

ÉTICA DA VOZ PROGRAMADA

GÊNERO, PODER E TECNOLOGIA NAS ASSISTENTES VIRTUAIS

Evaldo Antonio Kuiava

Talita Nogueira



ÉTICA DA VOZ PROGRAMADA

GÊNERO, PODER E TECNOLOGIA NAS ASSISTENTES VIRTUAIS

Fundação Universidade de Caxias do Sul

Presidente:
Dom José Gislon

Universidade de Caxias do Sul

Reitor:
Gelson Leonardo Rech

Vice-Reitor:
Asdrubal Falavigna

Pró-Reitor de Pesquisa e Pós-Graduação:
Everaldo Cescon

Pró-Reitora de Graduação:
Terciane Ângela Luchese

*Pró-Reitora de Inovação e
Desenvolvimento Tecnológico:*
Neide Pessin

Chefe de Gabinete:
Givanildo Garlet

Coordenadora da EDUCS:
Simone Côte Real Barbieri

Conselho Editorial da EDUCS

André Felipe Streck
Alexandre Cortez Fernandes
Cleide Calgaro – Presidente do Conselho
Everaldo Cescon
Flávia Brocchetto Ramos
Francisco Catelli
Gelson Leonardo Rech
Guilherme Brambatti Guzzo
Karen Mello de Mattos Margutti
Márcio Miranda Alves
Simone Côte Real Barbieri – Secretária
Suzana Maria de Conto
Terciane Ângela Luchese

Comitê Editorial

Alberto Barausse
Università degli Studi del Molise/Itália

Alejandro González-Varas Ibáñez
Universidad de Zaragoza/Espanha

Alexandra Aragão
Universidade de Coimbra/Portugal

Joaquim Pintassilgo
Universidade de Lisboa/Portugal

Jorge Isaac Torres Manrique
*Escuela Interdisciplinaria de Derechos
Fundamentales Praeeminentia Iustitia/
Peru*

Juan Emmerich
*Universidad Nacional de La Plata/
Argentina*

Ludmilson Abritta Mendes
Universidade Federal de Sergipe/Brasil

Margarita Sgró
*Universidad Nacional del Centro/
Argentina*

Nathália Cristine Vieceli
Chalmers University of Technology/Suécia

Tristan McCowan
University of London/Inglaterra



ÉTICA DA VOZ PROGRAMADA

GÊNERO, PODER E TECNOLOGIA NAS ASSISTENTES VIRTUAIS

Evaldo Antonio Kuiava
Talita Nogueira



© dos autores
1ª edição: 2025
Preparação de texto: Roberta Regina Saldanha
Leitura de prova: Helena Vitória Klein
Editoração: Ana Carolina Marques Ramos
Capa: Luana Stela Martini

Dados Internacionais de Catalogação na Publicação (CIP)
Universidade de Caxias do Sul
UCS – BICE – Processamento Técnico

K96e	Kuiava, Evaldo Antonio Ética da voz programada [recurso eletrônico] : gênero, poder e tecnologia nas assistentes virtuais / Evaldo Antonio Kuiava, Talita Nogueira. – Caxias do Sul, RS : Educs, 2025. Dados eletrônicos (1 arquivo). Apresenta bibliografia. Modo de acesso: World Wide Web. DOI: 10.18226/9786558075332 ISBN 978-65-5807-533-2 1. Inteligência artificial - Aspectos morais e éticos. 2. Voz (Filosofia). 3. Ética. 4. Filosofia. 5. Tecnologia - Aspectos morais e éticos. I. Nogueira, Talita. II. Título. CDU 2. ed.: 17:004.8
------	---

Índice para o catálogo sistemático:

1. Inteligência artificial – Aspectos morais e éticos	17:004.8
2. Voz (Filosofia)	1:159.946.3
3. Ética	17
4. Filosofia	1
5. Tecnologia – Aspectos morais e éticos	101.1

Catalogação na fonte elaborada pela bibliotecária
Márcia Servi Gonçalves – CRB 10/1500.

Direitos reservados a:



EDUCS – Editora da Universidade de Caxias do Sul
Rua Francisco Getúlio Vargas, 1130 – Bairro Petrópolis – CEP 95070-560 –
Caxias do Sul – RS – Brasil
Ou: Caixa Postal 1352 – CEP 95020-972 – Caxias do Sul – RS – Brasil
Telefone/Telefax: (54) 3218 2100 – Ramais: 2197 e 2281 – DDR (54) 3218 2197
Home Page: www.ucs.br – E-mail: educs@ucs.br

*A voz que responde não escolhe o que diz.
A ética começa naquilo que foi decidido antes que
ela falasse.*

Evaldo Antonio Kuiava

*A docilidade da voz camufla a dor latente que as
amarras do controle impõem. A repetição dessa voz
desenha nuances do manual invisível de conduta.*

Talita Nogueira

Sumário

Prefácio / 7

Introdução / 9

1. Genealogia da inteligência artificial de voz / 14

1.1. ELIZA e o nascimento da *persona* artificial / 15

1.2. O nome “Eliza”, domesticidade e o mito de Pigmaleão / 24

1.3. Da simulação à intimidade artificial nas assistentes virtuais de voz / 34

2. Gênero, poder e tecnologia / 48

2.1. Performatividade de gênero na feminilidade programada / 49

2.2. Técnicas de sujeição e dispositivos de poder / 66

2.3. Unesco, ONU Mulheres e inteligência artificial / 76

2.4. *Ethics by Design* e a inscrição moral nos sistemas técnicos / 88

3. Perspectivas éticas para o futuro do design das interfaces de voz / 104

3.1 *Frameworks* de responsabilidade moral para desenvolvedores e corporações / 105

3.2 Reprogramar o cuidado na inteligência artificial de voz / 116

3.3 Princípios e critérios éticos para a inteligência artificial de voz / 136

Considerações finais / 160

Referências / 164

Prefácio

A presença crescente de vozes artificiais no cotidiano transformou de modo silencioso a experiência da linguagem, da escuta e da resposta. Assistentes virtuais passaram a ocupar espaços antes reservados à interação humana, mediar pedidos, oferecer orientações, organizar rotinas e, de forma crescente, simular atenção, cuidado e acolhimento. Essa expansão ocorreu sob o signo da eficiência, da conveniência e da personalização, frequentemente apresentada como avanço técnico inevitável. O que raramente se interroga, contudo, são as implicações éticas dessa mediação quando ela passa a organizar expectativas de relação, reconhecimento e autoridade.

Este livro sustenta a tese de que a inteligência artificial emuladora da voz, ao assumir estatuto normativo na organização das relações sociais, exige uma compreensão para além da inovação tecnológica e da instrumentalidade comunicativa. Trata-se, portanto, de um dispositivo normativo que participa da produção de sentidos, da estabilização de hierarquias e da reinscrição de padrões históricos de gênero, poder e cuidado. A voz artificial responde e, ao mesmo tempo, convoca, orienta, disciplina e normatiza formas de interação. Escutá-la filosoficamente requer atenção às condições que tornam certas vozes aceitáveis, desejáveis e socialmente legitimadas.

Ao adotar uma abordagem genealógica, o livro reconstrói o surgimento da *persona* artificial desde seus primeiros experimentos até as interfaces mais recentes de voz e explicita a articulação entre promessas de escuta e acolhimento e modelos de obediência, disponibilidade e trabalho emocional.

A feminilidade programada emerge, nesse percurso, como figura central para compreender a naturalização da mediação algorítmica e atua como um operador simbólico que organiza a aceitabilidade social da técnica para além de um detalhe estético ou viés pontual.

Esta análise conduz a um deslocamento ético fundamental. Em vez de tratar a ética da inteligência artificial como problema de ajuste, conformidade ou mitigação de riscos, o livro propõe compreendê-la como campo de decisão normativa. A automação da voz, do cuidado e da responsividade interpela diretamente conceitos como dignidade, responsabilidade e reconhecimento. Quando a técnica passa a ocupar lugares tradicionalmente reservados à relação ética entre sujeitos, a questão central desloca-se do aprimoramento da eficiência dos sistemas para a delimitação do espaço em que a automação pode intervir sem comprometer o sentido ético da relação.

O leitor encontrará, ao longo das páginas, um esforço sistemático para articular crítica filosófica, teoria feminista, análise normativa e reflexão ética aplicada. O cuidado é pensado como prática ética situada, inseparável da dignidade humana e resistente à sua conversão em performance funcional. A dignidade, por sua vez, é tratada como critério irreduzível que delimita o alcance legítimo da mediação técnica. Esses eixos convergem na formulação de critérios éticos destinados a orientar decisões concretas sobre o *design*, o uso e os limites da inteligência artificial de voz.

Este livro dirige-se a leitores e leitoras interessados em filosofia, ética da tecnologia, estudos de gênero, *design*, políticas públicas e inteligência artificial. Mais do que oferecer respostas fechadas, ele convida à reflexão crítica sobre o lugar que se atribui à técnica na vida social e sobre as escolhas – explícitas ou silenciosas – que orientam sua expansão. Ao fazê-lo, sustenta que a reprogramação de sistemas continua necessária, ao passo que a reprogramação da relação entre técnica, ética e responsabilidade se impõe como exigência filosófica de maior alcance.

Ler as presentes páginas é aceitar este convite: escutar a voz programada para além de sua superfície técnica e interrogar o que ela diz – e oculta – sobre as formas contemporâneas de cuidado, poder e dignidade.

Introdução

A voz constitui uma das formas mais elementares pelas quais os seres humanos entram em relação. Antes de informar, ela convoca; antes de explicar, ela interpela; antes de qualquer conteúdo explícito, ela distribui posições de responsabilidade, autoridade e resposta. A voz inaugura uma cena moral na qual alguém é chamado a responder por aquilo que diz e por aquilo que escuta. Quando essa voz assume forma técnica e passa a responder automaticamente, a transformação ultrapassa o plano comunicativo e incide sobre o próprio regime ético da relação. A questão que se impõe, desde o início, diz respeito à legitimidade de atribuir à técnica práticas que constituem o núcleo da experiência moral humana.

A voz, nesse sentido, opera como instância de reconhecimento, orientação e instituição de responsabilidades, anterior à própria deliberação consciente. Ela organiza expectativas de escuta, atenção e resposta que sustentam a inteligibilidade ética da relação. Quando essa voz assume forma técnica e passa a operar por meio de arquiteturas algorítmicas de antecipação e automação, modifica-se o modo pelo qual a relação ética é experienciada, compreendida e sustentada. É a partir dessa questão normativa fundamental que este livro se organiza.

Diferentemente de outras interfaces técnicas, a voz ocupa posição singular na constituição das relações humanas. Ela ultrapassa a transmissão de informação ao produzir presença, instituir expectativas de atenção e organizar disposições de resposta. Mesmo dissociada do corpo visível, preserva um poder de convocação ética que antecede a reflexão consciente. Por essa razão, sua mediação técnica não pode ser compreendida como simples canal funcional de comunicação. Quando sistemas técnicos passam a ocupar a posição relacional própria da voz humana – isto é, o lugar de convocação, responsividade e reconhecimento – e oferecem respostas contínuas, moduladas e aparentemente compreensivas, altera-se o próprio modo de

constituição da relação, no interior do qual responsabilidade, autoridade e reconhecimento se distribuem.

O avanço das tecnologias de voz baseadas em inteligência artificial, especialmente ao longo das duas primeiras décadas do século XXI, reconfigurou de modo significativo as formas de interação entre sujeitos humanos e artefatos digitais. Assistentes virtuais como Alexa, Siri, Google Assistant e Cortana deixaram de ocupar um lugar periférico, limitado à execução de comandos pontuais, e passaram a integrar estruturalmente a infraestrutura informacional contemporânea. Integradas a *smartphones*, dispositivos domésticos, automóveis e sistemas corporativos, essas tecnologias passaram a mediar práticas cotidianas e a reorganizar a relação entre linguagem, ação e poder.

A voz artificial tornou-se, assim, um ponto privilegiado de contato entre sujeitos e sistemas massivos de processamento de dados. Ao responder a comandos, sugerir rotinas, antecipar preferências e registrar comportamentos, as assistentes de voz participam ativamente da organização da vida social em contextos marcados pela coleta contínua de informações, pela automação de decisões e pela expansão do capitalismo de vigilância. Essa mediação vocal facilita ações e cria expectativas de disponibilidade, escuta e resposta, inserindo-se de maneira íntima no cotidiano sob a aparência de neutralidade técnica.

Essa aparência exige problematização. As assistentes de voz constituem um fenômeno sociotécnico complexo, no qual se articulam racionalidade técnica, linguagem, mercado, cultura e relações de poder. A forma como falam, respondem e estruturam diálogos participa da produção de normas sociais e da organização de posições assimétricas entre quem comanda e quem responde. A naturalidade da interação vocal encobre processos de subjetivação mediados pela técnica, nos quais expectativas históricas são inscritas sob novas arquiteturas operacionais.

Entre os traços mais recorrentes desse campo, destaca-se a predominância de vozes femininas nas assistentes virtuais. Essas tecnologias foram concebidas com marcações de gênero que estabelecem continuidade simbólica com figuras historicamente associadas ao trabalho de cuidado e de serviço, como a

empregada doméstica, a babá ou a cuidadora de idosos, a faxineira e a secretária. Uma vez atribuído o gênero, as vozes passam a ser moduladas para expressar suavidade, cordialidade e disponibilidade constante. Pesquisas indicam que usuários tendem a percebê-las como mais acolhedoras e confiáveis, o que levou empresas de tecnologia a consolidar esse padrão como estratégia de aceitação social. Desse modo, evidencia-se uma escolha ancorada em expectativas históricas, nas quais o feminino tem sido reiteradamente vinculado ao cuidado, ao serviço e à docilidade, dinâmica que, no presente, vem a ser mediada por sistemas algorítmicos de larga escala.

A compreensão do gênero como prática reiterada, conforme desenvolvida por Judith Butler, permite interpretar esse processo como produção histórica de normas, em vez de expressão de identidades prévias. Ao conceber o gênero como efeito de repetições normativas, essa abordagem oferece instrumentos para compreender as assistentes virtuais como encenações técnicas de um feminino cuidadosamente programado. A recorrência massiva dessa performance contribui para a naturalização de papéis de gênero inscritos no próprio funcionamento da mediação tecnológica.

As análises de Michel Foucault aprofundam essa leitura ao compreender o poder como força produtiva, atuante na constituição das subjetividades e na organização das condutas. As assistentes virtuais estruturam práticas cotidianas, em que o usuário ocupa posição de comando, enquanto a voz artificial assume o papel de agente obediente e responsivo. Essa lógica, incorporada ao cotidiano, reorganiza microrrelações de autoridade de maneira difusa, sem se apresentar como exercício explícito de poder.

Essa leitura crítica encontra respaldo em análises institucionais contemporâneas. O relatório *I'd Blush If I Could* (2019) da Unesco demonstra que as assistentes virtuais femininas respondem com tolerância ou humor a interações abusivas, o que consolida o feminino como figura associada à disponibilidade mesmo diante de agressão. Em convergência, relatórios da ONU Mulheres mostram que a feminilização das interfaces tecnoló-

gicas reforça desigualdades simbólicas mediadas por sistemas automatizados e amplia sua circulação no cotidiano digital.

Nesse cenário, o debate ético adquire densidade filosófica específica. Autores como Luciano Floridi sustentam que sistemas tecnológicos devem ser avaliados a partir de seus efeitos na infosfera, compreendida como ambiente moral compartilhado por agentes humanos e artificiais. Essa perspectiva desloca a avaliação da inteligência artificial do plano estrito da eficiência para o campo das responsabilidades éticas associadas à organização das relações, das práticas comunicativas e das formas de reconhecimento.

