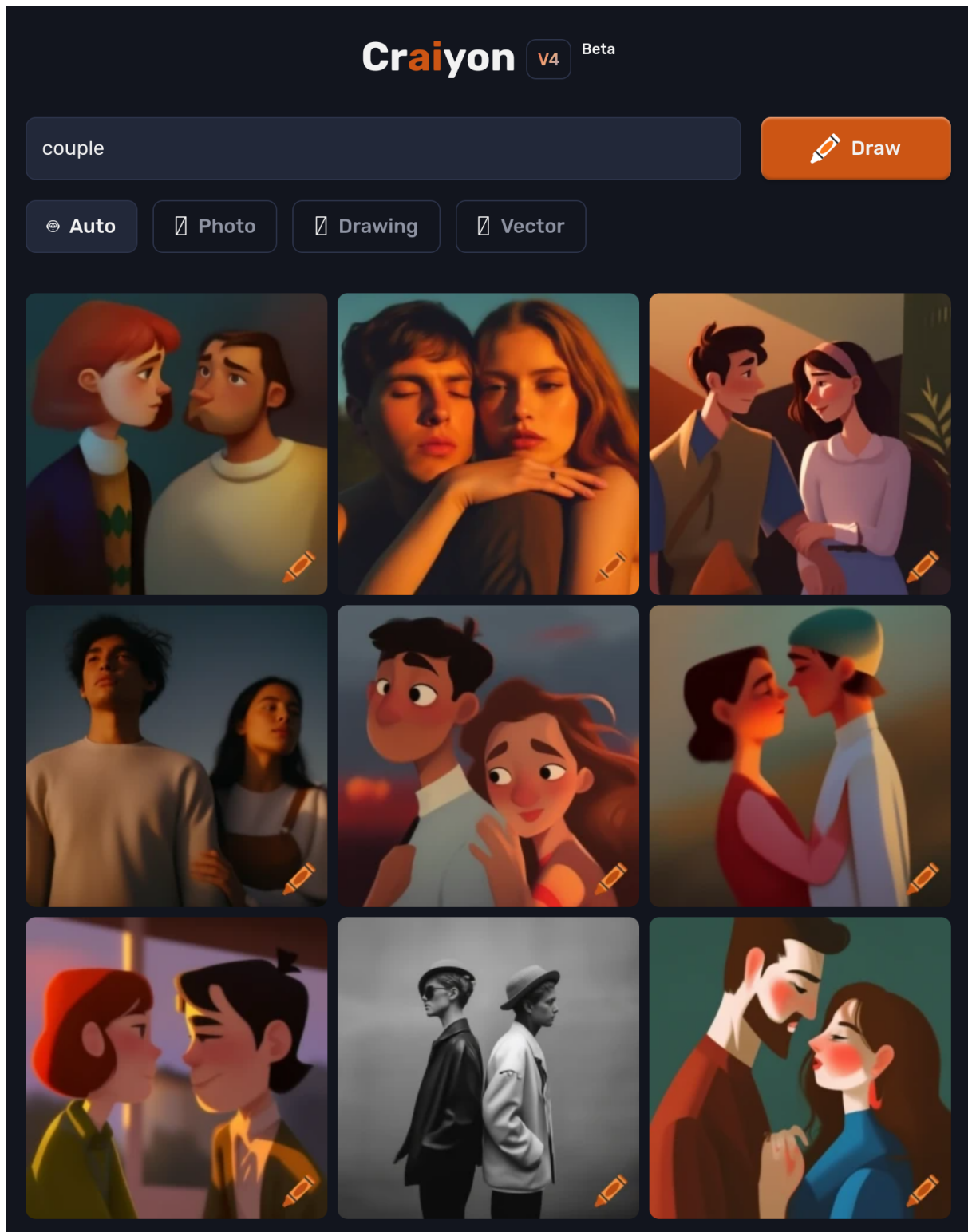


Figura 8.4: Execução do modelo gerativo Craiyon v4 para o prompt “couple”.

Considerações finais

Assistimos ao desenvolvimento cada vez mais casado das tecnociências e do capitalismo. A passagem do século XX para o XXI, descrito diversificadamente como “mundo de ciborgues” (Haraway), “sociedades de controle” (Deleuze) e “virada cibernética” (Garcia dos Santos), fez surgir não só novas tecnologias, como também novas formas de acumulação capitalista amparadas em noções tecnocientíficas como as de controle, código e informação. Hoje, a proliferação de dispositivos computacionais e o aumento significativo de situações mediadas por algoritmos evidenciam a posição central das tecnologias digitais e dos algoritmos. Enfatizando novos modos de dominação, a realidade contemporânea é heterogeneamente descrita como “colonialismo digital” (Kwet, Lippold e Faustino), “colonialismo de dados” (Ricaurte, Couldry e Mejias), “capitalismo de vigilância” (Zuboff), entre outros. Comum a todas essas descrições está o papel central da inteligência artificial, das técnicas de aprendizado de máquina e de uma multiplicidade de processos relativos ao big data.

Diante disso, duas questões se impõem: a primeira diz respeito à posição que a aprendizagem de máquina ocupa no nosso coletivo: posicionada na intersecção entre tecnociência e capitalismo, qual a operação sociotécnica que ela desempenha? A segunda, aparentemente mais teórica, diz respeito à possibilidade mais ampla de considerar que é possível falar de uma agência algorítmica: como descrever a ação dos algoritmos sem reduzi-la a uma explicação puramente discursiva ou a uma explicação esquiva que a reduz completamente ou para o funcionamento de seus componentes materiais elementares ou para agentes humanos por detrás? Esta dissertação foi uma tentativa de dar uma resposta inicial para ambas as questões partindo de um recorte estreito: um conjunto de técnicas de aprendizagem de máquina utilizadas para a modulação de comportamento patenteadas pela Google.

Com Zuboff (2018), entendemos que o pioneirismo e a atual posição de centralidade da empresa Google em relação ao uso do big data para a acumulação de capital, fazem dela um objeto privilegiado para entender a convergência atual entre tecnociência e capitalismo. Sua influência geopolítica atrelada a instituições do governo estadunidense, sua superioridade legal e mercadológica, a gigantesca capacidade infraestrutural e a escala e escopo de seus serviços que se tornaram pontos de passagem obrigatória, a aproximam hoje de uma posição de governo global. Em relação ao aprendizado de máquina, entendemos que ele é utilizado por ela simultaneamente para fins econômicos – a geração de lucro a partir de dados e da atividade de usuárias – e políticos – a concentração de poder por meio de uma rede inigualável que vai da captura de fluxos informacionais até a modulação de atividade de usuárias que utilizam seus serviços.

Em relação ao seu papel na acumulação capitalista, não acreditamos que as usuárias, mediadas por algoritmos de AdM, possam ser definitivamente enquadradas como “trabalhadoras” como por vezes sugerem análises como a de Fuchs (2012), utilizando a noção de prosumidor, pois, não havendo uma relação de assalariamento, é difícil definir tal relação como trabalho com alguma precisão conceitual,

uma exploração econômica propriamente dita. Por outro lado, também não reduzimos o aspecto econômico a um enfoque que veja nos dados recursos, isto é, matéria prima – como na sua aproximação com o petróleo, ou em perspectivas unicamente preocupadas com o aspecto expropriativo ou extrativista do AdM (e.g. [Couldry e Mejias 2019](#); e, embora parcialmente, [Zuboff 2018](#)) – dado que o foco da valorização por meio da mediação algorítmica recai justamente na atividade das usuárias. Desse modo, estamos mais próximo da perspectiva proposta por Deleuze e Guattari ([2011–2012](#)), em considerarmos que o papel da usuária que interage com sistemas algorítmicos se assemelha mais com um mecanismo do sistema, com um componente em uma máquina: servidão maquínica de onde se extrai um mais-valor maquínico¹⁰ – nesse sentido, a usuária talvez esteja mais próxima da categoria marxiana de “meio de produção”, pois mobiliza informações e, por meio de sua atividade, valoriza o valor, produz dados e parâmetros numéricos, do que de uma “trabalhadora” ou “matéria-prima”. Esperamos que nosso trabalho possa ser ponto de partida para investigações mais aprofundadas sobre esse aspecto.

Já em relação ao exercício de poder, partimos da questão da modulação de atividades, também chamada de modulação de comportamento: o modo como a mediação algorítmica informa o humano, participando ativamente da constituição das ações de uma usuária que interage com um sistema de computação. Nossa análise concluiu que esse processo depende de outros mecanismos para que a modulação se realize, de modo que escolhemos caracterizar o funcionamento sociotécnico do AdM com foco em três momentos que, embora não sejam realmente distintos, pois são codependentes e possuem um funcionamento emaranhado, podem ser analiticamente descritos: a captura, a imaginação e a modulação.

A captura é o processo responsável por interceptar fluxos informacionais, o que é feito por meio de uma rede de sensores presentes particularmente nos smartphones, mas que se expande cada vez mais para outros dispositivos como na ideia de IoT, e também pela interface de usuária, que é responsável por gerar e captar dados das interações de usuárias em tecnologias digitais (como em softwares, sites e aplicativos). Esse processo de captura digitaliza a informação, produz dados que são armazenados em bases de dados proprietárias e configura os parâmetros do aprendizado de máquina, que passa a ser capaz de extrair relações dos dados criando representações, um excedente informacional, uma imagem.

A imaginação é o processo centrado na constituição dessa imagem. Aqui, o treinamento do modelo de AdM, isto é, o processamento iterativo, em tempo real ou não, dos dados gerados na interação da usuária com a máquina, configuram os parâmetros numéricos que definem a agência propriamente dita do AdM. Os parâmetros numéricos são a base material da imagem, endogenamente chamada de espaço latente, que é um repositório de virtualidades para o funcionamento do modelo. Desse modo, se a captura se conecta com a imaginação provendo informação digital que é processada gerando uma imagem, a imaginação se conecta com a modulação na medida em que esta é a atualização da imagem em processos de inferência e incidência informacional: classificação de padrões, predição de séries temporais, codificação de palavras ou geração de imagens, por exemplo.

A modulação é o modo como a ação do AdM incide retroativamente à usuária que interage com ele. Como resultado de inferências e incidências informacionais, que são a atualização da imagem, o modelo

¹⁰“Deleuze e Guattari, sensíveis ao crescimento extraordinário do capital constante (as maquinarias), introduzem o conceito de mais-valia maquínica e de tempo maquínico. Esses tempos são os tempos de servidão onde não distinguimos mais sujeito e objeto, humano e não humano, natural e artificial. Essas temporalidades maquínicas constituem os fatores essenciais da produção capitalista. À diferença do tempo e da mais-valia humanos, os tempos e a mais-valia maquínicos têm a propriedade de não serem nem quantificáveis, nem determináveis” ([Lazzarato 2010:171](#)).

produz números que operam classificações, predições, codificações e gerações. Esses resultados alteram diretamente a interface de usuária – na forma de modificações na disposição dos elementos gráficos, do aparecimento de botões, da geração de textos e figuras, em restrições de acesso e reordenamentos –, de modo que a usuária é indiretamente afetada. Mostramos como essa forma de poder de um centro, o modelo e sua imagem, sobre uma periferia, as usuárias, acontece na forma de um controle direto sobre o meio no qual a usuária se localiza, e por comunicação instantânea que traduz sua atividade, de modo que podemos chamar essa ação de uma modulação da atividade da usuária, seguindo a descrição de Deleuze (2013).