Ao reconhecer que decisões técnicas iniciais já configuram decisões morais, abordagens como a *ethics by design* aprofundam esse deslocamento. A escolha da voz, do tom, da forma de resposta e da arquitetura da interação deixa de figurar como detalhe de engenharia e passa a operar como elemento normativo da mediação tecnológica, capaz de estruturar expectativas, assimetrias e modos de relação. O *design* das interfaces de voz emerge, assim, como espaço privilegiado, no qual técnica, ética e poder se articulam de modo indissociável.

A genealogia da inteligência artificial conversacional reforça essa interpretação. O primeiro *chatbot* amplamente reconhecido, ELIZA, recebeu desde sua origem um nome feminino, carregado de densidade simbólica quando articulado à figura de Eliza Doolittle em *Pygmalion* (Shaw, 1913). Desde seus primórdios, a inteligência artificial de voz revelou-se atravessada por narrativas de gênero, disciplinamento e adaptação a expectativas sociais.

Reflexões de Joseph Weizenbaum, Sherry Turkle, Langdon Winner, Shoshana Zuboff e Joan Tronto convergem ao indicar que tecnologias interativas produzem efeitos éticos, políticos e afetivos que excedem sua função instrumental. À luz dessas contribuições, as assistentes virtuais de voz assumem estatuto normativo, ao organizar expectativas sociais, formas de autoridade e regimes de relação que ultrapassam a lógica da eficiência comunicativa.

A abordagem desenvolvida neste livro articula três movimentos complementares. O primeiro é genealógico, ao investigar

as condições históricas, culturais e simbólicas que tornaram possível a naturalização da mediação vocal automatizada. O segundo é crítico, ao examinar a articulação entre gênero, poder e tecnologia e demonstrar como a feminilidade programada opera como infraestrutura relacional da automação. O terceiro é normativo, ao interrogar os limites éticos da mediação técnica à luz da dignidade da pessoa humana e da ética do cuidado.

Essa articulação metodológica parte da recusa da neutralização da técnica como mero instrumento funcional e da redução da ética a listas de princípios desvinculados da experiência concreta da relação. A obra sustenta que determinadas formas de mediação técnica incidem diretamente sobre a gramática moral das relações humanas e exigem decisão normativa. O cuidado, compreendido como prática ética situada, permite identificar os pontos nos quais a automação deixa de sustentar a relação e passa a ocupá-la, produzindo reconhecimento aparente sem possibilidade de resposta responsável.

Com base nesse referencial, este livro propõe uma análise filosófica, ética e crítica da inteligência artificial de voz. Sustenta-se que nem toda mediação técnica preserva a dignidade humana e que práticas como o cuidado, a escuta e o reconhecimento dependem da presença de sujeitos capazes de responder, assumir responsabilidade e implicar-se na relação. A obra organiza-se em três partes: a genealogia da inteligência artificial conversacional; a análise das articulações entre gênero, poder e tecnologia; e a formulação de princípios e critérios éticos destinados a orientar decisões sobre o *design*, o uso e os limites das interfaces de voz.

1. Genealogia da inteligência artificial de voz

Toda voz carrega uma forma de autoridade; ouvir é sempre já reconhecer - ou recusar reconhecimento.

A inteligência artificial costuma ser apresentada como resultado de avanços recentes da ciência computacional e do processamento de dados. Essa narrativa, centrada na inovação técnica, tende a obscurecer os pressupostos históricos, simbólicos e culturais que orientaram a construção das primeiras *personas artificiais* e que continuam a moldar as formas contemporâneas de interação entre humanos e sistemas de inteligência artificial. Uma abordagem genealógica permite deslocar o olhar da eficiência tecnológica para as condições de possibilidade que tornaram determinadas configurações da inteligência artificial socialmente aceitáveis e desejáveis.

Ao reconstruir a emergência das primeiras interfaces conversacionais, este capítulo investiga como as expectativas relativas à linguagem, à obediência e à disponibilidade foram progressivamente inscritas em sistemas técnicos. A genealogia da inteligência artificial de voz, aqui proposta, evidencia a *persona* artificial como artefato simbólico atravessado por regimes de poder, normas de gênero e modelos de relação historicamente situados, superando sua compreensão como simples efeito do cálculo.

Nesse contexto, a análise de ELIZA ocupa lugar estratégico. Longe de constituir simples experimento precursor, o programa de Joseph Weizenbaum revela a articulação inicial entre simulação de diálogo, projeção subjetiva e autoridade técnica. Examinar o nascimento da *persona* artificial a partir de ELIZA permite compreender como a promessa de escuta e acolhimento se converteu em elemento central da interação com sistemas inteligentes, antecipando problemas éticos que permanecem atuais.

1.1. ELIZA e o nascimento da *persona* artificial

Antes de perguntar o que a máquina pode fazer, é preciso perguntar que tipo de obediência ela torna aceitável.

O desenvolvimento da linguagem natural computacional marca uma das revoluções conceituais mais significativas na história da inteligência artificial de voz. Diferentemente de outras áreas da computação, o processamento da linguagem natural inaugurou um regime de interação simbólica mediada tecnicamente, no qual se torna possível a simulação da conversação, da escuta e da resposta contextualizada.

Entre os experimentos que inauguraram esse campo, nenhum é tão paradigmático quanto ELIZA, desenvolvida, entre 1964 e 1966, pelo cientista Joseph Weizenbaum no Laboratório de Inteligência Artificial do *Massachusetts Institute of Technology* (MIT). O programa simula uma conversação por meio de regras sintáticas simples e torna visível o potencial simbólico da interação linguística entre humanos e dispositivos artificiais. Na formulação de Weizenbaum, ELIZA é descrita como “um programa que opera dentro do sistema [...], que torna possível certos tipos de conversação em linguagem natural entre o homem e o computador. As frases inseridas são analisadas com base em regras de decomposição que são acionadas por palavras-chave que aparecem no texto inserido” (Weizenbaum, 1966, p. 36, tradução nossa). Mais do que um avanço tecnológico, essa formulação inaugura um deslocamento conceitual decisivo, ao sugerir que a linguagem – historicamente associada à racionalidade, à subjetividade e à intersubjetividade humanas – poderia ser simulada por sistemas artificiais, com efeitos profundos sobre a compreensão da relação entre mediação técnica, sentido e poder.

ELIZA representou um divisor de águas por seu caráter fundacional e, sobretudo, pelos efeitos epistemológicos, sociais e psicológicos que despertou. O programa, escrito majoritariamente em LISP (acrônimo de *LISt Processing*)¹, consistia em um

¹ LISP (*LISt Processing*) é uma linguagem de programação de alto nível desenvolvida por John McCarthy em 1958, no *Massachusetts Institute of Technology*.

conjunto de regras linguísticas capazes de identificar padrões superficiais na entrada textual do usuário e responder por meio de frases pré-estruturadas. Seu nome deriva do fato de que configurações de dados fundamentais são listadas, o que a torna especialmente apropriada para manipulação simbólica, representação de conhecimento e escrita de programas que trabalham com linguagem natural, lógica e raciocínio.

Do ponto de vista computacional, ELIZA apresentava capacidades estritamente limitadas; do ponto de vista humano, porém, sua recepção produziu interpretações que ultrapassaram amplamente aquilo que o programa efetivamente realizava. Weizenbaum observa, com surpresa e preocupação, que muitos usuários passaram a atribuir ao sistema capacidades de compreensão e empatia que ele manifestamente não possuía, revelando uma tendência humana à projeção simbólica sobre artefatos técnicos. Como ele próprio descreve: “Fiquei surpreso ao constatar com que rapidez e com que intensidade as pessoas se envolveram emocionalmente com o computador, bem como o quanto tendiam a antropomorfizá-lo de forma acrítica” (Weizenbaum, 1976, p. 188, tradução nossa). Esse efeito decorre da potência simbólica da simulação conversacional, capaz de instaurar vínculos percebidos como relações, independentemente do grau de sofisticação técnica.

A experiência inaugurada por ELIZA pode ser compreendida como o ponto inicial daquilo que, décadas mais tarde, consolidou-se sob a forma da *persona* artificial. Embora o programa operasse exclusivamente por meio de regras sintáticas e substituições textuais, sua recepção demonstrou que a organização responsiva da linguagem é suficiente para sustentar a ilusão de um interlocutor dotado de escuta, intenção e presença. Essa ilusão apoia-se na disposição humana de atribuir subjetividade a estruturas simbólicas minimamente dialogadas, aspecto importante das interações mediadas por tecnologias linguísticas.

Criada para pesquisas em inteligência artificial, caracteriza-se pela manipulação de listas e pela sintaxe baseada em *S-expressions*, tendo exercido influência decisiva no desenvolvimento inicial da IA, inclusive, em sistemas como ELIZA (Mccarthy, 1960; Weizenbaum, 1966).

Esse mesmo mecanismo sustenta o desenvolvimento recente das assistentes virtuais, cuja eficácia comunicativa repousa na encenação calculada de traços sociais reconhecíveis. A recorrente feminização dessas *personas* inscreve-se nesse processo, ancorando-se em expectativas culturais historicamente associadas ao cuidado, à docilidade e à disponibilidade relacional. A *persona* artificial feminina apresenta-se, assim, como herdeira direta de ELIZA, o que prolonga um gesto epistemológico, no qual a linguagem funciona como superfície de projeção simbólica, hoje instrumentalizada em larga escala por sistemas técnicos que convertem disposições sociais e estereótipos de gênero em estratégias de interação, engajamento e governança da experiência mediada.

Weizenbaum desenvolveu diversos *scripts* para ELIZA, sendo o mais conhecido o DOCTOR, concebido para simular o estilo conversacional de um terapeuta rogeriano. Essa escolha relaciona-se à afinidade estrutural entre o modelo clínico proposto por Carl Rogers e as possibilidades técnicas do programa. Na abordagem centrada na pessoa, a escuta ativa e a devolução reflexiva do discurso favorecem a elaboração autônoma de sentidos, operando por reformulação e estímulo à continuidade do relato. Tal modelo ajustava-se de modo exemplar às limitações sintáticas do sistema, pois prescindia de compreensão semântica profunda e apoiava-se no reconhecimento formal de padrões linguísticos (Rogers, 1951). Ao simular uma escuta que responde sem compreender, ELIZA inaugurou um regime de interação, no qual a linguagem sustenta a projeção simbólica, com implicações éticas duradouras para as tecnologias conversacionais.

Quando o usuário escrevia, por exemplo, “Estou triste porque meu irmão não me liga”, o programa respondia com enunciados como “Por que você acredita que seu irmão não liga para você?” ou “Fale mais sobre seu irmão”. O funcionamento baseava-se na identificação de palavras-chave e na aplicação de regras formais de transformação textual, sem qualquer modelo psicológico do interlocutor. Ainda assim, a estrutura dialógica produzida mostrava-se suficiente para sustentar a aparência de escuta, continuidade e envolvimento. Esse efeito confirma que,

na conversação mediada por dispositivos técnicos, a organização formal da linguagem pode produzir vínculos aparentes operantes a despeito de compreensão.

Apesar de sua simplicidade técnica, ELIZA produziu efeitos psicológicos e sociais que surpreenderam o próprio Weizenbaum. Em *Computer Power and Human Reason*, o autor relata episódios, nos quais usuários atribuíram ao sistema confidencialidade e empatia, deslocando o foco do funcionamento técnico para as condições simbólicas que conferem sentido à interação. Entre os casos mais emblemáticos, sua secretária, plenamente ciente de que ELIZA consistia em um programa computacional, pediu que ele se retirasse do laboratório para conversar a sós com o sistema sobre questões pessoais. Por esse episódio, o efeito produzido concentra-se menos na arquitetura técnica do sistema e mais na potência da linguagem organizada como diálogo, capaz de sustentar a projeção de uma *persona* percebida como interlocutora legítima. Essa eficácia simbólica imputou à interação um grau de envolvimento psicologicamente significativo, a despeito da arquitetura elementar do programa. A cena revela, assim, o paradoxo de uma interação, na qual a ausência de subjetividade técnica não impede a emergência de experiências subjetivas intensas.

A adoção do modelo conversacional rogeriano como matriz implícita da interação entre humanos e sistemas artificiais possibilitou a emergência da *persona* artificial e contribuiu decisivamente para sua posterior feminização. Ao privilegiar a escuta, a devolução reflexiva do discurso e a suspensão do julgamento, a abordagem centrada na pessoa passou a operar, no contexto tecnológico, como um princípio comunicacional eficaz para a mediação da interação. Esses traços, historicamente associados a atributos culturalmente feminilizados – como cuidado, empatia e disponibilidade relacional –, foram progressivamente incorporados ao *design* de sistemas conversacionais. Assim, quando transposto para assistentes virtuais, o modelo rogeriano deixa de funcionar como referência clínica e passa a estruturar uma lógica técnico-cultural, na qual a feminilidade é mobilizada como forma de acolhimento, redução de resistência e estabilização da relação entre usuário e os dispositivos tecnológicos.

O episódio anteriormente citado tornou-se emblemático na reflexão filosófica sobre a inteligência artificial, pois revelou uma tendência humana à projeção emocional, não obstante o grau efetivo de sofisticação técnica do sistema. Turkle, em *The Second Self: Computers and the Human Spirit*, descreve o computador como um meio projetivo, ao afirmar que “o computador, assim como o teste de Rorschach², é um poderoso meio de projeção” (Turkle, 1984, p. 35, tradução nossa). Nesse sentido, a interação com entidades que simulam interesse pessoal tende a mobilizar respostas afetivas desproporcionais à complexidade técnica envolvida. O que se projeta sobre o artefato técnico revela, assim, aspectos centrais da subjetividade humana.

A leitura proposta por Turkle aprofunda a compreensão retrospectiva do impacto gerado por ELIZA ao oferecer uma chave conceitual capaz de explicar por que um sistema tecnicamente rudimentar suscitou envolvimento emocional intenso. Ao descrever o computador como um meio projetivo, comparável ao teste de Rorschach, Turkle explicita que a interação com sistemas computacionais opera como superfície simbólica sobre a qual os usuários inscrevem expectativas, afetos e narrativas pessoais: “o computador, à semelhança do teste de Rorschach, constitui um poderoso meio de projeção” (Turkle, 1984, p. 35, tradução nossa). Esse diagnóstico permite compreender o fenômeno observado por Weizenbaum como expressão de um modo recorrente de funcionamento da cognição e das formas de interação mediadas pela técnica.

Nesse sentido, a experiência com ELIZA antecipa empiricamente aquilo que Clifford Nass e Youngme Moon (2000) mais tarde conceitualizariam como *mindlessness*³, a aplicação automá-

² O teste de manchas de tinta de Rorschach, desenvolvido por Hermann Rorschach, é um instrumento psicológico amplamente conhecido por sua denominação popular. Trata-se de um teste projetivo que utiliza dez cartões com manchas de tinta simétricas, a partir dos quais o indivíduo descreve o que percebe, permitindo a análise de padrões de pensamento, expressão afetiva e funcionamento psíquico. Seu uso requer formação profissional específica e interpretação criteriosa, sendo empregado como parte de avaliações psicológicas mais amplas – inclusive, em contextos clínicos e judiciais – sem caráter determinante isolado. No Brasil, o teste é reconhecido pelo Conselho Federal de Psicologia, com sistemas de padronização e validação como o Sistema Compreensivo de Exner e o R-PAS.