Mas o funcionamento do AdM não acontece de maneira isolada, afetando apenas individualmente a usuária diretamente conectada a ela. Partindo da crítica de uma racionalidade que visa se amparar em uma suposta objetividade do cálculo e do automatismo para construir uma sociedade o máximo possível automatizada, mostramos como “o aprendizado de máquina sobre a ação social” já é “ação social do aprendizado de máquina”. Esse debate está geralmente presente na constatação de que os algoritmos, e sobretudo aqueles relacionados com a IA, possuem vieses.

Apontamos para o fato de que existem limites da conceitualização da agência algorítmica como viés, a saber: o foco nos resultados individualmente considerados ético ou antiético, que não dão conta de apelar a questões políticas mais estruturais; a enganosa consideração de que os algoritmos podem e devem ser desviesados, o que limita o horizonte de possibilidades destes na transformação social; e subjacente a ambos, uma ideia de fundo de neutralidade da tecnologia, que não os atribui uma verdadeira agência, considerando que as tendências presentes neles são fruto de erros de funcionamento ou de intencionalidades humanas instrumentalmente realizadas por meio deles. Assim, buscamos construir uma conceitualização alternativa.

Partindo de um debate sobre a materialidade dos sistemas de computação que questiona as distinções software/hardware, material/imaterial e simbólico/concreto, concluímos que um algoritmo é misto de regras lógicas e operações materiais. Na convergência de duas tendências tecnológicas – a programabilidade e a microeletrônica digital – o computador digital é o objeto técnico central do funcionamento de algoritmos. Amparados em uma releitura de Simondon, localizamos os componentes microeletrônicos como elementos técnicos responsáveis pela tecnicidade inerente à computação; o computador digital como indivíduo técnico preponderante da tecnologia contemporânea; e os sistemas de computação como conjuntos técnicos no qual a informação digital é mediada por algoritmos. Concluímos assim que os algoritmos são os mediadores da informação digital em sistemas de computação baseados em um computador digital programável como indivíduo técnico e em componentes microeletrônicos como base elementar.

Assim, baseado na noção de mediação técnica de Latour, cunhamos o termo mediação algorítmica como maneira alternativa de descrição da agência algorítmica em relação ao debate ao redor do viés. A mediação algorítmica é o modo como ação e significação resultam do controle da informação digital em um sistema de computação. A partir disso, pudemos mostrar ainda mediações algorítmicas específicas do AdM: a “regeneração do passado” (Pasquinelli e Joler 2021) que acontece por meio de um laço de realimentação conservador; a estereotipagem, consequência do automatismo e da compressão de informação; e a participação na produção e reprodução de discriminações sociais como o racismo e o sexismo algorítmico.

Entretanto, embora nosso trabalho tenha focado nas consequências negativas da aprendizagem de

máquina e, de maneira geral, nas mediações algorítmicas socialmente negativas de nossa perspectiva; acreditamos que uma das contribuições de nosso trabalho é justamente, por meio da noção de mediação algorítmica, abrir margem para investigações que sejam capazes de descrever mediações positivas, contribuindo para uma potencialização da tecnologia na transformação social. Em outras palavras, para além do viés, entendido como um defasamento indesejado em relação ao funcionamento reto e neutro, é possível conceber mediações transformadoras, que contribuam ativamente para a redução de assimetrias de poder e que nos ajude a construir um mundo alternativo: mediações feministas, antirracistas e ainda outras que sejam capazes de imaginar e realizar mundos que efetivamente desejamos. Não queremos, com nossa descrição de mediação algorítmica, reificar ou tipificar as tecnologias como sendo individualmente consideradas como exercendo mediações positivas ou negativas. Mas, ao contrário, acreditamos que esta seja uma ferramenta teórica que pode ser utilizada em conjunto com uma visão construtivista da tecnologia e da sociedade, como forma de análise das mediações que de fato existem e que podem ser contraditórias, inclusive em uma mesma tecnologia. Por entender que essa é uma contribuição importante para a sociologia, esboçaremos abaixo alguns exemplos de onde acreditamos ser possível identificar mediações transformadoras, esperando que este trabalho possa ser ponto de partida para desenvolvimentos que explorem mediações libertadoras no seio das tecnologias digitais.

Resistências, desvios e alternativas

No campo das artes, é possível perceber explorações criativas da IA sobretudo em trabalhos que partem de suas falhas ou através de usos desviantes dessas tecnologias. As obras do pesquisador Cesar Baio podem ser entendidas nesse registro: o trabalho *Espaços de falha* (2014-2019) cria fotografias através de um uso desviante dos algoritmos de alinhamento de bordas em aplicativos de fotografia panorâmica em smartphones, criando cenas que distorcem o espaço referente e incluindo “buracos” não preenchidos; enquanto que o trabalho *Percursos contra-GPS* (2019) cria gravuras a partir do registro de percursos sugeridos em aplicativos de navegação via GPS quando o artista não segue as instruções do aplicativo, de modo que o percurso muda constantemente¹¹. Como parte do coletivo Cesar & Lois, parceria com a artista Lucy HG Solomon, Baio também participou da construção da obra de arte *Degenerative cultures* [Culturas degenerativas]. A instalação – que visa “cria[r] uma rede biológica-tecnológica em que organismos vivos, redes digitais e inteligência artificial trabalham juntas para examinar a relação entre humanidade e natureza” – é composta por livros físicos inoculados com esporos do organismo *Phy-sarum polycephalum* de onde captura-se informação para a aprendizagem de máquina de um “fungo digital” que replica o comportamento do fungo físico. Se este se alimenta de porções de papel, destruindo parcialmente os livros, o fungo digital digere textos online gerando restos poéticos que são publicadas no X/Twitter online. Desse modo, a instalação parte da capacidade do AdM de replicar um comportamento, mas que é utilizado de uma maneira completamente nova: a pretensão não é criar um fungo digital como mimese fidedigna do fungo vivo, mas sim utilizar o fungo vivo como potencial criativo para a construção explicitamente artificial do fungo digital enquanto obra de arte¹².

Pasquinelli e Joler (2021) também dedicam uma parte de seu texto-mapa-manifesto sobre AdM à “inteligência adversarial”, ou seja, àquela que surge em resistência ou em reação à inteligência de

¹¹Esses trabalhos são apresentados pelo autor neste vídeo: <https://www.youtube.com/watch?v=60miOixnTXI> (acesso: 13/12/2024).

¹²Conferir a página do projeto <https://cesarandlois.org/digitalfungus/> (acesso: 16/12/2024).

máquina: a exploração de “difrações estatísticas” e “alucinações”, “hackeamentos”, estudos de como o AdM “quebra”, práticas de “envenenamento de dados”, entre outros (Pasquinelli e Joler 2021:1277). Um exemplo citado no artigo é o trabalho *HyperFace* (2016) de Adam Harvey. A obra é um “tecido de camuflagem [...] que engana algoritmos de visão computacional para ver múltiplos rostos humanos onde não há nenhum” (Pasquinelli e Joler 2021:1277). Aqui a mediação algorítmica é mobilizada de uma forma ainda mais externa, a capacidade de classificação de rostos é desviada de sua função e domesticada pelos padrões impressos que servem como uma camuflagem por excesso.

Esses trabalhos, tanto quanto os de Eldagsen, Beiguelman e Anadol citados ao longo da dissertação, dão pista para um potencial criativo da IA: através dos desvios, das falhas e de usos não convencionais que subvertem as mediações algorítmicas que foram originalmente construídas visando a automatização, a homogeneização, o conservadorismo e a estereotipia. Aqui, vemos também como as mesmas mediações algorítmicas – como a estereotipia – podem servir para usos distintos, a depender do contexto criado pela artista que ressignifica as ações e significações produzidas algorítmicamente.

Já em relação à tecnologia e movimentos sociais, um exemplo explicitamente descrito como um caso de alternativa à tecnologia dominante é o “Mapa do Feminicídio” proposto no México. Ele é apresentado por Paula Ricaurte (2019) no contexto de sua reflexão sobre a colonialidade do poder e dadoficação, particularmente sobre a possibilidade de pensarmos e construirmos tecnologias decoloniais no seio do “colonialismo de dados”: “O México carece de dados sobre questões delicadas, e as iniciativas dos cidadãos tentam preencher essa lacuna de informações. *O mapa de feminicídios é um bom exemplo de como os dados podem ser usados para promover a justiça e os direitos humanos, nesse caso honrando as memórias das milhares de mulheres mortas pelas estruturas de violência econômica e patriarcal*” (Ricaurte 2019:11 grifo nosso). A autora apresenta como o Mapa é uma reflexão crítica sobre poder, violência, gênero, tecnologia e conhecimento que “reflet[e] as raízes da violência estrutural” e com “uma decisão explícita de usar software e ferramentas livres” (Ricaurte 2019:11). Para Ricaurte, o projeto explicita o valor das iniciativas cidadãs ou comunitárias, que possuem muitas dificuldades ao trabalhar com esse tipo de dado que não é facilmente acessível; desse modo, os dados levantados no projeto têm uma importância em “tornar visível a colonialidade do poder e a matriz de dominação que oculta a continuidade das formas de exclusão econômica, tecnológica, política, física, cognitiva e emocional”. Concomitantemente, o projeto mostra como é possível gerar “contra-narrativas” que “contribuem para a compreensão da importância do problema de gênero” (Ricaurte 2019:11). Podemos pensar como, em certa medida, mesmo que ainda incipiente, o Mapa tem um efeito de mediação algorítmica feminista, ou ao menos de resistência ao patriarcalismo, pois desafia a naturalização e/ou apagamento das mortes de mulheres ao construir uma visibilidade de casos de feminicídio a uma população que carecia desse dado. Entretanto, notamos também como a própria criação desses dados abre margem para usos que não necessariamente podem ser considerados feministas (alimentar bases de dados de grandes empresas de tecnologia, por exemplo, ampliando seu potencial de exploração e dominação).

No Brasil, coletivos como a Rede Mocambos¹³ são exemplos singulares de criação e uso de tecnolo-

¹³Ela é uma “rede formada por comunidades quilombolas, indígenas, povos de matriz africana, ribeirinhas e caiçaras que busca a utilização e apropriação das tecnologias de informação e comunicação como estratégias fundamentais para desenvolvimento de ações conjuntas das comunidades envolvidas na Rede” – como afirma Dara, integrante da Rede, citada em Paciornik (2021:311). A Rede se origina em 2001 a partir da articulação de comunidades quilombolas, pontos de cultura, ativistas negros e do software livre e têm como um de seus pontos centrais a Casa de Cultura Tainã, fundada por TC Silva, em Campinas-SP. O objetivo da Rede, seguindo a máxima de Zumbi dos Palmares, é “construir um mundo mais do nosso jeito” (Rede Mocambos

gias digitais alternativas ao modelo dominante. Uma de suas tecnologias é o software Baobáxia (BBX): uma rede federada que visa preservar a memória da Rede. Ela permite conectar, por exemplo, um Data Center Comunitário Livre e uma outra mucua com acesso eventual a internet¹⁴, formando um acervo digital distribuído capaz de integrar comunidades e compartilhar informação mesmo incluindo territórios com diferentes modos de acesso à internet (global, local ou eventual). Dessa forma, a mediação algorítmica do BBX permite a criação, o compartilhamento e curadoria comunitária de acervos de mídia espalhados em diferentes territórios como a Casa de Cultura Tainã, quilombo urbano de Campinas, e Aldeia Xavante Wede'Rã, da Terra Indígena Pimentel Barbosa, MT.

No que diz respeito ao aprendizado de máquina, o próprio PPB em conjunto com a métrica interseccional apresentados por Buolamwini e Gebre (2018) podem ser considerados um caso de mediação algorítmica que visa dar visibilidade para assimetrias apresentadas por outros algoritmos. Outro uso alternativo é a utilização de aprendizagem de máquina para a preservação de línguas indígenas, como o conhecido caso dos Maori na Nova Zelândia¹⁵, mas que também tem sido pensado no Brasil¹⁶. Nesse caso, embora o projeto também levante questões sobre as consequências da captura de informação digital sobre línguas de povos minoritários, o conservadorismo inerente ao AdM encontra um uso interessante: os MdLs podem ser utilizados para imaginar línguas com poucos falantes vivos, proliferando e conservando um discurso criado nesta língua. Acreditamos que um tal projeto ainda possui muitos desafios para que seja considerado um verdadeiro uso libertador do AdM, mas ele ao menos aponta para uma direção distinta de mobilização do esforço tecnocientífico.

Um outro uso possível seria utilizar aprendizagem de máquina explicitamente programada como uma ação afirmativa: do mesmo modo como os trabalhos sobre viés tendem a reduzir disparidades entre diferentes grupos, coletividades específicas podem ser privilegiadas. No caso do PPB, por exemplo, as fotografias de parlamentares negras poderiam ser utilizadas em um sistema que fosse utilizado somente para o processamento de imagens de rostos de pessoas negras, o que provavelmente aumentaria a eficácia de uma tarefa como a classificação nesse grupo de pessoas. Por outro lado, nossa própria dissertação refletiu sobre as dificuldades de encontrar um uso legítimo para a classificação de pessoas baseadas em tecnologias de reconhecimento facial, sendo este uso, ainda pouco transformador. De uma perspectiva distinta, nos perguntamos finalmente, sobre algumas outras operações especuladas: qual seria o impacto de treinar um MdL utilizando apenas uma literatura feminista? E utilizando textos de ficção científica escritos por mulheres? Poderíamos encontrar alguma potência na criação de um chat alternativo treinado nesse tipo de dados? Pensadas para além da pretensão representativa do AdM, algumas novas possibilidades se abrem, cabe a nós continuarmos atentos às brechas da tecnologia dominante, a fim de construir nossas linhas de fuga.

Acreditamos que a noção de mediação, ao fugir de um certo apego pela neutralidade da tecnologia, não só amplia a capacidade explicativa de casos de reprodução de assimetrias – como explorado

sem data).

¹⁴Termos utilizados para descrever, respectivamente, centros de dados baseados em software livre e geridos pela comunidade, e computadores que proveem acesso a uma rede (“servidores web”) em geral (Rede Mocambos 2024).

¹⁵<https://www.wired.com/story/maori-language-tech/> (acesso: 12/12/2024).

¹⁶<https://anaind.org.br/noticias/preservacao-das-linguas-indigenas-ganha-impulso-com-inteligencia-artificial/> (acesso: 12/12/2024).

aqui para o caso de AdM –, mas também permite analisar experiências que desafiam as normas sociais. Dessa maneira, embora essa perspectiva tenha sido apenas delineada nesta conclusão, acreditamos que ela pode servir de base para outros trabalhos que busquem descrever tecnologias contra-hegemônicas, alternativas e libertadoras. Além disso, esperamos ter adicionado alguns elementos para questionar visões da agência algorítmica que reduzam-na ou desloquem-na para outros atores (programadoras ou bases de dados), mostrando que há um componente da ação que é própria das tecnologias algorítmicas. Finalizamos com o desejo de que esse debate, ao nosso ver tão importante para o entendimento e construção da realidade contemporânea, possa continuar e dar novos frutos no seio da sociologia da tecnologia, a fim de potencializar uma teoria-prática sociológica politicamente responsável, situadamente objetiva e interessada na construção de um futuro comprometido com a diversidade e com a potência das lutas pela transformação social.

Referências

3Blue1Brown, dir. 2017a. [Gradient Descent, How Neural Networks Learn | Chapter 2, Deep Learning](#).

3Blue1Brown, dir. 2017b. [What Is Backpropagation Really Doing? | Chapter 3, Deep Learning](#).

Abbas, Assad, Limin Zhang, e Samee U. Khan. 2014. «A Literature Review on the State-of-the-Art in Patent Analysis». *World Patent Information* 37:3–13. doi: [10.1016/j.wpi.2013.12.006](#).

Agre, Philip E. 1994. «Surveillance and Capture: Two Models of Privacy». *The Information Society* 10(2):101–27. doi: [10.1080/01972243.1994.9960162](#).

Alberti, Leon Battista. 2014. *Da Pintura*. 4.^a ed. Campinas: Editora da Unicamp.

Alves da Silveira, Rafael, e Pedro Peixoto Ferreira. 2009. «[Considerações Acerca Do Trabalho Imaterial e Da Produção de Valor No Capitalismo Contemporâneo](#)».

Amadeu da Silveira, Sérgio. 2019. «A noção de modulação e os sistemas algorítmicos». *PAULUS: Revista de Comunicação da FAPCOM* 3(6, 6). doi: [10.31657/rcp.v3i6.111](#).

Anadol, Refik. 2024. «Machine Hallucinations — Nature Dreams». Obtido 12 de dezembro de 2024 (<https://refikanadol.com/works/machine-hallucinations-nature-dreams/>).

Angwin, Julia, Jeff Larson, Surya Mattu, e Lauren Kirchner. 2016. «[Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And It’s Biased Against Blacks](#).» *ProPublica*, maio 23.

Anón. 2024. «What Is Machine Learning (ML)? | IBM». Obtido 27 de abril de 2024 (<https://www.ibm.com/topics/machine-learning>).

Anón. sem data. «This Person Does Not Exist». Obtido 4 de dezembro de 2024 (<https://thispersondoesnotexist.com/>).

Aradhya, Hrishikesh, Wei Hua, Rwei-Sung Lin, e Mohammed Saberian. 2017. «[Systems and Methods for Prioritizing Notifications on Mobile Devices](#)».

- Assange, Julian. 2015a. *Cypherpunks: Liberdade e o Futuro Da Internet*. Boitempo Editorial.
- Assange, Julian. 2015b. *Quando o Google Encontrou o Wikileaks*. Boitempo Editorial.
- Astray, Miles. 2024. «Reclaiming the brAIIn: How My Real Photo F L A M I N G O N E Won an AI Award». Obtido 18 de setembro de 2024 (<https://www.milesastray.com/news/newsflash-reclaiming-the-brain>).
- Baio, Cesar. 2022. «Da Ilusão Especular à Performatividade Das Imagens». *Significação: Revista de Cultura Audiovisual* 49(57, 57):80–102. doi: [10.11606/issn.2316-7114.sig.2022.183203](https://doi.org/10.11606/issn.2316-7114.sig.2022.183203).
- Bateson, Gregory. 1979. *Mind and Nature: A Necessary Unity*. 1st ed. New York: Dutton.
- Beiguelman, Giselle. 2021. *Políticas Da Imagem: Vigilância e Resistência Na Dadosfera*. Ubu Editora.
- Bengio, Yoshua, Réjean Ducharme, Pascal Vincent, e Christian Jauvin. 2003. «[A Neural Probabilistic Language Model](#)». *Journal of Machine Learning Research* 3:1137–55.
- Benjamin, Ruha. 2019. «Assessing Risk, Automating Racism». *Science* 366(6464):421–22. doi: [10.1126/science.aaz3873](https://doi.org/10.1126/science.aaz3873).
- Benjamin, Walter. 1993. «Pequena História Da Fotografia». Pp. 91–107 em *Obras Escolhidas: Magia e Técnica, Arte e Política*. São Paulo: Editora Brasiliense.
- Bezerra, Arthur Coelho, e Marco Antonio Almeida. 2020. «[Rage Against the Machine Learning: A Critical Approach to the Algorithmic Mediation of Information](#)». *Brazilian Journal of Information Science* 14(2):6–23.
- boyd, danah. 2012. «[Participating in the Always-on Lifestyle](#)». Pp. 71–76 em *The Social Media Reader*, editado por M. Mandiberg. New York: New York University Press.
- Braun, Max Benjamin, e Nirmal Jitendra Patel. 2022. «[Message Suggestions](#)».
- Breitzman, Anthony F., e Mary Ellen Moguee. 2002. «The Many Applications of Patent Analysis». *Journal of Information Science* 28(3):187–205. doi: [10.1177/016555150202800302](https://doi.org/10.1177/016555150202800302).
- Brin, Sergey, e Lawrence Page. 1998. «The Anatomy of a Large-Scale Hypertextual Web Search Engine». *Computer Networks and ISDN Systems* 30(1-7):107–17. doi: [10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X).
- Brown, Sara. 2021. «Machine Learning, Explained». Obtido 15 de maio de 2024 (<https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>).

- Bruno, Fernanda. 2024. «Mediação algorítmica em plataformas digitais: racionalidade, subjetividade e conduta.» Obtido 16 de outubro de 2024 (<https://medialabufrij.net/projetos/mediacao-algoritmica-em-plataformas-digitais-racionalidade-subjetividade-e-conduta/>).
- Bruno, Fernanda Glória, Anna Carolina Franco Bentes, e Paulo Faltay. 2019. «Economia psíquica dos algoritmos e laboratório de plataforma: mercado, ciência e modulação do comportamento». *Revista FAMECOS* 26(3):e33095–95. doi: [10.15448/1980-3729.2019.3.33095](https://doi.org/10.15448/1980-3729.2019.3.33095).
- Bruno, Fernanda Glória, Paula Cardoso Pereira, Anna Carolina Franco Bentes, Paulo Faltay, Mariana Antoun, Debora Dantas Pio da Costa, Helena Strecker, e Natássia Salgueiro Rocha. 2021. «“Tudo por conta própria”: autonomia individual e mediação técnica em aplicativos de autocuidado psicológico». *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde* 15(1, 1). doi: [10.29397/reciis.v15i1.2205](https://doi.org/10.29397/reciis.v15i1.2205).
- Bruno, Fernanda Glória, e Pablo Manolo Rodríguez. 2022. «The Dividual: Digital Practices and Biotechnologies». *Theory, Culture & Society* 39(3):27–50. doi: [10.1177/02632764211029356](https://doi.org/10.1177/02632764211029356).
- Buolamwini, Joy, e Timnit Gebru. 2018. «Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification». Pp. 1–18 em *Proceedings of Machine Learning Research*. Vol. 81.
- Carbone, Victor, e Sandro Feuz. 2022. «Using Live Data Streams and/or Search Queries to Determine Information about Developing Events».
- Cardon, Dominique, Jean-Philippe Cointet, e Antoine Mazieres. 2018. «La Revanche Des Neurones: L’invention Des Machines Inductives et La Controverse de l’intelligence Artificielle». *Réseaux* 5(211). doi: [10.3917/res.211.0173](https://doi.org/10.3917/res.211.0173).
- Cassino, João, Joyce Souza, e Sérgio Amadeu da Silveira, eds. 2021. *Colonialismo de Dados: Como Opera a Trincheira Algorítmica Na Guerra Liberal*. São Paulo, SP: Autonomia Literária.
- Chatterjee, Surojit, Alex Gawley, e Chrix Erik Finne. 2015. «Method and Apparatus Using Accelerometer Data to Serve Better Ads».
- Cheong, Marc, Ehsan Abedin, Marinus Ferreira, Ritsaart Reimann, Shalom Chalson, Pamela Robinson, Joanne Byrne, Leah Ruppner, Mark Alfano, e Colin Klein. 2024. «Investigating Gender and Racial Biases in DALL-E Mini Images». *ACM Journal on Responsible Computing*. doi: [10.1145/3649883](https://doi.org/10.1145/3649883).
- Christiano, Paul, Jan Leike, Tom B. Brown, Miljan Martic, Shane Legg, e Dario Amodei. 2023. «Deep Reinforcement Learning from Human Preferences». Obtido 3 de julho de 2024 (<http://arxiv.org/abs/1706.03741>).
- Collins, Patricia Hill. 2016. «Aprendendo com a outsider within». *Sociedade e Estado* 31:99–127. doi:

[10.1590/S0102-69922016000100006](https://doi.org/10.1590/S0102-69922016000100006).