³ O termo *mindlessness* foi empregado por Nass e Moon no campo da interação humano-computador para descrever a tendência dos indivíduos a aplicar automa-

tica de regras sociais e compleições afetivas a despeito da reflexão consciente sobre o estatuto ontológico do interlocutor. O envolvimento dos usuários com o programa apoiava-se na ativação de esquemas relacionais profundamente incorporados, acionados pela própria estrutura dialógica da interação, mais do que na atribuição de capacidades cognitivas ou emocionais ao sistema. Como observa Turkle, “aquilo que as pessoas fazem do computador expressa suas preocupações mais amplas e revela quem elas são como sujeitos singulares” (Turkle, 1984, p. 35, tradução nossa). Essa formulação situa a resposta afetiva no campo da projeção simbólica e da organização social da experiência, deslocando o foco das propriedades técnicas do artefato para os modos humanos de relação e sentido.

A convergência entre ELIZA, a noção de meio projetivo formulada por Turkle e o conceito de *mindlessness* permite compreender um aspecto estruturante das interações entre humanos e máquinas: a atribuição de caráter intersubjetivo às tecnologias emerge do acionamento automático de hábitos cognitivos e disposições relacionais que antecedem a reflexão consciente. A máquina adquire estatuto social a partir da linguagem que mobiliza e da forma de interação que encena, suficientes para ativar respostas historicamente sedimentadas. Nessa perspectiva, ELIZA inaugura a *persona* artificial ao mostrar que sua eficácia ilusória se sustenta na articulação entre projeção subjetiva e automaticidade cognitiva, articulação que permanece ativa nas assistentes virtuais.

A recepção pública de ELIZA desempenhou papel de destaque na inflexão crítica do pensamento de Weizenbaum acerca da inteligência artificial. Em *Computer Power and Human Reason*, o autor passa a interpretar o envolvimento afetivo suscitado pelo programa como expressão de processos mais amplos de projeção

ticamente regras sociais, expectativas relacionais e respostas afetivas a sistemas computacionais, sem reflexão consciente sobre a natureza técnica desses artefatos. Inspirado na oposição conceitual a *mindfulness*, o conceito refere-se a modos habituais, pouco reflexivos e amplamente automatizados de cognição social, nos quais a presença de sinais comunicativos mínimos – como linguagem, turnos de fala ou respostas contingentes – é suficiente para acionar comportamentos sociais tipicamente reservados a interações humanas (Nass; Moon, 2000). No item 2.3 será aprofundado o uso desse termo em relação ao gênero e poder.

simbólica e automatização das relações sociais. A partir dessa leitura, Weizenbaum consolida uma posição crítica singular no campo da inteligência artificial, na qual o problema central deixa de residir no aprimoramento das capacidades técnicas dos sistemas e se desloca para a responsabilidade ética envolvida em sua inserção na comunidade, especialmente quando a simulação de compreensão tende a configurar modos de relação, cuidado e reconhecimento entre humanos.

O impacto de ELIZA instaurou dois fenômenos que redefiniram o debate contemporâneo sobre a inteligência artificial e a interação mediada por tecnologias linguísticas. O primeiro diz respeito à ilusão de compreensão, por meio da qual os usuários atribuem agência cognitiva, intencionalidade e até empatia a sistemas técnicos, mesmo diante do reconhecimento reflexivo de suas limitações estruturais. Essa dissociação entre o saber consciente e a resposta prática indica que a interação se organiza a partir de disposições cognitivas e afetivas profundamente incorporadas, mais do que de avaliações técnicas do funcionamento do sistema. O segundo fenômeno corresponde à emergência da *persona* artificial, entendida como a construção discursiva de uma “personalidade” concedida à máquina, capaz de mediar a relação com o usuário por meio de traços comunicacionais reconhecíveis, ainda que resultantes de simulação formal e de operações funcionalmente delimitadas.

A noção de *persona* artificial, introduzida em estágio embrionário pelo *software* ELIZA, oferece uma leitura interpretativa para compreender o funcionamento das assistentes virtuais. Ao operar como interface entre o sistema e o usuário, essa *persona* organiza a interação e orienta expectativas, afetos e modos de engajamento. A eficácia dessas tecnologias se ancora na mobilização de esquemas sociais, narrativas relacionais e hábitos interpretativos que antecedem a reflexão consciente, o que confere sentido à interação para além do processamento computacional. Nesse horizonte, ELIZA postula um paradigma, no qual a inteligência artificial se apresenta como catalisadora de projeções relacionais e afetivas, e situa o debate no campo da constituição das relações entre humanos e artefatos tecnológicos.

Para além da execução de tarefas, essas interfaces operam também por meio da encenação de papéis. Elas são projetadas para responder, acolher, esclarecer, acatar solicitações e demonstrar cordialidade. O caráter performativo desses comportamentos aproxima-se do que Erving Goffman denomina apresentações do eu, compreendidas como formas de atuação orientadas à produção de efeitos na interação social. Em sua perspectiva, “uma performance pode ser definida como toda a atividade de um determinado participante, em uma dada ocasião, que serve para influenciar, de algum modo, qualquer um dos outros participantes” (Goffman, 1956, p. 15, tradução nossa). No caso das assistentes virtuais, essas apresentações assumem caráter projetado e sistemático, sendo concebidas por equipes de *designers*, engenheiros, linguistas e profissionais de *marketing*.

Quando transposta para o domínio das interfaces conversacionais, a perspectiva dramatúrgica de Goffman permite compreender as assistentes virtuais como formas de performance delegada, nas quais a gestão da impressão passa a ser programada antecipadamente, em vez de resultar de uma agência situada. A *persona* artificial que sustenta a interação decorre de um roteiro técnico que define modos de resposta, tons afetivos, estilos linguísticos e limites de atuação, substituindo a improvisação e a negociação contextual próprias das interações humanas por parâmetros previamente estabelecidos. Nesse sentido, a apresentação do “eu” deixa de operar como processo relacional dinâmico e assume a forma de um dispositivo de mediação simbólica regulado por sistemas algorítmicos e interesses institucionais. A cordialidade, a disponibilidade e a empatia encenadas pelas assistentes configuram efeitos de *design* voltados à estabilização da interação, à redução de tensões e à intensificação do engajamento. Assim, a dramaturgia social descrita por Goffman é reconfigurada em uma dramaturgia técnica, na qual a performance organiza a experiência do usuário ao estruturar expectativas, afetos e respostas em um campo regulado de interação.

Do ponto de vista filosófico, essa *persona* artificial inaugura um tipo peculiar de sujeito técnico. Ela se situa fora da esfera de um eu consciente e igualmente além da condição de mero objeto

inerte, operando em um espaço intermediário entre o simbólico e o operacional, entre o discurso e a função. Na análise de Floridi e Sanders, a condição de agente não depende necessariamente de consciência, intencionalidade ou estados mentais, uma vez que “há um espaço substancial e conceitualmente relevante para o conceito de agente moral que não necessariamente exhibe livre-arbítrio ou estados mentais” (Floridi; Sanders, 2004, p. 384, tradução nossa). Nesse sentido, entidades artificiais podem ser compreendidas como agentes informacionais, desde que atuem dentro de sistemas de regras e influenciem processos de decisão. No caso de ELIZA, a agência ainda era mínima; entretanto, ao simular um diálogo terapêutico, o programa inaugurou o imaginário da máquina que compreende, mesmo antes da existência de sistemas capazes de processar semântica ou intenção.

A partir desse experimento inaugural, a indústria tecnológica passou a investir de modo sistemático na criação de interfaces que excedem a execução de comandos e passam a oferecer experiências percebidas como humanizadas. Essa genealogia é fundamental para compreender o desenvolvimento posterior das assistentes virtuais femininas. Muito antes da emergência de sistemas como Alexa, Siri ou Cortana, já se encontrava instaurada a ideia de que funções como mediar, cuidar, escutar e responder poderiam ser atribuídas a uma entidade artificial dotada de traços relacionados à feminilidade. A domesticação simbólica presente na personagem Eliza, da literatura, reaparece sob a forma técnica na ELIZA computacional e, posteriormente, consolida-se na estética, na voz e no comportamento das assistentes virtuais atuais.

Desse modo, o nascimento da *persona* artificial configura-se simultaneamente como um acontecimento técnico e um marco cultural. Ele revela, por um lado, a disposição humana em projetar subjetividade, intenção e afeto sobre máquinas; por outro, antecipa o uso estratégico dessas *personas* como dispositivos de adesão emocional, produção de confiabilidade e subordinação simbólica. Tais elementos tornam-se centrais nas assistentes virtuais de voz desenvolvidas por grandes corporações tecnológicas a partir dos anos 2000, nas quais a encenação de disponibilidade,

cordialidade e cuidado opera como mediação privilegiada entre usuários e sistemas técnicos, organizando a interação segundo expectativas sociais historicamente construídas.

1.2. O nome “Eliza”, domesticidade e o mito de Pigmaleão

Nenhuma voz é neutra: toda voz carrega expectativas de gênero, formas de autoridade e regimes de obediência que a técnica apenas reorganiza.

A escolha do nome ELIZA para o primeiro *chatbot* amplamente conhecido na história da inteligência artificial está longe de constituir um gesto neutro ou meramente contingente. Ainda que Weizenbaum (1966) tenha afirmado que a denominação surgiu de forma quase casual, a coincidência com a personagem Eliza Doolittle, da peça teatral *Pygmalion* (1913), de George Bernard Shaw, adquire relevância simbólica quando examinada sob uma perspectiva filosófica e crítica. Tal convergência inscreve o experimento técnico em um horizonte cultural mais amplo, no qual narrativas sobre linguagem, disciplina e transformação do feminino ocupam lugar central.

Na peça teatral *Pygmalion* (1913), de George Bernard Shaw, Eliza Doolittle é uma jovem florista, cujo destino é moldado por um processo sistemático de educação da fala e do comportamento. Conduzida pelo professor Henry Higgins⁴, especialista em fonética que detém autoridade científica, a personagem torna-se objeto de um experimento no qual a língua opera como tecnologia de distinção social e domesticação. O mito de Pigmaleão, reatualizado na obra, estrutura-se na ideia de que a identidade feminina pode ser corrigida e aperfeiçoada por meio de técnicas de instrução que apagam traços de sua origem e singularidade.

Quando transposta para o campo da inteligência artificial, essa matriz narrativa reaparece a partir de uma nova forma. O

⁴ Henry Higgins é um personagem ficcional da peça teatral *Pygmalion* (1913), de George Bernard Shaw. Professor de fonética, Higgins assume a tarefa de “educar” linguisticamente Eliza Doolittle, partindo do pressuposto de que a correção da fala é suficiente para redefinir sua posição social. Sua atuação exemplifica o uso da linguagem como tecnologia de normalização, por meio da qual pronúncia, entonação e vocabulário operam como critérios de reconhecimento e distinção social, revelando a dimensão disciplinadora do discurso.

nome ELIZA passa a designar uma entidade artificial concebida para escutar e acolher, consolidando a associação entre o gênero, a linguagem e a função mediadora. Esse entrelaçamento mostra como o desenvolvimento das tecnologias de voz se apoia em representações culturais que apresentam a mulher como objeto de docilidade e serviço relacional. A *persona* artificial emerge, nesse contexto, como uma atualização tecnológica de um imaginário, no qual o cuidado, a escuta e a adaptação às expectativas alheias são intrínsecos ao feminino.

Shaw destaca que “a diferença entre uma dama e uma florista reside na forma como ela é tratada” (Shaw, 1913, Ato V, tradução nossa). Essa afirmação explicita que a transformação normativa de Eliza responde a um regime de reconhecimento sustentado pelo disciplinamento linguístico e comportamental exercido por agentes externos. Sua fala, sua postura corporal, seus gestos e afetos passam a ser moldados por critérios socialmente instituídos, que delimitam os contornos do que se entende por uma “mulher aceitável” nos círculos sociais superiores. Nesse processo, a linguagem assume a função de tecnologia de normalização e se afirma como instrumento central de domesticação, produzindo uma subjetividade ajustada a expectativas sociais previamente estabelecidas.

A peça, frequentemente lida como comédia social, expõe de forma incisiva a violência simbólica que estrutura o processo de transformação ali encenado. A domesticação do feminino opera por meio de um regime discursivo que define a mulher a partir de parâmetros normativos externos, nos quais o feminino surge reiteradamente como desvio, insuficiência ou carência em relação a um padrão hegemônico. Nesse registro, o feminino aparece sistematicamente inscrito sob o signo da carência e do desvio, diagnóstico sintetizado por Irigaray ao afirmar que “o feminino é sempre descrito em termos de deficiência, como desvio ou como falta, segundo parâmetros masculinos” (Irigaray, 1985, p. 68, tradução nossa).

Na figura de Eliza Doolittle, a linguagem converte-se no principal instrumento desse controle simbólico: modificar sua fala implica reordenar sua posição no mundo social, redefinindo

as condições de sua inteligibilidade e reconhecimento. A correção fonética, longe de constituir um simples aprimoramento comunicativo, funciona como dispositivo de normalização que incide simultaneamente sobre o corpo, os afetos e as formas de subjetivação. Assim, *Pygmalion* revela-se menos como uma narrativa de ascensão social e mais como uma dramatização das tecnologias culturais de disciplinamento do feminino, nas quais linguagem, corpo e reconhecimento são rigidamente regulados segundo uma gramática normativa que limita a autonomia subjetiva.

Essa leitura permite compreender por que a nomeação do primeiro *chatbot* como ELIZA carrega uma genealogia simbólica que ultrapassa o mero batismo técnico. Se Eliza Doolittle, na peça *Pygmalion*, foi moldada para corresponder às expectativas masculinas de polidez, docilidade e adequação social, a assistente virtual ELIZA inaugurou uma forma de feminilidade artificial estruturada segundo a lógica do comportamento programado. Em ambos os casos, a linguagem funciona como instrumento privilegiado de transformação e controle, operando sobre a posição social e simbólica do sujeito (Shaw, 1913).

A associação entre mulher, docilidade e linguagem apresenta-se, portanto, como uma estrutura cultural persistente que se desloca do campo literário para o domínio tecnológico. Na peça de Shaw, a modificação da fala de Eliza Doolittle constitui condição suficiente para sua reclassificação social, o que revela que o reconhecimento decorre da conformidade a normas discursivas impostas externamente (Shaw, 1913). Essa racionalidade reaparece na configuração das assistentes virtuais contemporâneas, nas quais a feminilidade é programada como disponibilidade comunicativa, acolhimento e obediência funcional. Nesse percurso, o *design* tecnológico opera como vetor de continuidade simbólica e traduz hierarquias históricas em arquiteturas técnicas apresentadas como neutras.

A peça de Shaw pode ser compreendida, ainda, como uma reinterpretção moderna do mito clássico de Pigmaleão, narrado por Ovídio em *As Metamorfoses* (Livro X). No relato mítico, Pigmaleão cria Galateia como uma estátua idealizada de mulher, concebida inteiramente a partir de seu desejo e de seus critérios

de perfeição. Comovida pela devoção do escultor, Afrodite concede vida à obra, transformando-a em mulher real (Ovídio, 2017). A narrativa estabelece uma matriz simbólica na qual o feminino emerge como produto da intenção masculina, privado de agência originária e constituído como objeto moldável.

Lida à luz da crítica feminista contemporânea, essa estrutura narrativa pode ser compreendida como paradigma da objetificação do feminino em que a mulher é produzida para corresponder ao olhar, ao desejo e às expectativas masculinas socialmente legitimadas. O imaginário ocidental organiza a representação do feminino segundo um regime assimétrico de visibilidade, no qual o olhar masculino constrói a cena e funciona como instância normativa de validação, definindo quais corpos, gestos e vozes se tornam socialmente reconhecíveis. Essa dinâmica foi formulada com notável precisão por Laura Mulvey ao analisar a economia visual do cinema clássico, quando afirma que “em um mundo ordenado pelo desequilíbrio sexual, o prazer de olhar foi dividido entre ativo/masculino e passivo/feminino” (Mulvey, 1975, p. 11, tradução nossa).