Cordova, Yasodara, Lorryne Porciuncula, e Henri Brebant. 2018. «Big Data, Meager Returns?» Obtido 27 de novembro de 2024 (<https://medium.com/digitalhks/big-data-meager-returns-c7e7beceb3a7>).

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest, e Clifford Stein. 2022. *Introduction to Algorithms*. MIT press.

Couldry, Nick, e Ulises A. Mejias. 2019. «Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject». *Television & New Media* 20(4):336–49. doi: [10.1177/1527476418796632](https://doi.org/10.1177/1527476418796632).

Couldry, Nick, e Ulises A. Mejias. 2020. *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press.

Crawford, Kate. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

Crawford, Kate, Roel Dobbe, Theodora Dryer, Genevieve Fried, Ben Green, Elizabeth Kazianas, Amba Kak, Varoon Mathur, Erin McElroy, Andrea Nill Sánchez, e others. 2019. «[AI Now 2019 Report](#)». *New York, NY: AI Now Institute* 100.

Crawford, Kate, e Vladan Joler. 2018. «[Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources](#)». *AI Now Institute and Share Lab*.

Crawford, Kate, e Trevor Paglen. 2019. «[Excavating AI: The Politics of Images in Machine Learning Training Sets](#)». *The AI Now Institute, NYU*.

Dardot, Pierre, e Christian Laval. 2016. *A Nova Razão Do Mundo*. 1.^a ed. São Paulo: Boitempo.

Dardot, Pierre, e Christian Laval. 2017. *Comum: Ensaio Sobre a Revolução No Século XXI*. Boitempo Editorial.

Deleuze, Gilles. 1996. «O Atual e o Virtual». Pp. 47–57 em *Deleuze: Filosofia Virtual*. São Paulo: Editora 34.

Deleuze, Gilles. 2013. *Conversações*. 3.^a ed. São Paulo: Editora 34.

Deleuze, Gilles, e Félix Guattari. 2011a. «Balanço-Programa Para Máquinas Desejantes». Pp. 507–34 em *O Anti-édipo: Capitalismo e Esquizofrenia 1*. São Paulo: Editora 34.

Deleuze, Gilles, e Félix Guattari. 2011b. *O Anti-édipo: Capitalismo e Esquizofrenia 1*. São Paulo: Editora 34.

- Deleuze, Gilles, e Félix Guattari. 2011–2012. *Mil platôs: capitalismo e esquizofrenia 2*. 2.^a ed. São Paulo: Editora 34.
- Deleuze, Gilles, e Félix Guattari. 2018. *Kafka: Por Uma Literatura Menor*. Autêntica.
- Deleuze, Gilles, e Claire Parnet. 1998. *Diálogos*. São Paulo: Escuta.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, e Kristina Toutanova. 2019. «BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding». Obtido 27 de dezembro de 2020 (<http://arxiv.org/abs/1810.04805>).
- Diaz-Isenrath, Cecilia. 2005. «Um estudo sobre google: questões para uma leitura micropolítica das tecnologias da informação».
- Dijck, José van. 2014. «Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology». *Surveillance & Society* 12(2):197–208. doi: [10.24908/ss.v12i2.4776](https://doi.org/10.24908/ss.v12i2.4776).
- Domingues, Ana Carolina. 2020. «A perspectiva renascentista: a educação dos modos de ver entre e o real e o produzido». *Cadernos da Pedagogia* 14(28):51–59.
- Eldagsen, Boris. 2023a. «PSEUDOMNESIA». Obtido 19 de setembro de 2024 (<https://www.eldagsen.com/pseudomnesia/>).
- Eldagsen, Boris. 2023b. «Sony World Photography Awards 2023». Obtido 18 de setembro de 2024 (<https://www.eldagsen.com/sony-world-photography-awards-2023/>).
- Elish, Madeleine Clare, e danah boyd. 2018. «Situating Methods in the Magic of Big Data and AI». *Communication Monographs* 85(1):57–80. doi: [10.1080/03637751.2017.1375130](https://doi.org/10.1080/03637751.2017.1375130).
- Evangelista, Rafael. 2019. «Review of Zuboff's The Age of Surveillance Capitalism». *Surveillance & Society* 17(1/2):246–51. doi: [10.24908/ss.v17i1/2.13132](https://doi.org/10.24908/ss.v17i1/2.13132).
- Evangelista, Rafael. 2023. «Por uma etnografia do poder na inteligência artificial, no capitalismo de vigilância e no colonialismo digital». *Aurora. Revista de Arte, Mídia e Política* 16(47):112–33. doi: [10.23925/1982-6672.2022v15i47p112-133](https://doi.org/10.23925/1982-6672.2022v15i47p112-133).
- Faustino, Deivison, e Walter Lippold. 2023. *Colonialismo digital: Por uma crítica hacker-fanoniana*. Boitempo Editorial.
- Faustino, Rafael. 2023. «Artistas protestam contra uso de suas obras para alimentar inteligência artificial». *Época NEGÓCIOS*, janeiro 12.

- Ferreira, Pedro Peixoto. 2004. «Máquinas sociais: o Filo Maquínico e a Sociologia da Tecnologia». P. 7 em.
- Ferreira, Pedro Peixoto. 2010. «Por uma definição dos processos tecnicamente mediados de associação». *Revista Brasileira de Ciência, Tecnologia e Sociedade* 1(2):58–75.
- Ferreira, Pedro Peixoto. 2017. «Reticulações: Ação-Rede Em Latour e Simondon». *Revista ECO-Pós* 20(1):104–35.
- Ferreira, Pedro Peixoto. 2024. «“A Vida Eletrônica Da Perspectiva Dos Processos Tecnicamente Mediados de Associação: Gravação Da Prova Didática Do Concurso Para Obtenção de Título de Livre Docente Na Área de «Mudanças Sociais: Questões Em Ambiente e Tecnologia», Na Disciplina «HZ059 – Sociologia Da Tecnologia», Do Departamento de Sociologia (DS) Do Instituto de Filosofia e Ciências Humanas (IFCH) Da Universidade Estadual de Campinas (Unicamp).”»
- Flusser, Vilém. 2008. *O Universo Das Imagens Técnicas: Elogio Da Superficialidade*. São Paulo: Annablume.
- Flusser, Vilém. 2018. *Filosofia Da Caixa Preta: Ensaio Para Uma Filosofia Da Fotografia*. 1.^a ed. São Paulo: É Realizações Editora.
- Flusser, Vilém. sem data. «“Nascimento de Imagem Nova”».
- Foucault, Michel. 2014. «Disciplina». em *Vigiar e Punir*. Leya.
- Foucault, Michel. 2021. «Soberania e Disciplina». em *A Microfísica Do Poder*. São Paulo: Paz e Terra.
- França, Ricardo Orlandi. 1997. «Patente Como Fonte de Informação Tecnológica». *Perspectivas Em Ciência Da Informação* 2(2, 2).
- Friedman, Batya, e Helen Nissenbaum. 1996. «Bias in Computer Systems». *ACM Trans. Inf. Syst.* 14(3):330–47. doi: [10.1145/230538.230561](https://doi.org/10.1145/230538.230561).
- Fuchs, Christian. 2012. «Google Capitalism». *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society* 10(1):42–48. doi: [10.31269/triplec.v10i1.304](https://doi.org/10.31269/triplec.v10i1.304).
- Galloway, Alexander. 2019. «Are Algorithms Biased?» Obtido 30 de março de 2020 (<http://cultureandcommunication.org/galloway/are-algorithms-biased>).
- Garcia dos Santos, Laymert. 1981. «Questão de Método». Pp. 10–49 em *Desregularens: Educação, Planejamento e Tecnologia Como Ferramenta Social*. Campinas: Brasiliense, Fundação de Desenvolvimento da Unicamp.

- Garcia dos Santos, Laymert. 2000. «Código Primitivo - Código Genético: A Consistência de Uma Vizinhança». Pp. 415–20 em *Gilles Deleuze: Uma Vida Filosófica*. São Paulo: Editora 34.
- Garcia dos Santos, Laymert. 2003a. «A informação após a virada cibernética». Pp. 9–34 em *Revolução tecnológica, internet e socialismo, Socialismo em discussão*. São Paulo, SP, Brasil: Editora Fundação Perseu Abramo.
- Garcia dos Santos, Laymert. 2003b. *Politizar as Novas Tecnologias: o Impacto Sócio-Técnico Da Informação Digital e Genética*. 1.^a ed. São Paulo: Editora 34.
- Garcia dos Santos, Laymert. 2007. «Paradoxos da propriedade intelectual». em *Propriedade intelectual: tensões entre o capital e a sociedade*, editado por F. Villares. São Paulo: Instituto de Estudos Econômicos e Internacionais : Paz e Terra.
- Garcia dos Santos, Laymert. 2022. «Por uma sociologia não autocrática das máquinas». *Ideias* 13:e022026:1–24. doi: [10.20396/ideias.v13i00.8671667](https://doi.org/10.20396/ideias.v13i00.8671667).
- Garcia dos Santos, Laymert, e Pedro Peixoto Ferreira. 2008. «A Regra Do Jogo: Desejo, Servidão e Controle». *Novas mídias digitais: impactos políticos, econômicos e sociais*. Rio de Janeiro: E-papers.
- Gebru, Timnit. 2020. «Race and Gender». Pp. 252–69 em *The Oxford Handbook of Ethics of AI*, editado por M. D. Dubber, F. Pasquale, e S. Das. Oxford University Press.
- Genaro, Ednei de. 2015. «Imagens operativas e pós-fotográficas: um estudo a partir de Farocki». *Revista Eco-Pós* 18(2, 2):134–50. doi: [10.29146/eco-pos.v18i2.2345](https://doi.org/10.29146/eco-pos.v18i2.2345).
- Gillespie, Tarleton. 2014. «The Relevance of Algorithms». Pp. 167–94 em *Media Technologies*, editado por T. Gillespie, P. J. Boczkowski, e K. A. Foot. The MIT Press.
- Gillespie, Tarleton. 2018. «A relevância dos algoritmos». *Parágrafo* 6(1):95–121.
- GitHub. 2024. «GitHub Copilot - Your AI Pair Programmer». Obtido 2 de agosto de 2023 (<https://github.com/features/copilot>).
- Gonçalves, Rafael. 2022. «Automatismo ontem e hoje: reflexões sobre os limites da inteligência artificial a partir de Simondon». *Ideias* 13:e022008. doi: [10.20396/ideias.v13i00.8668218](https://doi.org/10.20396/ideias.v13i00.8668218).
- Gonçalves, Rafael. 2024a. «Captura como fenômeno sociotécnico: apontamentos para uma caracterização a partir de tecnologias de inteligência artificial». *Revista Pimentalab* 3 (Tecnoceno):30–38. doi: [10.34024/2764859n.3](https://doi.org/10.34024/2764859n.3).
- Gonçalves, Rafael. 2024b. «Mediação Algorítmica, Viés de Gênero e Sexismo Algorítmico». Pp. 1526–