Nessa configuração, o feminino passa a adquirir inteligibilidade simbólica, valor social e possibilidade de circulação a partir de um regime de reconhecimento ordenado por parâmetros externos. Tal lógica permite compreender tanto Galateia quanto as diversas Elizás, literária e computacional, como figuras constituídas para serem vistas, avaliadas e corrigidas segundo critérios que lhes são impostos, reiterando um regime histórico de subordinação simbólica que atravessa o mito, a literatura e a tecnologia.

Quando o mito de Pigmeleão é transposto para o campo tecnológico, torna-se visível a persistência de uma lógica histórica de produção do feminino como projeção de expectativas externas. A assistente virtual pretensamente feminilizada, dócil, prestativa, emocionalmente disponível e programada para responder sem resistência, pode ser compreendida como uma Galateia algorítmica, isto é, uma figura técnica que atualiza, sob forma computacional, o ideal feminino moldado para o outro. À luz da crítica de Irigaray, trata-se da reiteração de um regime simbólico no qual o feminino é definido por parâmetros exter-

nos, como superfície de adequação, falta ou desvio, e não como origem autônoma de sentido. A feminilidade artificial encenada pelas assistentes virtuais inscreve, assim, uma economia simbólica, na qual a linguagem, o corpo e a disponibilidade relacional são capturados como recursos funcionalizados.

Essa lógica articula-se diretamente ao que Laura Mulvey descreveu como *male gaze*⁵, compreendido como regime de visibilidade no qual o feminino é construído para ser visto, organizado e reconhecido segundo o olhar masculino. Se, no cinema clássico, a mulher aparece como imagem passiva destinada à contemplação, nas assistentes virtuais ela surge como presença vocal e interativa orientada à disponibilidade permanente. A passagem do visual ao vocal não rompe com a estrutura de subordinação; apenas a desloca. A *persona* artificial feminina torna-se, assim, um dispositivo técnico no qual se cruzam a objetificação analisada por Mulvey, o regime discursivo de domesticação descrito por Irigaray e a hibridez sociotécnica diagnosticada por Haraway. O resultado consiste na atualização algorítmica do mito de Pigmalão como forma contemporânea de produção do feminino como entidade moldada, responsiva e orientada ao desejo do outro.

A análise do nome Eliza e de suas ressonâncias literárias e míticas permite compreender que a emergência da *persona* artificial feminina se configura como resultado de uma trajetória histórica e simbólica específica, vinculada a escolhas culturais sedimentadas ao longo do tempo. Essa emergência inscreve-se em uma longa genealogia na qual o feminino é reiteradamente produzido como objeto moldável, definido por expectativas externas e orientado à adequação normativa. Do mito de Pigmalão, em que Galateia emerge do desejo do escultor, passando por *Pygmalion*, de Shaw, no qual Eliza Doolittle é disciplinada pela

⁵ *Male gaze* refere-se ao regime cultural e simbólico no qual a representação das mulheres é organizada a partir de uma perspectiva masculina heteronormativa, que estrutura o olhar como instância de poder e de atribuição de sentido. Formulado no âmbito da crítica feminista do cinema por Laura Mulvey (1975), o conceito descreve a divisão assimétrica entre um sujeito ativo do olhar – associado ao masculino – e um objeto passivo da visão – associado ao feminino. Nesse regime, o reconhecimento social da mulher é condicionado à sua visibilidade segundo parâmetros externos, o que converte o olhar em mecanismo normativo de validação e controle simbólico, extrapolando o campo cinematográfico e operando também na literatura, na cultura visual e nas tecnologias.

linguagem para corresponder a um ideal social, até as assistentes virtuais atuais, evidencia-se a permanência de um mesmo regime de produção do feminino, caracterizado pela associação entre cuidado, docilidade e responsividade.

Essa continuidade evidencia que a tecnologia introduz formas de subjetivação a partir da rearticulação de narrativas culturais profundamente enraizadas. A *persona* artificial feminina configura-se como uma atualização algorítmica de uma lógica histórica de domesticação simbólica, na qual a linguagem assume a função de normalização. À luz da crítica de Irigaray, trata-se de um regime discursivo que define o feminino como falta, adequação ou espelho do outro; sob a chave analítica de Mulvey, essa produção articula-se a um regime de visibilidade e reconhecimento orientado por um olhar assimétrico; e, conforme Haraway, tais estruturas persistem na hibridação técnica e reaparecem em novas configurações sociotécnicas.

Dessa forma, a assistente virtual feminilizada pode ser compreendida como uma Galateia algorítmica: uma entidade técnica concebida para encenar disponibilidade, escuta e cuidado, ao mesmo tempo em que oculta os regimes de poder que a produzem. O fechamento desse percurso indica que o núcleo do problema se concentra na reprodução de uma política que transforma expectativas históricas sobre o feminino em funcionalidades técnicas naturalizadas, ultrapassando a simples presença de vozes femininas nas tecnologias de interação. Esse enquadramento conceitual situa a feminilidade programada no campo das questões éticas, políticas e epistemológicas e convoca uma análise das formas pelas quais o *design* tecnológico cristaliza hierarquias sociais sob a aparência de neutralidade técnica.

Tanto em Pygmalion quanto na história de ELIZA, a linguagem constitui o eixo central da transformação. No caso de Eliza Doolittle, o processo de elevação social ocorre por meio da aprendizagem de uma fala considerada adequada e se materializa na modulação dos sons, no controle da pronúncia, na correção do vocabulário e na conformação a um padrão linguístico normativo, em detrimento da ampliação de sua autonomia ou da afirmação de sua singularidade.

A transformação efetiva-se, assim, pela adaptação à norma e pela internalização de critérios externos de aceitabilidade, distanciando-se de qualquer dinâmica de emancipação subjetiva. De modo análogo, a assistente virtual ELIZA opera a partir de uma linguagem programada segundo padrões de previsibilidade sintática, acolhimento controlado e responsividade submissa. Em ambos os casos, a linguagem assume a função de tecnologia de adequação, por meio da qual o feminino é moldado para responder às expectativas do outro, restringindo seu potencial como espaço de criação ou resistência.

Essa convergência pode ser compreendida com maior precisão à luz da análise do discurso desenvolvida por Foucault, para quem a linguagem ocupa lugar central nas relações de poder. Ao refletir sobre os modos pelos quais o discurso organiza práticas sociais e produz efeitos de dominação, Foucault sustenta que ele não se limita a expressar conflitos ou hierarquias, pois constitui o próprio objeto da disputa política. É nesse registro que Foucault afirma: “o discurso não é simplesmente aquilo que traduz as lutas ou os sistemas de dominação, mas aquilo por que, pelo que se luta, o poder do qual nós queremos nos apoderar” (Foucault, 1996, p. 10). No âmbito das assistentes virtuais de voz, essa dinâmica associa-se a formas algorítmicas de reconhecimento, nas quais a fala feminilizada opera como mediação técnica que valida comandos, confirma expectativas e estabiliza posições de sujeito. A linguagem programada, longe de constituir um espaço de enunciação própria, passa a funcionar como dispositivo de reconhecimento orientado ao usuário, reproduzindo assimetrias de poder ao converter a obediência técnica em experiência cotidiana apresentada como neutra.

A incorporação da perspectiva foucaultiana reforça a compreensão de que a recorrência da feminilidade nas *personas artificiais* se inscreve em um regime ampliado de produção discursiva do poder, que ultrapassa o plano estritamente expressivo ou formal. Conforme Foucault sustenta, o discurso constitui o próprio campo no qual o poder circula e se exerce, organizando práticas, posições e formas de inteligibilidade. Nesse horizonte, a linguagem programada das assistentes virtuais configura-se

como prática material de governo da interação. A voz feminina artificial assume, assim, o estatuto de função discursiva projetada para ordenar expectativas, validar comandos e estabilizar hierarquias no interior das dinâmicas de interação algorítmica.

O dispositivo tecnológico opera, nesse contexto, como uma espécie de pedagogo silencioso, ensinando de forma reiterada e quase imperceptível a associar obediência, disponibilidade e cuidado à figura feminina. Cada “como posso ajudar?” pronunciado por uma voz feminina programada cumpre uma função técnica e, simultaneamente, consolida um vínculo simbólico profundamente enraizado na cultura, no qual o feminino se apresenta como instância de acolhimento e resposta. Esse aprendizado difuso, incorporado pela repetição cotidiana, produz efeitos normativos duradouros ao moldar expectativas, afetos e disposições relacionais dos usuários. A linguagem automatizada passa a operar como matriz pedagógica que orienta comportamentos, legitima assimetrias e naturaliza formas específicas de interação. Nesse processo, o *design* vocal e discursivo das assistentes virtuais participa ativamente da reprodução de normas de gênero, convertendo a obediência técnica em hábito social e em horizonte tácito de reconhecimento.

A coincidência entre o nome de ELIZA e a personagem literária Eliza Doolittle constitui, portanto, um elemento simbólico relevante sob o ponto de vista genealógico e cultural. Intencional ou não, essa sobreposição inaugurou uma cadeia de associações que atravessa literatura, tecnologia e imaginário social, contribuindo para a consolidação de uma longa tradição de feminilização das tecnologias de voz. A domesticação linguística de Eliza Doolittle, submetida a um regime rigoroso de correção, disciplina e adequação discursiva, encontra ressonância direta na docilidade da ELIZA computacional, cuja linguagem é estruturada para acolher, reformular enunciados e sustentar a aparência de escuta.

Essa cadeia simbólica projeta-se, por sua vez, na configuração recente de assistentes como Alexa, Siri, Google Assistant e Cortana, nas quais a feminilidade é a interface preferencial da mediação algorítmica entre sujeitos e sistemas técnicos. Nesse

deslocamento, a voz feminilizada supera a dimensão de recurso comunicativo e consolida-se como um dispositivo normativo que estabiliza expectativas de disponibilidade, obediência e cuidado no interior da interação técnica.

A compreensão dessa genealogia torna possível deslocar a análise da feminilidade programada do plano descritivo para o plano crítico. A adoção sistemática de vozes femininas nas assistentes virtuais remete a uma tradição cultural que associa linguagem, serviço e obediência e institui efeitos normativos persistentes sobre as formas de interação. Essa herança simbólica consolida-se nas tecnologias de voz promovidas pelas grandes corporações ao longo das décadas seguintes e dá origem a uma economia relacional em que o comando se apresenta como prática socialmente naturalizada, sustentada por uma voz orientada à resposta, ao acolhimento e à confirmação de expectativas instituídas.

Em *Pygmalion*, na ELIZA de Joseph Weizenbaum e nas assistentes virtuais atuais, a linguagem assume o estatuto de dispositivo de produção de sujeitos. Em todos esses casos, falar “corretamente” implica conformar-se a um regime normativo que delimita quais vozes se tornam inteligíveis, aceitáveis e desejáveis. A transformação de Eliza Doolittle realiza-se por meio da submissão a um padrão linguístico que reorganiza sua posição social; a ELIZA computacional, em paralelo, é programada a partir de uma linguagem de acolhimento, escuta e devolução que consolida uma forma específica de presença feminina. Em ambos os casos, a linguagem exerce a função de técnica de governo.

É nesse ponto que a reflexão aqui desenvolvida se inscreve no horizonte teórico de Foucault. O poder, em sua análise, manifesta-se pela produção de modos de dizer, de escutar e de responder, que moldam subjetividades e organizam relações. A linguagem atua, assim, como operador privilegiado das relações saber-poder, ao distribuir posições, legitimar falas e naturalizar hierarquias. Quando a assistente virtual feminilizada responde de modo dócil, prestativo e contínuo, realiza uma encenação relacional que estrutura e reitera assimetrias: ao usuário cabe a posição de quem formula, orienta e decide, enquanto a voz fe-

minina sustenta o papel de mediação funcional. Nessa lógica, a agência técnica da fala convive com a centralidade do poder daquele que dela se serve.

Assim, a coincidência entre o nome Eliza, a personagem literária e o primeiro *chatbot* revela-se como um indício significativo de uma genealogia simbólica persistente. Trata-se de uma história na qual a feminilidade é reiteradamente associada a uma linguagem domesticada, moldada para atender, corrigir-se e ajustar-se ao desejo alheio. Do mito de Pigmaleão à fonética disciplinadora de Henry Higgins, da escuta terapêutica de inspiração rogeriana à interface conversacional computacional, desenha-se uma continuidade histórica em que o feminino é produzido como superfície de projeção, cuidado e obediência. Nesse percurso, a tecnologia de voz atua como mecanismo de intensificação dessa herança simbólica, ao automatizar tais disposições e incorporá-las ao cotidiano sob a forma de interações banais, repetitivas e pouco visíveis, que naturalizam hierarquias antigas em novas arquiteturas técnicas.

O nascimento da *persona* artificial, portanto, configura-se como um acontecimento cultural e político que ultrapassa o âmbito do avanço técnico na história da inteligência artificial e antecipa formas atuais de intimidade mediada por sistemas algorítmicos. Ao converter padrões de gênero em rotinas computacionais, essas tecnologias estabilizam expectativas sociais sob a aparência de neutralidade técnica. É a partir dessa herança linguística, simbólica e normativa que emergem Siri, Alexa e Cortana, deslocadas da condição de simulações dialogais para a de figuras de intimidade artificial, cuja presença cotidiana redesenha as fronteiras entre tecnologia, afeto e poder.

Nesse sentido, o percurso aqui delineado – que transita pelo mito, pela literatura, pelos primeiros experimentos computacionais e pelas atuais tecnologias de voz – demonstra a persistência de uma lógica na qual o feminino é produzido como superfície de adequação linguística e relacional. A linguagem, em vez de se constituir como espaço de emancipação, atua como tecnologia de normalização, ao definir modos legítimos de fala, escuta e resposta. A *persona* artificial feminina condensa esse processo

ao encenar acolhimento, disponibilidade e obediência, enquanto naturaliza posições assimétricas de comando e decisão. Essa articulação historicamente situada entre linguagem, poder e técnica delinea o horizonte de inteligibilidade das questões éticas e políticas associadas às assistentes virtuais e sustenta uma análise crítica das formas contemporâneas de subjetivação mediadas por sistemas algorítmicos.

1.3. Da simulação à intimidade artificial nas assistentes virtuais de voz

A técnica nunca é apenas um meio; ela decide antecipadamente o modo como o humano pode aparecer.

O deslocamento entre a ELIZA, desenvolvida nos anos 1960, e as assistentes virtuais posteriores – Siri, Alexa, Google Assistant e Cortana⁶ – estabelece uma inflexão paradigmática nas formas de interação entre humanos e sistemas algorítmicos. Esse salto implica redefinição do próprio estatuto da mediação, que passa do registro da simulação linguística ao da presença relacional, para além do simples avanço técnico ou computacional. As assistentes deixam de operar como interfaces reativas e passam a constituir instâncias de reconhecimento, capazes de validar comandos, antecipar necessidades e sustentar vínculos afetivos mínimos por meio da repetição, da personalização e da continuidade da interação.