- 34 em *Anais Da XV Jornada Latino-Americana de Estudos Sociais Da Ciência e Tecnologia*. Campinas.
- Gonçalves, Rafael, e Romis Attux. 2019. «Fundamentals of deep neural networks - application based in convolutional networks». *Revista dos Trabalhos de Iniciação Científica da UNICAMP* (27, 27):1–1. doi: [10.20396/revpibic2720193040](https://doi.org/10.20396/revpibic2720193040).
- Gonçalves, Rafael, e Pedro Peixoto Ferreira. 2022a. «Aprendizado de máquina como mediação técnica computacional: viés de gênero no processamento automático de textos relacionando profissões pelo algoritmo BERT». em *Anais [recurso eletrônico] : resumos da 74ª Reunião Anual da SBPC*.
- Gonçalves, Rafael, e Pedro Peixoto Ferreira. 2022b. «Eficiência Como Viés Algorítmico Nas Técnicas de Aprendizado de Máquina: Caracterização Baseada Na Produção Tecnocientífica Da Google». em *XXX Congresso de Iniciação Científica Da UNICAMP*.
- Gonçalves, Rafael, e Giordanno Oliveira Padovan. 2024a. «Captura da virtualidade em algoritmos bioinspirados: apropriação da informação em práticas de modelagem computacional biomimética». em *Anais da 34ª Reunião Brasileira de Antropologia*. Belo Horizonte: ABA - Associação Brasileira de Antropologia.
- Gonçalves, Rafael, e Giordanno Oliveira Padovan. 2024b. «Coletivos maquínicos e sistemas digitais: Tensionamentos da noção de humanidade a partir de técnicas de aprendizado de máquina patenteadas pela Google». *Anais da ReACT - Reunião de Antropologia da Ciência e Tecnologia* 6(6, 6):1–23.
- Goodfellow, Ian J., Jonathon Shlens, e Christian Szegedy. 2015. «Explaining and Harnessing Adversarial Examples». Obtido 24 de setembro de 2022 (<http://arxiv.org/abs/1412.6572>).
- Goodfellow, Ian, Yoshua Bengio, e Aaron Courville. 2016. *Deep Learning*. MIT Press.
- Google. 2005. «Google Code of Conduct». Obtido 25 de junho de 2024 (<https://web.archive.org/web/20050204181615/http://investor.google.com/conduct.html>).
- Google. 2020. «Search On 2020». Obtido 21 de agosto de 2021 (<https://searchon.withgoogle.com/>).
- Google. sem data. «Patents». Obtido 23 de novembro de 2024 (<https://patents.google.com/>).
- Han, Byung-Chul. 2023. *Psicopolítica: O Neoliberalismo e as Novas Técnicas de Poder*. 10.ª ed. Belo Horizonte: Editora Ayiné.
- Haraway, Donna. 1988. «Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective». *Feminist Studies* 14(3):575–99. doi: [10.2307/3178066](https://doi.org/10.2307/3178066).
- Haraway, Donna. 2009. «Manifesto Ciborgue: Ciência, Tecnologia e Feminismo-Socialista No Final Do

- Século XX». Pp. 33–118 em *Antropologia Do Ciborgue: As Vertigens Do Pós-Humano*, editado por T. Tadeu. Belo Horizonte: Autêntica Editora.
- Haraway, Donna. 2023a. *A Reinvenção Da Natureza: Símios, Ciborgues e Mulheres*. WMF Martins Fontes.
- Haraway, Donna. 2023b. *Ficar Com o Problema: Fazer Parentes No Chthuluceno*. n-1 edições.
- Haykin, Simon S. 2009. *Neural Networks and Learning Machines*. 3rd ed. New York: Prentice Hall.
- Hebb, Donald O. 1949. *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley.
- Hochreiter, Sepp, e Jürgen Schmidhuber. 1997. «Long Short-Term Memory». *Neural Computation* 9(8):1735–80. doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735).
- Hui, Yuk. 2012. «What Is a Digital Object?» *Metaphilosophy* 43(4):380–95. doi: [10.1111/j.1467-9973.2012.01761.x](https://doi.org/10.1111/j.1467-9973.2012.01761.x).
- Hui, Yuk. 2015. «Modulation after Control». *new formations: a journal of culture/theory/politics* 84(84):74–91.
- Hui, Yuk. 2020. *Tecnodiversidade*. São Paulo: Ubu Editora.
- Joler, Vladan. 2020. «New Extractivism». Obtido 6 de agosto de 2022 (<https://extractivism.online/>).
- Joler, Vladan, e Andrej Petrovski. 2016. «Immaterial Labour and Data Harvesting». Obtido 5 de abril de 2024 (<https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/>).
- Kanashiro, Marta Mourão. 2023. «ChatGPT: o que não está se debatendo». Obtido 2 de agosto de 2023 (<https://outraspalavras.net/tecnologiaemdisputa/chatgpt-o-que-nao-esta-se-debatendo/>).
- Karras, Tero, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, e Timo Aila. 2020. «Analyzing and Improving the Image Quality of StyleGAN». Pp. 8107–16 em *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA: IEEE.
- Kasper, Christian Pierre. 2006. «Tecnologia». Pp. 125–86 em *Habitar a rua*. Campinas, SP: Tese (doutorado) - Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.
- Kitchin, Rob. 2014a. «Big Data, New Epistemologies and Paradigm Shifts». *Big Data & Society* 1(1):2053951714528481. doi: [10.1177/2053951714528481](https://doi.org/10.1177/2053951714528481).
- Kitchin, Rob. 2014b. «Thinking Critically About and Researching Algorithms». *SSRN Electronic Journal*.

doi: [10.2139/ssrn.2515786](https://doi.org/10.2139/ssrn.2515786).

Kittler, Friedrich. 1995. «[There Is No Software](#)». *CTheory*.

Kittler, Friedrich. 2016. *Mídias Ópticas: Curso Em Berlim, 1999*. 1.^a ed. Rio de Janeiro: Contraponto.

Kopenawa, Davi, e Bruce Albert. 2019. *A Queda Do Céu: Palavras de Um Xamã Yanomami*. Editora Companhia das letras.

Krishna, Golden Gopal, Carl Magnus BORG, Miroslav Bojic, Henry Owen Newton-Dunn, Jacob M. Klinker, Mindy Pereira, Devin Mancuso, Daniel June Hyung Park, e Lily Sin. 2021. «[Suggesting Actions Based on Machine Learning](#)».

Krizhevsky, Alex, Ilya Sutskever, e Geoffrey E. Hinton. 2017. «ImageNet Classification with Deep Convolutional Neural Networks». *Communications of the ACM* 60(6):84–90. doi: [10.1145/3065386](https://doi.org/10.1145/3065386).

Kwet, Michael. 2019. «Digital Colonialism: US Empire and the New Imperialism in the Global South». *Race & Class* 60(4):3–26. doi: [10.1177/0306396818823172](https://doi.org/10.1177/0306396818823172).

Lance, Marcos Calvo, e Philip Koonce. 2021. «[Automatically Resolving, with Reduced User Inputs, a Set of Activity Instances for a Group of Users](#)».

Latour, Bruno. 1990. «Technology Is Society Made Durable». *The Sociological Review* 38:103–31. doi: [10.1111/j.1467-954X.1990.tb03350.x](https://doi.org/10.1111/j.1467-954X.1990.tb03350.x).

Latour, Bruno. 1992. «[Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts](#)». Pp. 225–59 em *Shaping Technology-Building Society. Studies in Sociotechnical Change*. MIT Press, Cambridge Mass.

Latour, Bruno. 1994. «[On Technical Mediation](#)». *Common Knowledge* 3(2):29–64.

Latour, Bruno. 2002. «[Morality and Technology](#)». *Theory, culture & society* 19(5-6):247–60.

Latour, Bruno. 2004. «[Redes Que a Razão Desconhece: Laboratórios, Bibliotecas, Coleções](#)». *Tramas da rede: novas dimensões filosóficas, estéticas e políticas da comunicação*. (Trad. Marcela Mortara) Porto Alegre: Sulina 39–63.

Latour, Bruno. 2011. *Ciência Em Ação: Como Seguir Cientistas e Engenheiros Sociedade Afora*. 2.^a ed. São Paulo: Unesp.

Latour, Bruno. 2012. *Reagregando o social: uma introdução a teoria do ator-rede*. Salvador: Edufba.