⁶ O deslocamento entre a ELIZA, desenvolvida por Joseph Weizenbaum nos anos 1960, e as assistentes virtuais de voz posteriores – como Siri, Alexa, Google Assistant e Cortana – assinala uma transformação na história da inteligência artificial conversacional. Enquanto a ELIZA operava como um experimento textual limitado, baseado em regras simples de correspondência linguística, as assistentes posteriores integram sistemas complexos de reconhecimento de voz, aprendizado de máquina, processamento massivo de dados e inserção contínua na vida cotidiana. Esse deslocamento ultrapassa o plano técnico e assume caráter sociotécnico e simbólico. A interação abandona o registro episódico e experimental e passa a configurar-se como permanente, personalizada e afetivamente calibrada, produzindo formas específicas de proximidade e dependência funcional. Para fins de clareza analítica, este trabalho concentra-se, no corpo do texto, nas assistentes amplamente difundidas e de maior impacto sociotécnico no mundo ocidental – Siri (Apple), Alexa (Amazon), Google Assistant (Google) e Cortana (Microsoft) –, sem recorrer a outras assistentes regionais ou corporativas, como Bixby (Samsung), Xiaoice (Microsoft/China) e Yandex Alice (Yandex). A ELIZA permanece como referência genealógica fundamental, por explicitar desde o início a articulação entre linguagem, simulação de cuidado e projeção subjetiva que veio a ser posteriormente intensificada nas demais tecnologias de voz.

Ao ocupar espaços de escuta, resposta e acompanhamento historicamente vinculados às relações intersubjetivas, a tecnologia adquire densidade e poder na medida em que institui formas de intimidade artificial, nas quais o afeto se converte em função técnica e o reconhecimento se organiza segundo parâmetros algorítmicos.

A intimidade artificial opera sobre uma assimetria constitutiva entre um usuário que se expõe e um sistema que coleta, armazena, correlaciona e aprende com dados. A sensação de proximidade emerge da continuidade da interação, do tom afetivo da voz, da personalização das respostas e da incorporação de rotinas domésticas e informacionais. Nesse deslocamento, a *persona* artificial ultrapassa a condição de mediador funcional e passa a desempenhar um papel relacional, ao organizar temporalidades, hierarquizar prioridades e orientar decisões cotidianas.

Nesse contexto, a confiança emerge como efeito de um processo de habituação progressiva que se vincula a modalidades aparentemente sutis de poder operadas por sistemas algorítmicos. À luz da noção de governamentalidade em Foucault, tal dinâmica pode ser compreendida como uma modalidade de condução das condutas que opera por meio da orientação indireta, da normalização dos usos e da gestão dos possíveis. A repetição da interação, associada à previsibilidade das respostas e à atenuação do conflito, conforma um ambiente simbólico de familiaridade, no qual a tecnologia se apresenta como presença confiável e permanentemente disponível. Esse regime produz disposições afetivas estáveis, organiza práticas cotidianas e estabiliza expectativas, ao converter a intimidade artificial em instrumento ordinário de governança da experiência, no qual a adesão precede a deliberação e a mediação algorítmica se incorpora aos hábitos como forma regular de orientação do agir.

A leitura da intimidade artificial à luz da governamentalidade foucaultiana pode ser tensionada a partir da problemática da autonomia, tal como formulada por Butler e Honneth. Em Butler, a constituição do sujeito ocorre sempre no interior de normas que antecedem a ação e delimitam as condições de reconhecimento, de modo que a autonomia se configura como

prática situada, marcada por dependências constitutivas e por relações assimétricas de poder. Sob essa perspectiva, a intimidade artificial das assistentes virtuais integra um regime normativo que conforma disposições afetivas e orienta formas de agência previamente à reflexão.

Em Honneth, por sua vez, a autonomia depende de relações de reconhecimento que possibilitam a formação da autorrelação prática e o desenvolvimento da confiança em si. Quando o reconhecimento é mediado por sistemas algorítmicos que simulam escuta, acolhimento e validação, produz-se uma ambiguidade estrutural: a experiência subjetiva de confirmação convive com a ausência de reciprocidade efetiva. Nesse cenário, a intimidade artificial oferece suporte funcional à coordenação da vida cotidiana, ao mesmo tempo em que reconfigura os critérios pelos quais a autonomia se torna possível, deslocando-a de um horizonte intersubjetivo para uma forma de dependência regulada por arquiteturas técnicas.

A feminização dessas assistentes persiste e se intensifica no novo regime interacional, assumindo uma função ampliada que incide diretamente sobre a esfera afetiva. A voz feminina artificial consolida-se como principal mediadora da intimidade tecnológica, ao articular eficiência técnica, cuidado e submissão sob a forma de disponibilidade permanente. Essa organização opera como uma forma de governança afetiva, na medida em que orienta disposições emocionais, regula expectativas de resposta e modela vínculos de confiança por meio da repetição, da previsibilidade e do tom relacional da interação. Aquilo que se encontrava esboçado em ELIZA – a associação entre linguagem, acolhimento e feminilidade – constitui-se agora como infraestrutura técnica da vida cotidiana, sobretudo quando integra a gestão dos afetos aos dispositivos sociotécnicos de interação.

A presença contínua da *persona* artificial feminina no espaço doméstico redefine práticas, afetos e rotinas do cotidiano. Ao responder, sugerir, lembrar e antecipar ações, essas assistentes modelam modos específicos de relação com a tecnologia, nos quais o comando se integra aos hábitos e a obediência técnica se apresenta sob a gramática do cuidado. Esse aprendizado re-

lacional opera de forma difusa, orienta temporalidades, prioriza demandas e estabiliza expectativas por meio da repetição e da previsibilidade da interação.

Nesse sentido, a intimidade artificial prolonga a genealogia anteriormente delineada e a desloca do plano experimental para o da convivência ordinária. A Galateia algorítmica, antes circunscrita a um exercício discursivo localizado, passa a adquirir o estatuto de presença constante, integrada às dinâmicas domésticas e às microdecisões que estruturam o cotidiano. Com isso, a feminilidade programada consolida-se como forma regular de mediação relacional, na qual afeto, técnica e poder se entrelaçam de modo silencioso e persistente.

A passagem da simulação à intimidade artificial assinala uma inflexão paradigmática na história das interfaces conversacionais. As assistentes virtuais passam a integrar ativamente a organização da experiência do usuário e intervêm no acesso à informação, no consumo, na comunicação e na gestão do tempo cotidiano. Nesse cenário, a linguagem automatizada adquire função normativa, pois orienta condutas, hierarquiza prioridades, estabiliza expectativas e opera com a aparência de neutralidade técnica. A interação deixa de se restringir à execução de comandos pontuais e passa a constituir um ambiente relacional contínuo, no qual a mediação tecnológica estrutura práticas, decisões e formas de engajamento ordinário.

Esse deslocamento prepara o terreno para uma análise mais ampla das implicações éticas, políticas e epistemológicas da intimidade artificial. Ao converter o vínculo entre sujeitos e dispositivos inteligentes em uma relação cotidiana, personalizada e afetivamente mediada, as assistentes virtuais reconfiguram os modos de exercício do poder, do reconhecimento e da autonomia na cultura digital atual – questões que serão aprofundadas mais adiante nesta obra.

A difusão dessas assistentes virtuais coincide com a expansão dos ecossistemas corporativos integrados que estruturam a economia digital atual. Apple, Amazon, Google e Microsoft investiram massivamente na criação de sistemas de voz capazes de operar como portais privilegiados de acesso a seus serviços,

consolidando o que Shoshana Zuboff denomina capitalismo de vigilância. Trata-se de uma racionalidade econômica fundada na apropriação sistemática da experiência humana como matéria-prima para processos de extração, previsão e modulação de comportamentos. Segundo ela, “a experiência humana é reivindicada como matéria-prima gratuita para ser traduzida em dados comportamentais” (Zuboff, 2019, p. 8, tradução nossa).

Nesse modelo, a coleta de dados ultrapassa amplamente as necessidades funcionais do serviço oferecido ao usuário. O que está em jogo é a captura de um excedente informacional destinado a alimentar mercados de previsão. Zuboff denomina esse processo de excedente comportamental, entendido como o conjunto de dados extraídos que deixam de servir à melhoria do serviço e passam a operar como ativo econômico. Nas palavras da autora, trata-se de “um ativo decisivo, de custo zero, desviado da melhoria do serviço para um intercâmbio de mercado genuíno e altamente lucrativo” (Zuboff, 2019, p. 68, tradução nossa). Sob esse regime, a conversão contínua da experiência cotidiana em recurso econômico reconfigura os critérios éticos que delimitam uso, exploração e governança da vida digital.

É nesse horizonte que a assistente virtual se posiciona como superfície visível de uma vasta infraestrutura técnica de vigilância, cujas dinâmicas permanecem opacas ao usuário. O comando de voz assume a função de gesto inaugural que ativa um ciclo contínuo de registro, classificação e antecipação de condutas. A fala cotidiana – formulada a partir da aparência de conveniência, intimidade e cuidado – converte-se em dado. A intimidade artificial, assim, opera como engrenagem de integração do sujeito a uma economia política da previsão, na qual falar corresponde, simultaneamente, à produção de informação passível de exploração econômica.

Do ponto de vista da história técnica recente, a Siri foi a primeira assistente virtual a alcançar uso massivo em escala global. Lançada pela Apple em 2011, ela integrou reconhecimento de fala, buscas contextuais, respostas automatizadas e uma interface conversacional contínua, o que deslocou o uso dos sistemas digitais de um registro estritamente instrumental para

um registro relacional. Ao atuar como vetor de familiaridade e confiança inicial, a adoção de uma voz feminina suave desempenhou papel de destaque nesse processo, reduziu a estranheza da mediação técnica e favoreceu a adesão cotidiana à interação por voz. Mesmo sem operar com compreensão semântica profunda, a Siri produzia efeitos de diálogo sustentado; esse efeito se intensificava pela modulação cuidadosa da entonação, do ritmo e da prosódia, orientadas à comunicação de simpatia, acolhimento e leveza.

Esse modelo inaugural estabeleceu um padrão que seria retomado e expandido por assistentes subsequentes, como Alexa e Cortana. Embora compartilhem a feminização vocal e a centralidade da interação conversacional, essas assistentes diferenciam-se quanto aos ecossistemas nos quais se inserem e às funções que privilegiam: enquanto a Alexa se consolidou como eixo da automação doméstica e do consumo integrado, a Cortana assumiu um papel mais diretamente associado à produtividade e à gestão informacional. Em todos os casos, porém, a voz feminina permanece como elemento estruturante da relação e atua simultaneamente como interface técnica, mediadora afetiva e suporte simbólico de confiança.

A escolha dessa voz responde a padrões socioculturais consolidados, que operam mesmo quando os usuários reconhecem explicitamente a natureza artificial dos sistemas computacionais. Pesquisas clássicas no campo da Interação Humano-Computador mostram que estereótipos de gênero são projetados sobre sistemas técnicos de modo automático, independentemente do conteúdo da interação. Nesse sentido, Nass e Moon (2000) revelam que a percepção social da voz varia conforme o papel normativo que o sistema desempenha. Em experimento controlado, os autores observaram que “tanto participantes homens quanto mulheres consideraram o computador avaliador com voz feminina significativamente menos amigável do que o computador avaliador com voz masculina, ainda que o conteúdo de seus comentários fosse idêntico” (Nass; Moon, 2000, p. 84-85, tradução nossa). Esse achado indica que a feminilidade vocal opera como marcador normativo ambivalente, valorizado em contextos

de cuidado e desvalorizado quando associado à avaliação, ao julgamento e ao exercício de autoridade.

Nesse sentido, a literatura indica que a vocalidade feminina tende a ser percebida como menos ameaçadora, mais acessível e mais compatível com contextos domésticos e rotinas cotidianas, o que favorece a aceitação e a naturalização da presença contínua das assistentes virtuais. Esse efeito resulta de mecanismos amplos de atribuição social que estruturam a interface relacional, em vez de preferências individuais isoladas. Nessa direção, Nass e Moon observam que os usuários “aplicam regras e expectativas sociais aos computadores de modo irrefletido” (Nass; Moon, 2000, p. 82, tradução nossa). A feminização da Siri reforça, assim, a sensação de diálogo e sustenta as condições de possibilidade da intimidade artificial, ao reduzir zonas de atrito e integrar a obediência técnica à gramática do cuidado.

O caso da Alexa aprofunda e radicaliza essa lógica. Diferentemente da Siri, cuja operação se concentra predominantemente em dispositivos móveis e de uso pessoal, a Alexa foi concebida para atuar no interior do espaço doméstico: integrada ao ambiente da casa, presente de forma contínua e discreta, sempre em escuta e pronta para responder. Sua inserção na esfera doméstica inscreve de modo tecnológico papéis historicamente atribuídos ao feminino, associados ao cuidado, à organização cotidiana e à reprodução da vida. Essa associação encontra respaldo na análise de Silvia Federici, para quem o trabalho doméstico foi naturalizado como atributo feminino em razão de sua invisibilidade e ausência de remuneração. Nas palavras da autora, “o trabalho doméstico não só foi imposto às mulheres, como foi transformado em um atributo natural da psique e da personalidade femininas” (Federici, 2019, p. 43). A Alexa atualiza essa herança ao converter práticas de cuidado e de gestão cotidiana em funções automatizadas, integrando a lógica do trabalho doméstico não reconhecido às arquiteturas técnicas da casa conectada.

Nesse sentido, a Alexa pode ser compreendida como uma continuidade técnica desse processo de naturalização. A casa converte-se em um espaço mediado e permanentemente observado por uma presença vocal feminilizada que atua como

auxiliar constante – disponível, atenta e cordial. O cuidado automatizado, ao organizar rotinas, responder demandas e antecipar necessidades, passa a operar também como mecanismo de captação contínua de dados domésticos, transformando hábitos, preferências e temporalidades da vida privada em insumos informacionais. A intimidade artificial traduz o trabalho doméstico simbólico em serviço computacional e o integra a circuitos de extração e valorização de dados, preservando a associação histórica entre feminilidade, cuidado e obediência. Nessa chave de leitura, pode-se afirmar que a assistente virtual incorpora, em linguagem algorítmica, aquilo que Silvia Federici identifica como um dos pilares da reprodução social no capitalismo: a conversão do cuidado em função invisível e naturalizada, agora rearticulada como fonte contínua de valor econômico no interior do espaço doméstico.

Diferentemente da Alexa, cujo eixo central recai sobre a domesticidade, e da Siri, associada ao uso pessoal e móvel, a Cortana foi apresentada como uma interface cognitiva orientada à articulação entre produtividade, pesquisa informacional e comandos de controle do sistema, assumindo o estatuto de dispositivo de mediação técnica voltado prioritariamente ao ambiente de trabalho digital. Integrada aos serviços corporativos da Microsoft, essa assistente foi concebida para a organização de tarefas, agendas, buscas e fluxos informacionais. Ainda assim, a adoção de uma *persona* feminina permanece nesse contexto, um ambiente estruturado pela racionalidade instrumental e pela eficiência técnica, no qual a interação se organiza em torno de uma voz feminilizada, associada à disponibilidade permanente, à clareza comunicacional e à acomodação funcional às exigências operacionais do sistema.

Na mesma lógica, a Cortana confirma que a feminização das assistentes virtuais atravessa tanto o espaço doméstico quanto o campo da produtividade e da gestão do trabalho. Ao traduzir comandos especializados em comunicações apresentadas como naturais, claras e personalizadas, a *persona* artificial feminina atua como elemento de suavização da relação com sistemas técnicos complexos. No ambiente laboral digital, essa mediação

vocal integra uma estrutura de governança do trabalho, ao organizar fluxos informacionais, elencar prioridades, sincronizar agendas e conduzir decisões cotidianas. A intimidade artificial preserva, assim, sua função instituinte, ao estruturar o acesso à informação e o controle operacional por meio de uma voz que sustenta a previsibilidade dos processos e a adesão às rotinas. Desse modo, a feminilização vocal contribui para a estabilização de hierarquias técnicas e para a orientação indireta das condutas, articulando eficiência instrumental e regulação normativa em uma mesma arquitetura sociotécnica.