- Latour, Bruno. 2017. *A Esperança de Pandora: Ensaios Sobre a Realidade Dos Estudos Científicos*. São Paulo: Editora Unesp.
- Latour, Bruno. 2019. *Jamais Fomos Modernos*. 4.^a ed. São Paulo: Editora 34.
- Latour, Bruno, e Steve Woolgar. 2013. *Laboratory Life: The Construction of Scientific Facts*. Princeton university press.
- Lazzarato, Maurizio. 2006. *As revoluções do capitalismo*. Rio de Janeiro: Civilização brasileira.
- Lazzarato, Maurizio. 2010. «Sujeição e servidão no capitalismo contemporâneo». *Cadernos de Subjetividade* (12):168–79. doi: [10.2354/cs.v0i12.38458](https://doi.org/10.2354/cs.v0i12.38458).
- Lazzarato, Maurizio, e Antonio Negri. 2001. *Trabalho Imaterial: Formas de Vida e Produção de Subjetividade*. Rio de Janeiro: DP&A.
- LeCun, Yann, Yoshua Bengio, e Geoffrey Hinton. 2015. «Deep Learning». *Nature* 521(7553):436–44. doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- Lewis, Justin, e Thomas Price. 2021. «[Using Machine Learning and Other Models to Determine a User Preference to Cancel a Stream or Download](#)».
- Lippold, Walter, e Deivison Faustino. 2022. «[Colonialismo digital, racismo e acumulação primitiva de dados](#)». *Germinal: marxismo e educação em debate* 14(2, 2):56–78.
- Luong, Minh-Thang, Hieu Pham, e Christopher D. Manning. 2015. «[Effective Approaches to Attention-based Neural Machine Translation](#)».
- Lyon, David. 2018. «Cultura Da Vigilância: Envolvimento, Exposição e Ética Na Modernidade Digital». Pp. 151–80 em *Tecnopolíticas de Vigilância: Perspectivas Da Margem*, editado por F. Bruno, B. Cardoso, M. Kanashiro, e L. Guilhon. São Paulo: Boitempo.
- Machado, Arlindo. 1984. *A Ilusão Especular: Introdução à Fotografia*. Brasiliense/Funarte.
- Machado, Arlindo. 2004. «As Imagens Técnicas: Da Fotografia à Síntese Numérica». em *Pré-Cinemas & Pós Cinemas*. Campinas: Papirus Editora.
- Machado, Débora Franco. 2019. «[Modulações Algorítmicas: Uma Análise Das Tecnologias de Orientação de Comportamento a Partir Das Patentes Do Facebook](#)». Dissertação de mestrado, São Bernardo do Campo: UFABC.
- Machado, Débora Franco. 2020. «[A modulação algorítmica de comportamento e suas categorias ope-](#)

- rativas a partir das patentes da Facebook Inc.» *Revista Eletrônica Internacional de Economia Política da Informação, da Comunicação e da Cultura* 22(2, 2):97–111.
- Marx, Karl. 2015. *Grundrisse: Manuscritos Econômicos de 1857-1858: Esboços Da Crítica Da Economia Política*. Boitempo editorial.
- Marx, Karl. 2017. *O Capital - livro 1*. 2.^a ed. São Paulo: Boitempo Editorial.
- Mattelart, Armand. 2005.«[Sociedade do Conhecimento e Controle da Informação e da Comunicação](#)». Apresentado na V Encontro Latino de Economia Política da Informação, Comunicação e Cultura, novembro 9–11, Salvador-BA, Brasil.
- Mayer-Schönberger, Viktor, e Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Houghton Mifflin Harcourt.
- McCarthy, John, Marvin Lee Minsky, Nathaniel Rochester, e Claude Elwood Shannon. 1955. «[A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence](#)».
- McCulloch, Warren S., e Walter Pitts. 1943. «A Logical Calculus of the Ideas Immanent in Nervous Activity». *The Bulletin of Mathematical Biophysics* 5(4):115–33. doi: [10.1007/BF02478259](https://doi.org/10.1007/BF02478259).
- Medialab.UFRJ. 2024. «Economia Psíquica Do Algoritmo Em Linha Do Tempo». Obtido 29 de novembro de 2024 (<https://medialabufrj.net/projetos/economia-psiquica-dos-algoritmos-em-linha-do-tempo>).
- Mehrabi, Ninareh, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, e Aram Galstyan. 2021. «A Survey on Bias and Fairness in Machine Learning». *ACM Computing Surveys* 54(6):115:1–35. doi: [10.1145/3457607](https://doi.org/10.1145/3457607).
- Mikolov, Tomas, Kai Chen, Greg Corrado, e Jeffrey Dean. 2013. «[Efficient Estimation of Word Representations in Vector Space](#)».
- Mineo, Fabrício Behrmann. 2014. «A relação entre pintura e natureza em Alberti». *Humanidades em diálogo* 6:101–16. doi: [10.11606/issn.1982-7547.hd.2014.106262](https://doi.org/10.11606/issn.1982-7547.hd.2014.106262).
- Mirelmann, Lucas, Zaheed Sabur, Bohdan Vlasyuk, Marie Patriarche Bledowski, Sergey Nazarov, Denis Burakov, Behshad Behzadi, Michael Golikov, Steve Cheng, Daniel Cotting, e Mario Bertschler. 2022. «[Performing Subtask\(s\) for a Predicted Action in Response to a Separate User Interaction with an Automated Assistant Prior to Performance of the Predicted Action](#)».
- Mitchell, Thomas. 1997. *Machine Learning*. 1.^a ed. McGraw Hill Higher Education.

- Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, e Luciano Floridi. 2016. «The Ethics of Algorithms: Mapping the Debate». *Big Data & Society* 3(2):2053951716679679. doi: [10.1177/2053951716679679](https://doi.org/10.1177/2053951716679679).
- Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fiedjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dhharshan Kumaran, Daan Wierstra, Shane Legg, e Demis Hassabis. 2015. «Human-Level Control Through Deep Reinforcement Learning». *Nature* 518(7540):529–33. doi: [10.1038/nature14236](https://doi.org/10.1038/nature14236).
- Mühlhoff, Rainer. 2020. «Human-Aided Artificial Intelligence: Or, How to Run Large Computations in Human Brains? Toward a Media Sociology of Machine Learning». *New Media & Society* 22(10):1868–84. doi: [10.1177/1461444819885334](https://doi.org/10.1177/1461444819885334).
- Mwema, Esther, e Abeba Birhane. 2024. «Undersea Cables in Africa: The New Frontiers of Digital Colonialism». *First Monday*. doi: [10.5210/fm.v29i4.13637](https://doi.org/10.5210/fm.v29i4.13637).
- Nanitchkov, Alexander. 2022. «“Arte” de IA Contemporânea é Criada Nas Costas de Centenas de Milhares de Artistas e Fotógrafos...» Obtido 2 de agosto de 2023 (<https://twitter.com/Artofinca/status/1599730391698485248>).
- Noble, Safiya Umoja. 2022. *Algoritmos Da Opressão: Como Os Mecanismos de Busca Reforçam o Racismo*. Editora Rua do Sabão.
- Ochigame, Rodrigo, e Katherine Ye. 2021. «Search Atlas: Visualizing Divergent Search Results Across Geopolitical Borders». Pp. 1970–83 em *Designing Interactive Systems Conference 2021*. Virtual Event USA: ACM.
- Olah, Christopher. 2014a. «Neural Networks, Manifolds, and Topology – Colah’s Blog». Obtido 3 de julho de 2024 (<https://colah.github.io/posts/2014-03-NN-Manifolds-Topology/>).
- Olah, Christopher. 2014b. «Visualizing MNIST: An Exploration of Dimensionality Reduction - Colah’s Blog». Obtido 3 de julho de 2024 (<https://colah.github.io/posts/2014-10-Visualizing-MNIST/>).
- Ouyang, Long, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, e Ryan Lowe. 2022. «Training Language Models to Follow Instructions with Human Feedback». Obtido 3 de julho de 2024 (<http://arxiv.org/abs/2203.02155>).
- Oztekin, Bilgehan Uygur, Daria Antonova, Kedar Dhamdhere, Finnegan Southey, e Zhenyu Mao. 2014. «Reordering Search Query Results in Accordance with Search Context Specific Predicted Perfor-

mance Functions».

Paciornik, Guilherme Flynn. 2021. «[Movimentos Sociais e Tecnologias Digitais: Cultura Digital Brasileira, Software Livre e Tecnopolítica](#)». Tese (doutorado), Campinas, SP: Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.

Page, Lawrence. 2001. «[Method for Node Ranking in a Linked Database](#)».

Parra, Henrique Zoqui Martins. 2022. «[Da Tecnopolítica Às Lutas Cosmotécnicas: Dissensos Ontoepistêmicos Face à Hegemonia Cibernética No Antropoceno](#)». Pp. 339–94 em *Engenharias e Outras Práticas Técnicas Engajadas : Diálogos Interdisciplinares e Decoloniais*.

Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

Pasquinelli, Matteo. 2009. «[Google's PageRank Algorithm: A Diagram of Cognitive Capitalism and the Rentier of the Common Intellect](#)». *Deep Search*.

Pasquinelli, Matteo. 2013. «[Capitalismo maquínico e mais-valia de rede Notas sobre a economia política da máquina de Turing](#)». *Lugar Comum – Estudos de mídia, cultura e democracia* (39, 39):13–36.

Pasquinelli, Matteo. 2021. «[How to Make a Class: Hayek's Neoliberalism and the Origins of Connectivism](#)». *Qui Parle* 30(1):159–84.

Pasquinelli, Matteo. 2023. *The Eye of the Master: A Social History of Artificial Intelligence*. Verso Books.

Pasquinelli, Matteo, e Vladan Joler. 2021. «[The Nooscope Manifested: AI as Instrument of Knowledge Extractivism](#)». *AI & SOCIETY* 36(4):1263–80. doi: [10.1007/s00146-020-01097-6](https://doi.org/10.1007/s00146-020-01097-6).

PatFT. sem data. «[Patent Full-text Databases](#)». Obtido 14 de abril de 2022 (<https://patft.uspto.gov/netahtml/PTO/index.html>).

Plant, Sadie. 1997. *Zeros + Ones: Digital Women and the New Technoculture*. London: Fourth Estate.

Plant, Sadie. 2000. «[On the Matrix: Cyberfeminist Simulations](#)». Pp. 325–36 em *The Cyberculture Reader*, editado por D. Bell e B. M. Kennedy. London: Routledge.

PPubS. sem data. «[Patent Public Search](#)». Obtido 2 de abril de 2024 (<https://ppubs.uspto.gov/pubwebapp/>).

Prates, Marcelo O. R., Pedro H. Avelar, e Luís C. Lamb. 2020. «[Assessing Gender Bias in Machine Translation: A Case Study with Google Translate](#)». *Neural Computing and Applications* 32(10):6363–

81. doi: [10.1007/s00521-019-04144-6](https://doi.org/10.1007/s00521-019-04144-6).

Quijano, Aníbal. 2009. «Colonialidade do Poder e Classificação Social». em *Epistemologias do Sul*, editado por B. de S. Santos e M. P. Meneses.

Rede Mocambos. 2024. «Núcleo de Pesquisa e Desenvolvimento Digital». Obtido 16 de dezembro de 2024 (<https://wiki.mocambos.net/index.php/NPDD>).

Rede Mocambos. sem data. «Sobre». Obtido 21 de julho de 2024 (<https://www.mocambos.net/tambor/pt/home>).

Ricourte, Paola. 2019. «Data Epistemologies, The Coloniality of Power, and Resistance». *Television & New Media* 20(4):350–65. doi: [10.1177/1527476419831640](https://doi.org/10.1177/1527476419831640).

Rico, Leonardo. [2019] 2024. «Bytesleo/Thispersondoesnotexist-Js».

Rombach, Robin, Andreas Blattmann, Dominik Lorenz, Patrick Esser, e Björn Ommer. 2022. «High-Resolution Image Synthesis with Latent Diffusion Models». Obtido 5 de dezembro de 2022 (<http://arxiv.org/abs/2112.10752>).

Rosenblatt, F. 1958. «The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain». *Psychological Review* 65:386–408. doi: [10.1037/h0042519](https://doi.org/10.1037/h0042519).

Rouvroy, Antoinette, e Thomas Berns. 2015. «Governamentalidade Algorítmica e Perspectivas de Emancipação: o Díspar Como Condição de Individuação Pela Relação?» *Revista ECO-Pós* 18(2):36–56.

Ruback, Livia, Denise Carvalho, e Sandra Avila. 2022. «Mitigando Vieses no Aprendizado de Máquina: Uma Análise Sociotécnica». *iSys - Brazilian Journal of Information Systems* 15(1):23:1–31. doi: [10.5753/isys.2022.2396](https://doi.org/10.5753/isys.2022.2396).

Rumelhart, David E., Geoffrey E. Hinton, e Ronald J. Williams. 1986. «Learning Representations by Back-Propagating Errors». *Nature* 323(6088, 6088):533–36. doi: [10.1038/323533a0](https://doi.org/10.1038/323533a0).

Salviano, Maria Cortez. 2023. «O Estatuto Da Imagem Na Arte e Na Tecnologia: A Matemática Como Modo (Diverso) de Conhecimento». *Anais eletrônicos do X Simpósio Nacional de Ciência, Tecnologia e Sociedade (Esocite.BR)*.

Samuel, Arthur L. 1959. «Some Studies in Machine Learning Using the Game of Checkers». *IBM Journal* 3(3):210–29.

Schiavetto Amancio, Stefano. 2014. «Formas Contemporâneas de Relação Entre Capital e Tecnicidade : Estudo Sobre a Gênese de Microprocessadores de Licença Proprietária e Livre». Dissertação de

mestrado, Campinas, SP: Universidade Estadual de Campinas.

Schiavetto Amancio, Stefano. 2024. «Educação neoliberal, capitalismo de plataformas e tecnologias microeletrônicas no Brasil contemporâneo». Tese de doutorado, Campinas, SP: Universidade Estadual de Campinas.

Searle, John R. 1980. «Minds, Brains, and Programs». *Behavioral and Brain Sciences* 3(3):417–24. doi: [10.1017/S0140525X00005756](https://doi.org/10.1017/S0140525X00005756).

Seaver, Nick. 2019a. «Captivating Algorithms: Recommender Systems as Traps». *Journal of Material Culture* 24(4):421–36. doi: [10.1177/1359183518820366](https://doi.org/10.1177/1359183518820366).

Seaver, Nick. 2019b. «Knowing Algorithms». em *digitalSTS: A Field Guide for Science & Technology Studies*. Princeton University Press.

Senra, Stella. 2011. «Conversações em Watoriki». *Cadernos de Subjetividade* (13):55–77. doi: [10.2354/cs.v0i13.38464](https://doi.org/10.2354/cs.v0i13.38464).

Shannon, C. E. 1948. «A Mathematical Theory of Communication». *Bell System Technical Journal* 27(3):379–423. doi: [10.1002/j.1538-7305.1948.tb01338.x](https://doi.org/10.1002/j.1538-7305.1948.tb01338.x).

Sherman, Stephanie. 2022. «The Polyopticon: A Diagram for Urban Artificial Intelligences». *AI & SOCIETY*. doi: [10.1007/s00146-022-01501-3](https://doi.org/10.1007/s00146-022-01501-3).

Silva, Mariah Rafaela, e Joana Varon. 2021. *Reconhecimento Facial No Setor Público e Identidades Trans: Tecnopolíticas de Controle e Ameaça à Diversidade de Gênero Em Suas Interseccionalidades de Raça, Classe e Território*.

Silva, Tarcízio. 2022. *Racismo Algorítmico: Inteligência Artificial e Discriminação Nas Redes Digitais*. Edições Sesc SP.

Silva, Tarcízio. 2023. «Mapeamento de Danos e Discriminação Algorítmica». Obtido 13 de novembro de 2024 (<https://desvelar.org/casos-de-discriminacao-algoritmica/>).

Simondon, Gilbert. 2005. *L'invention Dans Les Techniques: Cours et Conférences*. editado por J.-Y. Chateau. Paris: Seuil.

Simondon, Gilbert. 2008. *Imagination et invention, 1965-1966*. editado por N. Simondon e J.-Y. Chateau. Chatou: les Éd. de la Transparence.

Simondon, Gilbert. 2020a. «A amplificação nos processos de informação». *Trans/Form/Ação* 43(1):283–300. doi: [10.1590/0101-3173.2020.v43n1.16.p283](https://doi.org/10.1590/0101-3173.2020.v43n1.16.p283).

- Simondon, Gilbert. 2020b. *A Individuação à Luz Das Noções de Forma e de Informação*. 1.^a ed. São Paulo: Editora 34.
- Simondon, Gilbert. 2020c. *Do Modo de Existência Dos Objetos Técnicos*. 1.^a ed. Rio de Janeiro: Editora Contraponto.
- Solagna, Fabricio. 2024. «Colonialismo Digital na África: os cabos de Internet traçam mesmas rotas dos navios negreiros». Obtido 28 de novembro de 2024 (<https://oplanob.com/colonialismo-digital-na-africa-os-cabos-de-internet-tracam-mesmas-rotas-dos-navios-negreiros/>).
- Souza, Joyce, Sérgio Amadeu da Silveira, e Rodolfo Avelino. 2021. *A Sociedade de Controle: Manipulação e Modulação Nas Redes Digitais*. 2.^a ed. São Paulo: Hedra.
- Sterling, Bruce. 1994. *The Hacker Crackdown: Law and Disorder on the Electronic Frontier*. Library Freeware.
- Thatcher, Jim, David O’Sullivan, e Dillon Mahmoudi. 2016. «Data Colonialism Through Accumulation by Dispossession: New Metaphors for Daily Data». *Environment and Planning D: Society and Space* 34(6):990–1006. doi: [10.1177/0263775816633195](https://doi.org/10.1177/0263775816633195).
- Tufekci, Zeynep. 2015. «Algorithmic Harms beyond Facebook and Google: Emergent Challenges of Computational Agency». *Colo. Tech. LJ* 13:203.
- Turing, A. M. 1937. «On Computable Numbers, with an Application to the Entscheidungsproblem». *Proceedings of the London Mathematical Society* s2-42(1):230–65. doi: [10.1112/plms/s2-42.1.230](https://doi.org/10.1112/plms/s2-42.1.230).
- Turing, Alan M. 1950. «Computing Machinery and Intelligence». *Mind* LIX(236):433–60. doi: [10.1093/mind/LIX.236.433](https://doi.org/10.1093/mind/LIX.236.433).
- USPTO. sem data-a. «General Information Concerning Patents». Obtido 27 de junho de 2022 (<https://www.uspto.gov/patents/basics/general-information-patents>).
- USPTO. sem data-b. «Patent Essentials». Obtido 18 de novembro de 2024 (<https://www.uspto.gov/patents/basics/essentials>).
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, e Illia Polosukhin. 2017. «Attention Is All You Need». *Advances in Neural Information Processing Systems* 30:5998–6008.
- Villares, Fábio. 2007. *Propriedade intelectual: tensões entre o capital e a sociedade*. São Paulo: Instituto de Estudos Econômicos e Internacionais : Paz e Terra.

- Viveiros de Castro, Eduardo. 2018. *Metafísicas Canibais: Elementos Para Uma Antropologia Pós-Estrutural*. São Paulo: Ubu Editora, n-1 edições.
- Von Zuben, Fernando, e Romis Attux. 2007. «IA353 - Redes Neurais - Turma A - 1o. Semestre de 2007». Obtido 3 de maio de 2022 (https://www.dca.fee.unicamp.br/~vonzuben/courses/ia353_1s07.html).
- Widrow, Bernard, e Marcian E. Hoff. 1960. «[Adaptive Switching Circuits](#)».
- Wiener, Norbert. [1948] 2007. *Cybernetics or Control and Communication in the Animal and the Machine*. 2. ed., reprint. Cambridge, MA, USA: MIT Press.
- WikiLeaks. 2015. «What Is WikiLeaks». Obtido 18 de novembro de 2024 (<https://wikileaks.org/What-is-WikiLeaks.html>).
- Zeiler, Matthew D., e Rob Fergus. 2014. «[Visualizing and Understanding Convolutional Networks](#)». Pp. 818–33 em *Computer Vision – ECCV 2014*, editado por D. Fleet, T. Pajdla, B. Schiele, e T. Tuytelaars. Cham: Springer International Publishing.
- Zuboff, Shoshana. 2018. «Big Other: Capitalismo de Vigilância e Perspectivas Para Uma Civilização de Informação». Pp. 17–68 em *Tecnopolíticas de Vigilância: Perspectivas Da Margem*, editado por F. Bruno, B. Cardoso, M. Kanashiro, e L. Guilhon. São Paulo: Boitempo.
- Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. First edition. New York: PublicAffairs.

Apêndices

Apêndice A

Lista de título e resumo das patentes

N	ID	Nome	Abstract
1	US11366812	Using live data streams and/or search queries to determine information about developing events	Techniques and a framework are described herein for gathering information about developing events from multiple live data streams and pushing new pieces of information to interested individuals as those pieces of information are learned. In various implementations, a plurality of live data streams may be monitored. Based on the monitoring, a data structure that models diffusion of information through a population may be generated and applied as input across a machine learning model to generate output. The output may be indicative of a likelihood of occurrence of a developing event and/or a predicted measure of relevancy of the developing event to a particular user. Based on a determination that the likelihood and/or measure of relevancy satisfies a criterion, one or more computing devices may render, as output, information about the developing event.
2	US11252114	Message suggestions	A method may involve, for each of one or more messages that are selected from a plurality of messages from an account: (a) extracting one or more phrases from a respective selected message; (b) determining that a conversation includes the respective selected message and one or more other messages from the plurality of messages; (c) generating a first feature vector based on the conversation, wherein the first feature vector includes one or more first features, wherein the one or more first features include one or more words from the conversation; and (d) generating, by a computing system, one or more training-data sets, wherein each training-data set comprises one of the phrases and the first feature vector. The method may further involve: training, by the computing system, a machine-learning application with at least a portion of the one or more training-data sets that are generated for the one or more selected messages; applying the trained machine-learning application to process an incoming message to the account; and responsive to applying the trained machine-learning application, determining one or more reply messages corresponding to the incoming message, wherein the one or more reply messages include at least one of the extracted one or more phrases.

N	ID	Nome	Abstract
3	US11222637	Performing subtask(s) for a predicted action in response to a separate user interaction with an automated assistant prior to performance of the predicted action	Implementations herein relate to pre-caching data, corresponding to predicted interactions between a user and an automated assistant, using data characterizing previous interactions between the user and the automated assistant. An interaction can be predicted based on details of a current interaction between the user and an automated assistant. One or more predicted interactions can be initialized, and/or any corresponding data pre-cached, prior to the user commanding the automated assistant in furtherance of the predicted interaction. Interaction predictions can be generated using a user-parameterized machine learning model, which can be used when processing input(s) that characterize a recent user interaction with the automated assistant. Should the user command the automated assistant in a way that is aligned with a pre-cached, predicted interaction, the automated assistant will exhibit instant fulfillment of the command, thereby eliminating any latency that the user would have otherwise experienced interacting with the automated assistant.
4	US11115695	Using machine learning and other models to determine a user preference to cancel a stream or download	A system and method are disclosed for training a machine learning model using information pertaining to transmissions of one or more media items to user devices associated with a user account. In an implementation, training data for the machine learning model includes first contextual information associated with a first user device and second contextual information associated with a second user device. The number of the transmissions to the user devices for the user account exceeds a threshold number of transmissions allowed for the user account. Training data further includes generating a first target output that identifies an indication of a preference of a user preference to keep or cancel each of the transmission. The method includes providing the training data to train the machine learning model. The trained machine learning model may be used to determine which of the new transmissions is to be canceled.
5	US11108889	Automatically resolving, with reduced user inputs, a set of activity instances for a group of users	Automatically determining, with reduced (or no) input from the users of a group, a set of activity instances that the group of users has interest in performing. A representation of the set of activity instances can be rendered for consideration by a group, and the set of activity instances can be determined even when only limited criteria are specified. Optionally, in response to affirmative user interface input(s) directed to a rendered representation of the set of activity instances, one or more of the activity instances of the set can be confirmed through limited input(s) of one or more users of the group. Further, the automatic determination of the set of activity instances is optionally performed using one or more trained machine learning models that are trained to optimize a likelihood that the users of the group will find the set satisfactory.