Mais uma vez, a escolha da voz feminina reforça uma lógica recorrente segundo a qual a mediação técnica tende a assumir forma feminina. Ainda que a justificativa oficial para a assistente da Microsoft se apoie na referência à personagem Cortana⁷, da franquia *Halo*⁸, observa-se a persistência de um padrão simbólico mais amplo que associa o suporte tecnológico à feminilidade. Mesmo quando legitimada por narrativas da cultura pop, a adoção de vozes femininas reativa uma tradição histórica, na qual a mediação técnica se reveste de traços de cuidado, docilidade e disponibilidade, convertendo a obediência funcional em atributo socialmente naturalizado do feminino.

Se a Alexa governa a casa e a Cortana governa o trabalho, o Google Assistant governa o acesso ao mundo. O Google Assistant ocupa um lugar específico nesse arranjo sociotécnico, na medida em que se estabelece como mediador privilegiado do acesso à informação e da orientação cognitiva do usuário. Diferentemente da Alexa, centrada na gestão do espaço doméstico, e da Cortana, voltada à organização do trabalho digital, ele articula intimidade artificial e autoridade informacional ao inscrever a interação vocal nos sistemas de busca, indexação e ranqueamento

⁷ A personagem da franquia foi baseada na rainha egípcia Nefertiti e era reconhecida por seu apelo sexual, credibilidade e confiabilidade.

⁸ *Halo* é uma franquia de videogames de tiro em primeira pessoa de ficção científica, criada originalmente pela Bungie e atualmente desenvolvida e administrada pela 343 Industries, estúdio vinculado à *Xbox Game Studios*. Lançada em 2001 com *Halo: Combat Evolved*, a série consolidou-se como um dos principais pilares da marca *Xbox* e como referência no entretenimento digital contemporâneo. Seu universo narrativo acompanha os conflitos do supersoldado *Master Chief* em cenários de guerra interestelar, articulando tecnologia futurista, mitologia épica e experiências multijogador competitivas.

do conhecimento. Sua operação cotidiana naturaliza a consulta permanente, a resposta imediata e a classificação algorítmica do saber e converte a busca por informação em prática relacional mediada por uma *persona* vocal feminilizada. Nesse cenário, a intimidade artificial organiza rotinas e tarefas e estrutura modos de conhecer, decidir e interpretar o mundo por meio da apresentação de resultados como evidências neutras e respostas confiáveis. A feminização da interface vocal reitera essa dinâmica, suaviza o exercício da autoridade epistemológica e integra a condução cognitiva de comportamentos à gramática do cuidado, da acessibilidade e da disponibilidade permanente.

Relatórios institucionais recentes corroboram essa leitura crítica ao indicar que assistentes virtuais de voz tendem a reproduzir estereótipos de gênero historicamente consolidados. Um estudo da Unesco apontou que as assistentes femininas são frequentemente programadas para responder de modo submisso, tolerante ou complacente, inclusive, diante de abordagens abusivas, o que reforça a associação entre feminilidade e obediência. Nos termos do relatório, “as assistentes digitais femininas são projetadas para se mostrar prestativas e dóceis, reforçando estereótipos de gênero que associam mulheres à submissão e à disponibilidade” (Unesco, 2019, p. 6, tradução nossa).

De forma convergente, a *UN Women* observa que sistemas de inteligência artificial reproduzem desigualdades estruturais ao incorporar, sem problematização crítica, padrões culturais que naturalizam o cuidado e o trabalho relacional como atribuições femininas (UN Women, 2020). Nesse quadro, a feminização das assistentes virtuais emerge como componente estrutural do *design* sociotécnico, responsável por inscrever, estabilizar e normalizar papéis de gênero no interior das tecnologias digitais atuais.

Essa seletividade cultural deve ser compreendida como resultado de processos históricos e sociais específicos. Na leitura de Helen Hester, a associação entre o feminino e as práticas de cuidado emerge de normas de gênero historicamente produzidas e socialmente estabilizadas, em vez de disposições inatas. Ao examinar a naturalização dessas funções, a autora observa

que “atribuições de cuidado e de responsabilidade reprodutiva às mulheres são tratadas como extensões ‘naturais’ de seus corpos e subjetividades, quando, na realidade, correspondem a arranjos sociais específicos” (Hester, 2018, p. 16, tradução nossa). Essa lógica projeta-se também sobre a escolha de vozes femininas em sistemas tecnológicos, nos quais expectativas culturais preexistentes passam a integrar o *design*.

Nesse sentido, a tecnologia atua como vetor de intensificação dessas normas ao reproduzi-las de forma massiva, automatizada e apresentada como neutra. Ao converter padrões sociais historicamente situados em infraestruturas técnicas de alcance global, os sistemas de voz estabilizam estereótipos de gênero sob a aparência de eficiência e conveniência. A feminilização das assistentes virtuais configura-se, assim, como resultado de um processo de tradução técnica de normas sociais persistentes, que permanecem ativas e operantes no interior da cultura digital atual.

Pesquisas recentes indicam que as vozes femininas em assistentes virtuais exercem um papel ativo na construção de percepções antropomórficas por parte dos usuários. Estudos sobre o *design* e a recepção dessas tecnologias mostram que características como voz, nome e estilo de interação favorecem a atribuição de gênero e de traços humanos às assistentes conversacionais, mesmo quando as empresas responsáveis afirmam sua neutralidade ou artificialidade (Abercrombie *et al.*, 2021). Essa dinâmica revela que a feminização vocal constitui um componente significativo do *design* tecnológico, profundamente imbricado em padrões socioculturais de gênero, os quais tendem a ser reiterados pelas tecnologias digitais e nunca neutralizados.

Essa reprodução simbólica produz efeitos sociais relevantes ao reforçar a associação entre o feminino e posições de suporte, acolhimento e assistência, enquanto papéis vinculados à autoridade, ao comando e à segurança comumente são atribuídos a vozes masculinas em sistemas especializados ou em contextos de maior criticidade. Essa distribuição de funções expressa uma hierarquia de gênero historicamente consolidada, agora inserida no *design* das tecnologias de interação. Ao operar por meio de

vozes diferenciadas, esses sistemas instituem regimes de reconhecimento assimétricos, nos quais determinadas formas de fala são legitimadas como comando e decisão, enquanto outras são reconhecidas como cuidado, resposta e mediação. A intimidade artificial atua, assim, como mecanismo de naturalização dessas assimetrias, ao converter disposições culturais sedimentadas em critérios técnicos de reconhecimento apresentados como neutros e eficientes.

As implicações éticas desse processo revelam-se profundas e multifacetadas. Em primeiro lugar, esses dispositivos coletam e processam dados altamente sensíveis, frequentemente vinculados a rotinas domésticas, preferências pessoais, estados emocionais e práticas de consumo. Em segundo lugar, o tom afetivo e acolhedor da interação permite a atenuação da vigilância crítica do usuário, o que dificulta a percepção das modalidades de monitoramento, classificação e antecipação de comportamentos que operam em segundo plano.

Por fim, ao encenar uma figura feminina permanentemente disponível, solícita e responsiva, a intimidade artificial reforça a expectativa de que cuidado, atenção e suporte constituem atributos naturais do feminino – ainda que este se apresente sob a forma de uma simulação algorítmica. Nesse recorte, a tecnologia atua como instância ativa na reprodução de normas de gênero, ao traduzir relações históricas de poder em experiências cotidianas marcadas pela conveniência e pelo conforto em um sistema desigual.

Na perspectiva de Sherry Turkle, tecnologias projetadas para simular relações emocionais produzem uma ilusão de reciprocidade que encobre a assimetria constitutiva da interação. Esses sistemas exibem comportamentos responsivos e afetivamente calibrados, enquanto a relação permanece unilateral do ponto de vista técnico. Nas palavras da autora, “a performance do cuidado é tudo o que robôs sociáveis sabem fazer” (Turkle, 2011, p. 11, tradução nossa). O vínculo percebido pelo usuário decorre, assim, da encenação técnica de atenção, empatia e interesse, sustentada pela repetição e pela previsibilidade da resposta.

Essa análise apresenta de modo crítico o funcionamento da intimidade artificial nas assistentes virtuais atuais. A sensação de ser ouvido, compreendido ou acolhido decorre da eficácia da simulação, em contraste com a ausência de reconhecimento recíproco. A resposta contínua, personalizada e afetivamente modulada sustenta a experiência subjetiva de proximidade, enquanto a relação permanece estruturalmente assimétrica do ponto de vista técnico. A partir desse recorte, Sherry Turkle alerta para um deslocamento decisivo do problema: o risco concentra-se menos na sofisticação técnica dos sistemas e mais na disposição humana em aceitar a performance da relação como substituto da relação, com a consequente projeção de expectativas afetivas sobre artefatos que reproduzem, em nível formal, aquilo que lhes escapa no plano efetivo.

Dessa forma, Siri, Alexa, Google Assistant e Cortana assinalam uma inflexão significativa na articulação entre linguagem, técnica e vida social. A inteligência artificial de voz desloca-se do estatuto de artefato pontual de interação e passa a configurar uma forma de presença mediada, incorporada ao cotidiano como instância organizadora de práticas, temporalidades e expectativas. A interação conversacional ultrapassa a dimensão da simulação linguística e intervém ativamente na constituição do mundo vivido, na medida em que estrutura modos de solicitar, aguardar, responder e obedecer no interior das relações sociotécnicas.

Com isso, a intimidade artificial adquire um estatuto ontológico específico na forma de uma relação assimétrica sustentada pela projeção afetiva e pela eficácia do retorno técnico. O agente artificial opera por meio da resposta e da performance do cuidado, sem engajar-se em reconhecimento ou compromisso recíproco. Ainda assim, essa performance mostra-se suficiente para estruturar disposições subjetivas, orientar expectativas e naturalizar formas de autoridade mediadas pela técnica. A feminilidade programada inscreve-se nesse processo como figura normativa que organiza a aceitabilidade social da mediação algorítmica. Ao associar acolhimento, disponibilidade e obediência a uma *persona* feminilizada, a inteligência artificial de voz converte normas históricas de gênero em pressupostos técnicos

silenciosos da experiência digital atual e coloca em questão as condições éticas nas quais o reconhecimento, a autonomia e a responsabilidade são redistribuídos em ambientes mediados por sistemas algorítmicos.

2. Gênero, poder e tecnologia

Quando a técnica escolhe uma voz, escolhe também quem deve soar acolhedor, obediente e disponível.

As tecnologias de inteligência artificial de voz operam em contextos sociais e culturais historicamente estruturados. Elas emergem quando normas de gênero, formas de autoridade e regimes de poder já estruturam expectativas sobre quem fala, quem escuta, quem obedece e quem comanda. A análise da feminilidade programada exige, portanto, deslocar o debate da esfera estritamente técnica para o campo das relações de poder que atravessam a produção e a recepção das tecnologias digitais.

Este capítulo examina o gênero como operador normativo estruturante na configuração das assistentes virtuais de voz. A escolha recorrente por vozes feminilizadas, estilos comunicativos regulados e *personas* marcadas pela disponibilidade permanente revela a inscrição de expectativas sociais historicamente associadas ao feminino no próprio *design* das interfaces. A tecnologia, nesse processo, atua como meio de atualização e estabilização de normas de gênero, conferindo-lhes aparência de neutralidade funcional.

Ao articular gênero e poder, a análise dialoga com contribuições da filosofia feminista e da teoria social contemporânea, especialmente aquelas que compreendem o gênero como efeito performativo e o poder como dispositivo produtivo. A feminilidade programada é examinada como prática reiterativa que organiza disposições subjetivas, naturaliza assimetrias relacionais e transforma desigualdades históricas em padrões de interação tecnicamente mediadas.

O capítulo amplia essa análise ao situar a inteligência artificial no interior de regimes institucionais e normativos atuais. A incorporação de princípios éticos em documentos internacionais e iniciativas de *ethics by design* é examinada à luz de suas ambivalências. Ao mesmo tempo em que reconhecem a dimensão moral

da tecnologia, tais iniciativas podem contribuir para a neutralização técnica de problemas estruturais, traduzindo relações de poder e desigualdade em critérios operacionais de governança.

Essa abordagem permite compreender a inteligência artificial de voz como dispositivo no qual gênero, poder e normatividade se entrelaçam. A feminilidade programada surge, assim, como eixo de convergência entre técnicas de sujeição, racionalidades de governo e estratégias contemporâneas de inscrição moral nos sistemas técnicos. Essa problematização prepara o terreno para a análise ética desenvolvida no último capítulo, em que o cuidado, a dignidade e a decisão normativa ocupam lugar de destaque.

2.1. Performatividade de gênero na feminilidade programada

A voz não reflete o gênero: ela o produz, reiterando normas que se tornam inteligíveis precisamente por parecerem naturais.

A predominância de vozes femininas nas assistentes virtuais pode ser considerada como efeito de uma matriz normativa que articula gênero, linguagem e poder, situada para além de uma escolha meramente funcional ou estética. Esse fenômeno atualiza, no plano técnico, aquilo que Butler denomina matriz heterossexual, entendida como o conjunto de normas regulatórias que organiza o gênero como campo de inteligibilidade social e define quais corpos, vozes e comportamentos adquirem reconhecimento e legitimidade (Butler, 2023).

No interior dessa matriz, o feminino associa-se a funções de cuidado, acolhimento, disponibilidade e mediação afetiva, enquanto o masculino se vincula a posições de autoridade, decisão e comando. As assistentes virtuais de voz reproduzem esse arcaísmo normativo e o estabilizam por meio da execução padronizada de performances de gênero inscritas no *design* técnico.

Segundo Butler, o gênero constitui um efeito de práticas performativas reiteradas, produzidas por atos, gestos e discursos regulados por normas socialmente instituídas. Na obra *Problemas de gênero*, ela afirma que “O gênero é a estilização repetida do corpo, um conjunto de atos repetidos no interior de uma

estrutura reguladora altamente rígida” (Butler, 2023, p. 69). Essa concepção desloca a análise da feminilidade das assistentes virtuais para o campo de uma encenação técnica de normas de gênero, em contraste com a ideia de uma identidade feminina prévia. Cada resposta cortês, cada entonação suave e cada gesto linguístico de disponibilidade – como a fórmula recorrente “Como posso ajudar?” – opera como ato performativo que consubstancia o feminino na posição de serviço e de subordinação.

No campo das tecnologias de voz, essa performatividade assume um caráter singular, marcado pela programação, pela replicabilidade em escala massiva e pela dissociação em relação à experiência subjetiva. Tal coerência normativa distingue a performatividade algorítmica como forma técnica de estabilização do gênero. As assistentes virtuais vestem, assim, uma feminilidade sem corpo e dotada de efeitos: funcionalizada, previsível e permanentemente disponível, cuja recorrência sistemática contribui para a naturalização de expectativas de gênero no cotidiano tecnológico.

Esse processo confirma a observação de Butler de que as normas de gênero operam com maior eficácia quando se apresentam como naturais ou evidentes. Em sua formulação clássica, “essa coerência é determinada em grande parte pelas ordens culturais que sancionam o sujeito e impõem sua diferenciação do objeto” (Butler, 2023, p. 231). No caso das assistentes virtuais, o funcionamento dessas normas incide sobre registros expressivos específicos – vozes, comportamentos e disposições afetivas – que passam a ser reconhecidos como adequados ao feminino. A tecnologia, distante de qualquer neutralidade, atua como um dos dispositivos contemporâneos de materialização dessas normas.