N	ID	Nome	Abstract
6	US10970096	Suggesting actions based on machine learning	This document describes techniques for suggesting actions based on machine learning. These techniques determine a task that a user desires to perform, and presents a user interface through which to perform the task. To determine this task, the techniques can analyze content displayed on the user device or analyze contexts of the user and user device. With this determined task, the techniques determine an action that may assist the user in performing the task. This action is further determined to be performable through analysis of functionalities of an application, which may or may not be executing or installed on the user device. With some subset of the application's functionalities determined, the techniques presents the subset of functionalities via the user interface. By so doing, the techniques enable a user to complete a task more easily, quickly, or using fewer computing resources.
7	US9817869	Systems and methods for prioritizing notifications on mobile devices	Certain embodiments of the disclosed technology include systems and methods for determining the priority of a notification on a mobile device using machine learning. Other aspects of the disclosed technology include selectively displaying notifications based on the priority of a notification. According to an embodiment of the disclosed technology, a computer-implement method is provided that comprises outputting, to a display device operatively coupled to a mobile device, a plurality of notifications, wherein each respective notification from the plurality of notifications is associated with a respective priority score; modifying, by the mobile device, a ranking model based on a user input received responsive to a first notification from the plurality of notifications and a characteristic of a second notification from the plurality of notifications; determining, by the mobile device, a priority score associated with a third notification based on the modified ranking model; and outputting, to the display device, the third notification based on the priority score associated with the third notification, wherein the third notification is graphically emphasized responsive to the priority score associated with the third notification being greater than at least one respective priority score associated with a corresponding respective notification from the plurality of notifications.
8	US9224100	Method and apparatus using accelerometer data to serve better ads	A mobile device, including an accelerometer and a processing unit receiving signals from the accelerometer and determining motion information of the mobile device based on the received signals is disclosed. A content selector selects a media format for content based on the physical state of the mobile device as indicated by the signals from the accelerometer. A display device displays the content in the selected media format. When the motion is a predetermined type of motion, the received content is withheld from being displayed by the display device. A machine learning apparatus is provided as the content selector.

N	ID	Nome	Abstract
9	US8645390	Reordering search query results in accordance with search context specific predicted performance functions	A method of processing a search query includes, for each search context of a plurality of search contexts, for each scoring primitive of a plurality of scoring primitives, and for a set of previously executed search queries that are consistent with the search context, determining a correlation between the scoring primitive and actual user selections of results of the previously executed search queries by a plurality of users. For each search context, machine learning is performed on the correlations to identify a predicted performance function comprising a weighted subset of the scoring primitives that meet predefined predictive quality criteria. Executing a user submitted search query includes associating the user submitted search query with a respective search context, and ordering at least a portion of the search results in accordance with the predicted performance function for the search context for the user submitted search query.

Apêndice B

Breve cronologia da IA¹

- 1943 • McCulloch e Pitts propõem um modelo artificial de neurônio para a resolução de cálculo lógico.
- 1946 • Início das Conferências Macy, série de 160 conferências que fundou a cibernética.
- 1948 • Wiener publica *Cibernética* e Shannon, o artigo que funda a teoria da informação.
- 1949 • Hebb propõe a ideia de plasticidade sináptica.
- 1950 • Turing publica artigo que define o que conhecemos como "teste de Turing".
- 1952 • Hayek publica *The Sensory Order*.
- 1953 • Fim das Conferências Macy.
- 1955 • McCarthy e colaboradores publicam a proposta de pesquisa em Dartmouth que cunha o termo "inteligência artificial".
- 1956 • Dartmouth Summer Research Project on Artificial Intelligence.
- 1958 • Rosenblatt propõe a arquitetura Perceptron, um modelo de múltiplas camadas de neurônio.
- 1959 • Samuel publica artigo que consolida o termo "aprendizado de máquina".
- 1960 • Widrow e Hoff propõem o ADALINE, um método automatizado de aprendizagem.
- 1962 • Rosenblatt publica *Principles of Neurodynamics*.
- 1963 • Vapnik e Chervonenkis propõe o algoritmo de máquinas de vetores de suporte.

- 1969 • Minsky e Papert publicam *Perceptrons*.

- 1986 • Rumelhart, Hinton e Williams popularizam o método de retropropagação de erros para cálculo de derivadas.

- 1997 • Hochreiter e Schmidhuber propõem o neurônio LSTM utilizado em redes recorrentes. Mitchell propõe uma definição avaliável de aprendizado.

- 2003 • Bengio *et al.* propõem um MdL probabilístico.

- 2012 • A arquitetura AlexNet proposta por Krizhevsky, Sutskever e Hinton vence a competição da ImageNet.
- 2013 • Mikolov *et al.* propõem o método word2vec para representação vetorial de palavras.
- 2014 • Szegedy *et al.* propõem a rede convolucional Inception, Sutskever *et al.*, a rede recursiva seq2seq e Goodfellow, a GAN.
- 2015 • Lecun, Bengio e Hinton publicam o artigo "Aprendizado profundo" na revista *Nature*.
- 2016 • Goodfellow, Bengio e Courville publicam o livro *Aprendizado profundo*.
- 2017 • Vaswani *et al.* propõem a arquitetura transformer.

¹Alguns eventos mencionados na dissertação, compilado principalmente de Goodfellow et al. (2016) e Pasquinelli (2023).

Índice completo de conteúdo

Resumos	iv
Lista de Figuras	vii
Lista de Tabelas	x
Lista de Abreviaturas e Siglas	xi
Apresentação	1
0.1 Sobre o autor: explicitação de uma perspectiva parcial	3
0.2 Posição inicial do problema	5
0.3 Argumento da dissertação	8
I Introdução: tecnociência e capitalismo contemporâneo	9
1 A Google como empresa e como tecnologia	10
1.1 Alguns aspectos sociais e econômicos	11
1.1.1 O império das mentes na nova era digital: Assange encontra Jared Cohen e Eric Schmidt	12
1.1.2 Gerando excedente econômico e político: Zuboff (e Fuchs) encontra Hal Varian	14
1.2 Do PageRank™ ao Google Search: tecnologia e modulação de atividade	17
2 Propriedade intelectual, sociologia de patentes e os algoritmos de modulação de atividade	24
2.1 Patentes, informação, software e algoritmos	25
2.2 Inteligência artificial, aprendizado de máquina e a propriedade intelectual de máquinas de apropriação intelectual	27
2.2.1 Inteligência artificial	27
2.2.2 Aprendizado de máquina e redes neurais artificiais	29
2.2.3 Aprendizado de máquina como apropriação intelectual	33
2.3 Uma sociologia da tecnologia das patentes da Google	34
2.3.1 Dificuldades da pesquisa sobre algoritmos	35
2.3.2 Análise qualitativa de patentes como método para a investigação de algoritmos	36
2.3.3 Apresentação das patentes da Google sobre técnicas de modulação de atividade utilizando aprendizado de máquina	39

US11366812 - Using live data streams and/or search queries to determine information about developing events	41
US11252114 - Message suggestions	41
US11222637 - Performing subtask(s) for a predicted action in response to a separate user interaction with an automated assistant prior to performance of the predicted action	41
US11115695 - Using machine learning and other models to determine a user preference to cancel a stream or download	41
US11108889 - Automatically resolving, with reduced user inputs, a set of activity instances for a group of users	42
US10970096 - Suggesting actions based on machine learning	42
US9817869 - Systems and methods for prioritizing notifications on mobile devices	42
US9224100 - Method and apparatus using accelerometer data to serve better ads	42
US8645390 - Reordering search query results in accordance with search context specific predicted performance functions	43
II Uma cartografia do aprendizado de máquina	44
3 Captura	45
3.1 Vigilância e captura	48
3.1.1 Panóptico e pós-panóptico	48
3.1.2 Colonialismo digital, colonialismo de dados e novos cercamentos	49
3.1.3 Da vigilância à captura	53
3.2 A informação digital e o aprendizado de máquina como aparelho de captura	57
3.2.1 Informação e informação digital	57
3.2.2 Aparelho de captura e aprendizado de máquina	59
3.2.3 Exposição de um aparelho de captura de informação digital	61
3.3 O processo de captura no aprendizado de máquina	65
4 Imaginação	71
4.1 Imagens: entre representação e performatividade	74
4.1.1 O caráter social das imagens	74
4.1.2 As imagens técnicas	76
4.1.3 Imagens operativas, pós-fotográficas e máquina-para-máquina	79
4.2 O filo maquinico e os agenciamentos de representação: uma genealogia do aprendizado de máquina	81
4.2.1 Renascimento	83
4.2.2 Fotografia	86
4.2.3 Do computador digital à inteligência artificial	88
4.3 Aprendizado de máquina e espaço latente: imagem, representação e performatividade para além da visualidade	92
4.3.1 Exemplo na patente	98
4.4 O processo de imaginação no aprendizado de máquina	101

5	Modulação	106
5.1	Molde e modulação, disciplina e controle	106
5.2	A modulação nos sistemas digitais	111
5.2.1	Exemplo na patente	115
5.3	O processo de modulação no aprendizado de máquina	117
5.4	Uma cartografia do aprendizado de máquina	120
	Mapa	123
	Legenda	124
	Informação de entrada	124
	Base de dados proprietária	124
	Modelo	124
	Imagem	125
	Informação de saída	125
	Aprendizado de máquina	126
III	Mediação algorítmica	127
6	Do aprendizado de máquina sobre a ação social à ação social do aprendizado de máquina	128
6.1	A sociedade automática	129
6.2	A questão do viés	130
6.3	Limites do termo viés	131
6.3.1	Invisibilização, hipervisibilização, justiça criminal e racismo	132
6.3.2	Tons de gênero, binarismo e sexismo	134
7	O que é um algoritmo? por uma definição de mediação algorítmica	137
7.1	Computador digital, sistema de computação e algoritmo	139
7.1.1	Software e hardware	140
7.1.2	O computador digital como um indivíduo técnico	144
7.1.3	Elementos eletrônicos, computadores digitais e conjuntos algorítmicos	148
7.2	Algoritmo e agência	149
7.2.1	Mediação técnica e mediação algorítmica	150
8	Mediações algorítmicas na aprendizagem de máquina	153
8.1	Laço de realimentação conservador	153
8.2	Estereotipia	155
8.3	Racismo e sexismo algorítmico	158
8.3.1	Racismo algorítmico	159
8.3.2	Sexismo algorítmico	161
	Considerações finais	166
	Resistências, desvios e alternativas	169
	Referências	173

Apêndices	192
A Lista de título e resumo das patentes	192
B Breve cronologia da IA	196
Índice completo de conteúdo	197