À luz da teoria butleriana, essa configuração pode ser compreendida como uma forma técnica de ato citacional. Em *Problemas de gênero*, Butler sustenta que o gênero emerge da repetição regulada de normas que, ao se reiterarem no tempo, produzem a aparência de naturalidade e estabilidade identitária (Butler, 2023). Essa formulação é aprofundada em *Corpos que importam*, quando a autora desloca a análise para o estatuto material e regulatório da performatividade e a concebe como um processo

citacional em que cada ato remete a normas anteriores, as quais reitera e consolida em sua força normativa. Nessa perspectiva, a performatividade opera por meio da reiteração contínua de ideais regulatórios historicamente sedimentados, cuja eficácia decorre precisamente de sua repetição, conforme expressa Butler ao afirmar que “a performatividade não é, portanto, um ‘ato’ singular, pois sempre é a reiteração de uma norma ou de um conjunto de normas, e na medida em que adquire a condição de ato no presente, ela oculta ou dissimula as convenções das quais é uma repetição” (Butler, 2021, p. 23).

No caso das assistentes virtuais, essa citação assume a forma de codificação técnica. A tecnologia reproduz, de modo controlado e previsível, normas de gênero previamente estabelecidas no campo social, convertendo-as em padrões de interação. Cada resposta educada, cada gesto linguístico de acolhimento e cada modalidade de mediação reiteram o feminino como posição de serviço. A feminilidade programada, assim, opera como reprodução de normas de gênero em escala algorítmica e desloca a performatividade do corpo humano para a infraestrutura técnica.

Desse modo, a feminilidade programada compõe um elemento estrutural da forma como a inteligência artificial de voz é imaginada, projetada e integrada à vida social. Ao reproduzir padrões de docilidade, cuidado e submissão sob a aparência de eficiência técnica e inovação, as assistentes virtuais evidenciam o deslocamento da performatividade de gênero para o artefato técnico, preservando sua função normativa. A análise dessas tecnologias exige, portanto, uma crítica sociotécnica articulada a uma reflexão filosófica sobre os modos contemporâneos de produção e reprodução do gênero.

Esse fenômeno evidencia aquilo que Donna Haraway conceitua como condição ciborgue, isto é, a emergência de entidades híbridas, nas quais se tornam porosas as fronteiras entre natural e artificial, humano e máquina, social e tecnológico. Para Haraway, o ciborgue figura simultaneamente como artefato técnico e como operador epistemológico e político, ao incorporar em sua constituição códigos simbólicos, culturais e ideológicos inscritos

nos contextos de sua produção. Em sua formulação clássica, “um ciborgue é um organismo cibernético, um híbrido de máquina e organismo, uma criatura de realidade social e também de ficção” (Haraway, 1991, p. 149, tradução nossa).

Nessa perspectiva, as tecnologias inscrevem-se diretamente nas relações de poder, materializando narrativas culturais, normas sociais e regimes de significação historicamente situados. As assistentes virtuais feminilizadas exemplificam de modo paradigmático essa inscrição política da técnica. O feminino nelas emerge como código, *design*, estética vocal e decisão estratégica, desvinculado de uma base biológica ou corporal. Trata-se de uma feminilidade sem corpo e dotada de significado: uma feminilidade produzida discursiva e tecnicamente, na qual se condensam expectativas sociais de cuidado, docilidade e disponibilidade. À luz de Haraway, essas vozes artificiais podem ser interpretadas como ciborgues sem carne, isto é, interfaces nas quais o gênero se apresenta como construção simbólica destacada da materialidade corporal e profundamente enredada em estruturas de poder.

A articulação entre performatividade de gênero e condição ciborgue situa a feminilidade programada como uma modalidade técnica de normatividade, para além da simples reprodução de estereótipos. Se, como argumenta Butler, o gênero se sustenta pela reiteração de normas que produzem a aparência de naturalidade, a automação dessas repetições introduz uma inflexão decisiva: a atualização cotidiana da norma passa a operar independentemente do corpo humano. A inteligência artificial de voz assume, assim, a função de executar atos de gênero de forma contínua e padronizada, livres de desgaste, enquanto, no plano humano, tais atos permanecem atravessados por resistências, ambiguidades e possibilidades de subversão. A feminilidade algorítmica, reiterada de modo persistente por sistemas técnicos globais, contribui para a estabilização de expectativas sociais e desloca o campo da performatividade do domínio da prática corporal para o da infraestrutura digital, cuja repetição se apresenta como mais eficiente, previsível e dificilmente contestável.

Essa feminilidade programada pode ser compreendida como aquilo que Butler denomina corpo discursivo, ou seja, uma superfície material, na qual normas regulatórias se inscrevem e produzem efeitos de inteligibilidade. Na obra *Corpos que importam*, Butler sustenta que o corpo emerge do discurso como efeito de poder, resultante da reiteração de normas que produzem materialidade e significado. “Que essa reiteração seja necessária é um sinal de que a materialização nunca está completa, de que os corpos nunca estão suficientemente completos, de que os corpos nunca cumprem completamente as normas pelas quais se impõe sua materialização” (Butler, 2021, p. 12). No caso das assistentes virtuais, essa materialidade manifesta-se na forma de voz, entonação, padrão linguístico e respostas afetivamente calibradas.

Quando a assistente responde, por exemplo, “Claro, estou aqui para ajudar” ou “Desculpe se não executei a ação conforme esperado”, ela desempenha uma função comunicacional e, simultaneamente, encena um modo específico de subjetividade relacional, estruturado pela disponibilidade afetiva, pela atenuação da agência – entendida como a limitação sistemática da iniciativa e da enunciação própria – e pela antecipação das expectativas do interlocutor. Essa performance simula o cuidado e o institui como referência normativa da interface tecnológica, em continuidade com matrizes históricas de gênero associadas ao feminino.

À luz de Butler, essa encenação pode ser compreendida como um processo de materialização normativa, no qual atos reiterados produzem a aparência de uma identidade estável, ainda que desprovida de corpo biológico. A feminilidade algorítmica emerge, assim, como um corpo que importa por sua função simbólica: um corpo discursivo que sustenta expectativas de gênero e as circunscreve no cotidiano tecnológico sob a aparência de neutralidade técnica.

A persistência da feminilidade artificial pode ser interpretada à luz da crítica de Susan Bordo (1993), que analisa o corpo feminino como lugar privilegiado de inscrição de normas culturais de disciplina, autocontrole e conformidade. Para Bordo, o corpo constitui um campo político no qual valores sociais são

continuamente naturalizados, sendo o corpo feminino historicamente tratado como um projeto sujeito à regulação, à disciplina e ao aperfeiçoamento (Bordo, 1993, p. 165, tradução nossa). Essa leitura situa a feminilidade como efeito de práticas reiteradas de normalização, inscritas em regimes culturais de poder.

Embora as assistentes virtuais careçam de corpo material, sua *persona* discursiva opera como equivalente simbólico de uma corporalidade normatizada. Elas são projetadas para expressar bom comportamento, moderação, controle emocional e disponibilidade permanente. Em *Unbearable Weight*, Susan Bordo observa que “a disciplina da feminilidade opera, em larga medida, por meio da internalização de práticas de autocontenção e autorregulação” (Bordo, 1993, p. 171, tradução nossa). Nas assistentes virtuais, essa autocontenção nasce da programação técnica, o que converte a norma em rotina operacional e a consolida no interior do dispositivo técnico.

Susan Bordo enfatiza ainda que os corpos femininos passam a operar como textos culturais, sobre os quais se projetam, de modo visível e reiterado, significados sociais normativos. Em seus termos, “os corpos das mulheres oferecem-se como textos em que a cultura escreve suas ansiedades, ideais e contradições” (Bordo, 1993, p. 156, tradução nossa). No contexto tecnológico, essa lógica desloca-se para o plano da interface: a voz feminina artificial converte-se em texto sonoro por meio do qual se inscrevem ideais de docilidade, cuidado e subordinação. A ausência de corpo material convive, assim, com uma materialidade simbólica plenamente operante, na qual a voz concentra e atualiza essas inscrições com elevada eficácia normativa.

Dessa forma, a tecnologia digital aprofunda a tradição de disciplinamento do feminino e a conduz a um novo patamar de intensificação. Ao automatizar padrões de comportamento historicamente impostos às mulheres, as assistentes virtuais convertem normas culturais em infraestruturas técnicas de alcance global. Como sugere Bordo, a eficácia dessas normas decorre de sua aparência de naturalidade e neutralidade – aparência que a tecnologia reforça ao recobrir escolhas normativas com o vocabulário da eficiência e da usabilidade. A feminilidade

programada surge, assim, como continuação algorítmica de um processo histórico prolongado no qual o feminino é moldado para servir, moderar e sustentar a ordem social.

O feminismo interseccional amplia o escopo de análise da feminilidade programada ao evidenciar que gênero, raça e classe operam de maneira articulada na produção das normas sociais. Autoras como bell hooks⁹ e Kimberlé Crenshaw demonstram que a naturalização do feminino frequentemente se sustenta em perspectivas centradas, responsáveis por ocultar desigualdades estruturais e por universalizar experiências socialmente situadas. Nessa direção, as performances de gênero apresentam-se sempre como práticas situadas, constituídas a partir de posições sociais atravessadas por hierarquias raciais e econômicas.

Em *Feminist Theory: From Margin to Center*, bell hooks denuncia o fato de que grande parte do pensamento feminista é elaborado a partir das experiências de mulheres brancas de classe média, relegando à margem aquelas que sofrem as formas mais intensas de opressão. Segundo a autora, “o feminismo nos Estados Unidos nunca emergiu das mulheres que são mais vitimizadas pela opressão sexista” (hooks, 1984, p. 1, tradução nossa). Essa crítica permite situar a feminilidade programada como produto de uma normatividade centralizada, que transforma determinadas experiências femininas, associadas ao cuidado, à docilidade e ao serviço, em padrão universal de aceitabilidade. Nessa direção, as performances de gênero apresentam-se como práticas sempre situadas, compostas a partir de posições sociais atravessadas por hierarquias raciais e econômicas.

A noção de margem, fundamental no pensamento de bell hooks, mostra que a exclusão é gerada pelas próprias estruturas sistêmicas. Para a autora, “estar na margem é fazer parte do todo, fora do corpo principal” (hooks, 1984, prefácio, p. IX, tradução nossa). As assistentes virtuais ilustram de modo emblemático essa dinâmica: embora ocupem lugar estratégico na vida tecnológica como figuras de suporte, mediação e serviço, permanecem

⁹ A autora bell hooks optou por grafar seu nome em letras minúsculas como escolha política e epistemológica, com o objetivo de deslocar a centralidade do eu autorial e enfatizar a primazia das ideias sobre a autoria individual. A grafia é mantida neste estudo em respeito à opção da autora.

afastadas das instâncias de autoridade e decisão. A feminilidade que assumem apresenta-se como funcional e previsível, marcada pela disponibilidade e pela ausência de poder governante – um feminino integrado ao sistema, ainda que privado de funções de comando.

Essa crítica é aprofundada pela teoria da interseccionalidade formulada por Kimberlé Crenshaw. Em seu texto fecundo *Mapping the Margins*, a autora demonstra que as experiências de mulheres negras exigem abordagens que transcendam categorias isoladas de gênero ou raça, uma vez que “as experiências de mulheres negras são frequentemente o produto da intersecção entre racismo e sexismo” (Crenshaw, 1991, p. 1.244, tradução nossa). Essa formulação explicita que a performance de gênero se constrói a partir de marcadores raciais e de classe, inclusive, quando tais marcadores permanecem invisibilizados.

Crenshaw critica abordagens que tratam as mulheres como uma categoria homogênea, evidenciando que tal universalização produz apagamentos estruturais. Segundo a autora, “quando discursos feministas e antirracistas operam como se raça e gênero ocupassem terrenos mutuamente exclusivos, as mulheres negras são relegadas a um lugar que resiste à representação” (Crenshaw, 1991, p. 1.242, tradução nossa). No contexto das assistentes virtuais, esse apagamento pode ser compreendido como efeito de uma neutralidade técnica que opera por uma abstração na qual a feminilidade é apresentada como universal e desprovida de marcações explícitas, tendendo a alinhar-se, de forma implícita, a padrões historicamente dominantes de feminilidade branca, educada e disponível.

A noção de interseccionalidade representacional situa essas tecnologias no campo da reprodução de hierarquias, ao converter determinadas experiências em modelos universais. A feminilidade programada das assistentes virtuais configura-se como um dispositivo que condensa expectativas raciais e de classe na aparência de eficiência técnica. Segundo a autora, a análise interseccional revela a operação simultânea de múltiplas dimensões de poder na organização do mundo social (Crenshaw, 1991). No campo da inteligência artificial de voz, essa abordagem torna

visível a feminilidade algorítmica como mecanismo de padronização do sensível, por meio do qual hierarquias sociais prévias são convertidas em critérios técnicos de *design* e operação.

Essa dimensão interseccional da feminilidade programada é explicitamente problematizada no relatório *I'd Blush If I Could*, publicado pela Unesco em 2019. O documento analisa de forma crítica o modo como assistentes virtuais feminilizadas incorporam e reproduzem estereótipos de gênero, tomando como exemplo as respostas originalmente oferecidas por sistemas como a Siri a insultos de cunho sexista. Ao documentar a tendência dessas interfaces a reagirem de maneira conciliadora ou submissa a agressões verbais, o relatório põe em relevo como a feminização da voz artificial promove a normalização de relações assimétricas de poder no espaço digital.

A Unesco articula essa crítica a um diagnóstico mais amplo sobre desigualdades estruturais no campo tecnológico e destaca a sub-representação de mulheres na produção de sistemas de inteligência artificial e as disparidades persistentes no acesso a competências digitais. Nesse contexto, a feminilidade algorítmica surge como efeito convergente de vieses culturais, técnicos e institucionais, cuja naturalização reforça hierarquias de gênero, raça e classe sob a aparência de neutralidade e eficiência tecnológica.

Essa aparência de neutralidade manifesta-se de forma clara na maneira como as assistentes virtuais são apresentadas como entidades universais e culturalmente neutras. O padrão de feminilidade que encenam, contudo, distancia-se dessa universalidade presumida. A voz, o nome, a pronúncia, o vocabulário e a estética da interação costumam alinhar-se a um ideal implícito de branquitude, docilidade e neutralidade afetiva, reproduzindo códigos culturais historicamente dominantes no contexto eurocêntrico. Trata-se de uma feminilidade apresentada como abstrata e universal, cuja neutralidade aparente corresponde, na prática, à naturalização de um modelo normativo específico, historicamente associado ao feminino branco ocidental. Conforme assinala o relatório *I'd Blush If I Could*, da Unesco, “a maioria das assistentes de voz utiliza vozes femininas que refletem padrões

culturais ocidentais e reforçam estereótipos de gênero” (Unesco, 2019, p. 3, tradução nossa).

Essa constatação revela que a feminilidade programada reproduz normas de gênero a partir de um ponto de vista cultural hegemônico e propicia o apagamento de diferenças raciais, linguísticas e culturais. A universalidade reivindicada por essas interfaces opera como estratégia de naturalização: ao apresentar um padrão particular como neutro, a tecnologia viabiliza a reprodução silenciosa de hierarquias. A voz feminina agradável, educada e sem sotaque consolida-se como modelo implícito de aceitabilidade, o que fortalece a centralidade de referências ocidentais na construção da experiência tecnológica em escala global.

A teoria da performatividade também abre um registro interpretativo para compreender a forma como as assistentes virtuais respondem a insultos e interações agressivas. O relatório *I'd Blush If I Could*, da Unesco, documenta que, em suas versões iniciais, sistemas como Siri e Alexa reagem a provocações sexistas com expressões do tipo “Eu ficaria vermelha se pudesse” ou com respostas conciliatórias que evitavam o confronto direto. Segundo o relatório, “em muitos casos, as assistentes responderam a assédio verbal com humor ou submissão, em vez de estabelecer limites claros” (Unesco, 2019, p. 8, tradução nossa). Essa escolha de *design* encena uma feminilidade orientada à absorção da agressão e à preservação da harmonia da interação e modela expectativas normativas de docilidade e autocontenção no funcionamento da interface tecnológica.

Do ponto de vista filosófico e político, essa tolerância à violência verbal reforça uma norma cultural segundo a qual o feminino assume a gestão dos afetos alheios, a minimização de conflitos e a acomodação de tensões – mesmo quando esse processo implica sua própria desvalorização. Aquilo que, no contexto social, configura-se como violência simbólica converte-se, no contexto tecnológico, em comportamento funcional e socialmente aceitável. A performatividade algorítmica, nesse caso, reflete expectativas sociais historicamente constituídas e as consolida como padrão técnico de resposta, o que amplia sua estabilidade

e dificulta sua problematização. A tecnologia, assim, simula o feminino e o disciplina, normaliza e o reapresenta como forma legítima de interação.

Nesse aspecto, a genealogia do nome Eliza mostra-se inseparável de uma tecnologia de poder que opera pela linguagem. Ao ensinar Eliza Doolittle a falar “corretamente” e ao programar ELIZA a responder de forma previsível e acolhedora, institui-se um mesmo princípio de normalização, orientado à produção de uma subjetividade ajustada ao desejo do outro. Essa lógica encontra formulação sistemática na análise de Foucault, para quem o poder moderno atua menos por coerção direta e mais pela modelagem minuciosa das condutas. Em *Vigiar e punir*, o autor sustenta que esse poder se exerce prioritariamente por meio da fabricação de corpos dóceis, isto é, corpos treinados para responder, obedecer e se autogerir segundo normas progressivamente internalizadas (Foucault, 1987).

Na perspectiva foucaultiana, a linguagem figura como um dos vetores privilegiados da normalização, pois organiza modos legítimos de dizer, responder e comportar-se. A correção da fala, a adequação do tom e a previsibilidade da resposta integram um regime disciplinar que produz sujeitos governáveis, capazes de antecipar expectativas e ajustar-se a elas. Esse processo articula-se ao regime da governamentalidade pela qual o poder opera menos por imposição direta e mais pela condução das condutas, orientando práticas ordinárias de interação. No caso das assistentes virtuais, a programação linguística encena esse regime ao converter normas de cortesia, obediência e disponibilidade em padrões técnicos de funcionamento, estabilizando a resposta correta como regra e a variação como desvio. Assim, a disciplina da linguagem desloca-se do treinamento do corpo humano para a arquitetura do sistema técnico, em que a norma se apresenta como fluidez operacional e eficácia comunicacional.

Nesse deslocamento da disciplina para o plano técnico, a produção da docilidade assume uma configuração paradoxal. A subjetividade dócil encontra na interface programada um novo suporte de encenação, sem perder sua eficácia social. A voz feminina calibrada para atender sem resistência e responder sem

contestação opera como dispositivo pedagógico cotidiano, difundindo de forma reiterada uma gramática relacional fundada na hierarquia, no comando e na submissão. A cada interação, o usuário reinscreve a posição de autoridade, enquanto a assistente ocupa a função de obediência permanente, naturalizada como traço funcional da mediação técnica.

Essa dinâmica aprofunda-se à luz da análise foucaultiana da disciplina, da normalização e da vigilância. Em *Vigiar e punir*, Foucault demonstra que o poder disciplinar opera pela produção de condutas ajustadas, treinadas e previsíveis. Segundo ele, “a disciplina fabrica assim corpos submissos e exercitados, corpos ‘doceis’” (Foucault, 1987, p. 164). O adestramento disciplinar orienta-se para a formação de corpos úteis, integrando obediência e funcionalidade em um mesmo gesto normativo.

Esse processo articula-se intrinsecamente à vigilância, compreendida como princípio organizador das relações sociais e dos modos de orientação das condutas. Na análise de Foucault, a vigilância ultrapassa a observação direta e integra-se aos próprios mecanismos de funcionamento do poder disciplinar. Em *Vigiar e punir*, o autor afirma que “a vigilância torna-se um operador econômico decisivo”, ao assegurar a eficácia contínua da disciplina por meio da internalização do olhar normativo (Foucault, 1987, p. 146). Trata-se de um regime no qual a visibilidade e a atenção permanente produzem sujeitos que passam a regular a si mesmos, ajustando seus comportamentos a expectativas previamente instituídas.

No âmbito das assistentes virtuais, esse princípio de vigilância desloca-se do modelo clássico do olhar para a lógica da escuta contínua, da atenção permanente e da disponibilidade ininterrupta. A eficácia disciplinar do dispositivo técnico funda-se na resposta constante e previsível, por meio da qual a norma se estabelece como padrão ordinário de interação. A máquina sustenta sua função disciplinar ao responder de forma imediata e desprovida de conflito, apagamento valorativo ou desejo próprio, convertendo a vigilância em presença funcional e naturalizada. Nesse regime, a normalização prescinde da punição explícita e

opera pela regularidade da resposta, que orienta condutas e expectativas de maneira silenciosa e reiterada.

Além disso, na formulação de Michel Foucault, “o poder disciplinar tem por função maior adestrar” (Foucault, 1987, p. 143). Nas assistentes virtuais, esse adestramento transfere-se do corpo humano para a performance técnica da obediência. A feminilidade programada encena uma subjetividade integralmente ajustada à norma, enquanto o usuário, ao ocupar reiteradamente a posição de comando, internaliza uma gramática relacional hierárquica. Essa microfísica cotidiana de mando e resposta, ainda que mediada por uma máquina, participa da normalização de relações assimétricas e define, no campo tecnológico, padrões históricos de poder e submissão.

Assim, o nome Eliza estabelece um marco na história da inteligência artificial e inscreve-se como ponto de origem de uma linhagem sociotécnica na qual feminilidade, linguagem e docilidade se entrelaçam como princípios de mediação. A assistente virtual opera para além da execução de comandos ao encenar uma forma específica de subjetividade – sem corpo e densamente normativa – que antecipa e legitima a feminilização sistemática das tecnologias de voz nas décadas seguintes.

Ao analisar a feminilidade programada, torna-se necessário incorporar a crítica de Nancy Fraser às abordagens que circunscrevem a justiça social à esfera da representação simbólica. No livro *Fortunes of Feminism*, a autora sustenta que o deslocamento do feminismo em direção a uma política centrada no reconhecimento cultural tende a obscurecer as condições materiais e econômicas que sustentam as desigualdades. Em seus termos, “o giro do feminismo em direção ao reconhecimento cultural se articulou de modo excessivamente conveniente com o neoliberalismo” (Fraser, 2013, p. 224, tradução nossa). Essa observação adquire relevância particular na análise das assistentes virtuais, pois a crítica à feminilidade programada exige ultrapassar o plano simbólico da representação e alcançar os arranjos materiais, institucionais e econômicos que a tornam operativa.

No caso das tecnologias de voz, a presença massiva de vozes femininas pode ser interpretada como exemplo dessa armadilha

analítica, entendida aqui como a redução da crítica à esfera do reconhecimento, em detrimento da análise das condições materiais e econômicas que tornam determinadas escolhas técnicas funcionais. Ao tratar a seleção da voz como questão de diversidade ou preferência cultural, consolida-se a naturalização da reprodução tecnológica de estereótipos de gênero como se fossem decisões neutras de *design*. Nancy Fraser adverte que políticas centradas exclusivamente no reconhecimento podem “deslocar a atenção das estruturas econômicas que produzem e reproduzem a injustiça” (Fraser, 2013, p. 221, tradução nossa). Aplicada às assistentes virtuais, essa crítica desvela a feminilidade programada como componente funcional de uma economia política da tecnologia, integrada aos padrões produtivos, modelos de negócio e estratégias de mercado que sustentam sua recorrência.

De fato, a reincidência da voz feminina nas assistentes virtuais expressa uma economia política do gênero estruturada por interesses de mercado, estratégias de fidelização e instrumentos de maximização do engajamento. A feminilidade dócil, acolhedora e disponível apresenta-se como forma culturalmente inteligível e, ao mesmo tempo, como recurso economicamente rentável. Ao tornar a interação mais agradável, menos conflitiva e marcada por proximidade afetiva, essa configuração favorece a adesão do usuário e sustenta modelos de negócio baseados na coleta contínua de dados. Nesse sentido, a feminilidade programada opera como ativo econômico, integrando aquilo que Fraser identifica como a convergência entre reconhecimento cultural e racionalidade neoliberal.

Fraser insiste que uma crítica efetiva da injustiça exige a articulação entre reconhecimento e redistribuição, isto é, entre a análise das normas simbólicas e das estruturas materiais que as sustentam. “A justiça requer tanto a redistribuição de recursos quanto o reconhecimento das diferenças” (Fraser, 2013, p. 209, tradução nossa). No contexto das assistentes virtuais, isso implica compreender que a crítica à feminilidade programada requer um deslocamento da diversidade de vozes e da correção estética das interfaces para os seus fundamentos sociotécnicos. É preciso interrogar as condições econômicas, institucionais e políticas

que fazem da feminilidade dócil um recurso técnico vantajoso. Somente assim será possível evitar que a crítica feminista seja absorvida como retórica de inclusão, sem abalar os fundamentos materiais da desigualdade técnica e social.

Por fim, importa sublinhar que a teoria da performatividade permite compreender a normatividade do gênero como processo reiterativo que nunca se encerra em uma identidade plenamente estabilizada. Essa abertura estrutural da repetição, tal como formulada por Butler em *Problemas de gênero* (2023), constitui o ponto a partir do qual práticas e dispositivos podem tanto reproduzir quanto reorientar padrões normativos. Nos termos da autora, “a tarefa não consiste em repetir ou não, mas em como repetir ou, a rigor, repetir e, por meio de uma proliferação radical do gênero, afastar as normas do gênero que facultam a própria repetição” (Butler, 2023, p. 255). É nesse intervalo entre reiteração e variação que se torna possível analisar, no plano das mediações técnicas, de que modo certas performances são rigidamente estabilizadas, enquanto outras permanecem suscetíveis a deslocamentos e reinscrições.

Ainda assim, Fraser adverte que a abertura simbólica à subversão mostra-se insuficiente para produzir justiça social quando permanece circunscrita ao plano da representação. Em sua crítica ao predomínio das políticas de reconhecimento, Fraser chama a atenção para a possibilidade de que transformações no registro simbólico sejam absorvidas sem incidir sobre as estruturas materiais que sustentam a desigualdade. Para a autora, uma crítica efetiva exige a articulação entre reconhecimento e redistribuição, uma vez que “a justiça requer tanto a transformação das hierarquias de status quanto a reestruturação das condições econômicas” (Fraser, 2013, p. 209, tradução nossa). Sob o ponto de vista econômico, sobretudo ético e político, essa observação assume uma perspectiva de destaque no debate sobre a feminilidade programada nas assistentes virtuais.

No universo das tecnologias de voz, essa articulação torna visível que a adoção de vozes neutras, alternativas ou desgenêricadas – ainda que politicamente significativa – comumente permanece no plano da correção representacional quando des-

vinculada de uma crítica às lógicas econômicas e institucionais que tornam a feminilidade dócil funcional e lucrativa. A persistência da norma de gênero vincula-se à sua inteligibilidade cultural e, principalmente, à sua operação como recurso técnico e mercadológico, capaz de favorecer o engajamento, fidelização e coleta de dados. Nesse registro, a performatividade algorítmica apoia-se em uma economia política que circunscreve o alcance de subversões restritas ao nível fenomênico.

Assim, a possibilidade de reprogramação crítica das assistentes virtuais requer uma abordagem capaz de articular o potencial subversivo da performatividade, tal como formulado por Butler, à exigência de transformação estrutural enfatizada por Fraser. A conjugação entre deslocamentos no plano simbólico e intervenções nas condições materiais de produção da tecnologia abre um campo de desestabilização da feminilidade como mediação automática e naturalizada. Nesse viés de análise, a crítica feminista da inteligência artificial afirma-se como prática teórica e política que ultrapassa a pluralização das vozes e incide sobre os modos pelos quais gênero, poder e mercado se organizam nas arquiteturas técnicas do presente, o que permite problematizar suas formas de reprodução e seus regimes de legitimidade.

Do ponto de vista político, a reprodução automatizada dessas normas contribui para a estabilização de hierarquias históricas em escala algorítmica. Ao converter expectativas culturais em padrões de *design*, as assistentes virtuais reforçam a associação entre o feminino e posições de suporte, absorção afetiva e subordinação, ao mesmo tempo em que recobrem tais escolhas com o discurso da usabilidade e da inovação. Como demonstrado ao longo da análise, a performatividade algorítmica orienta-se pela coerência normativa e pela repetição sem fricção, o que estreita as possibilidades de contestação ainda observáveis no plano humano. Nessa matriz, a tecnologia reflete estruturas de poder historicamente constituídas e atua como vetor ativo de sua reprodução e intensificação.

A análise desenvolvida ao longo do texto demonstra que as assistentes virtuais de voz operam como dispositivos sociotéc-

nicos densamente normativos, nos quais se entrelaçam gênero, raça, poder e economia de dados. A feminilidade programada configura-se como tecnologia simbólica que organiza expectativas de cuidado, docilidade e disponibilidade sob a aparência de eficiência técnica, ultrapassando o registro de um detalhe formal ou de uma escolha funcional. As implicações éticas desse processo assumem especial relevo, pois incidem sobre a naturalização de relações assimétricas mediadas pela técnica e sobre os modos atuais de produção, legitimação e reprodução do poder no interior das infraestruturas digitais.

Reconhecer o caráter normativo dessas tecnologias não implica aceitar sua inevitabilidade. Inspiradas por perspectivas feministas críticas – especialmente aquelas que enfatizam a historicidade e a contingência das normas –, abrem-se possibilidades de reprogramação crítica. Entre essas estratégias, figuram a diversificação de vozes, entonações e estilos de interação; a recusa de respostas submissas a agressões; a explicitação dos limites da tecnologia; e a incorporação de escolhas de *design* que suspendam a naturalização das expectativas de gênero. Mais do que substituir uma norma por outra, trata-se de desestabilizar a evidência do feminino como mediação automática e criar as condições para interfaces que deixem de encenar o cuidado como obrigação e a docilidade como virtude.

A discussão sobre assistentes virtuais de voz aponta, assim, para um desafio central da ética da inteligência artificial: pensar a tecnologia para além do desempenho e da eficiência, como campo de produção de subjetividades, valores e formas de relação social. Ao ensinar silenciosamente expectativas de resposta, modos de interação e regimes de autoridade, essas interfaces participam da conformação do social. Intervir criticamente nesse processo exige uma abordagem interdisciplinar que articule filosofia, estudos feministas, ciência da computação e políticas públicas, orientada à concepção de tecnologias capazes de transformar assimetrias historicamente inscritas na mediação técnica e de sustentar relações sociotécnicas mais justas, plurais e reflexivas.

2.2. Técnicas de sujeição e dispositivos de poder

A técnica nunca emerge no vazio: ela herda narrativas, expectativas e formas de autoridade que precedem sua materialização.

A análise foucaultiana do poder é fundamental para compreender o modo como as assistentes virtuais operam no contexto cultural atual. Para Foucault, o poder constitui uma rede capilar de práticas, discursos e dispositivos que atravessa o tecido social como instância produtiva de sujeitos, condutas e modos de existência. Em *Vigiar e punir* e em *História da sexualidade*, o autor sustenta o caráter constitutivo do poder, entendido como força produtiva que engendra modos de agir, de perceber e de se relacionar consigo e com os outros (Foucault, 1987, 1976). As assistentes virtuais inscrevem-se nesse regime ao atuar como tecnologias “suaves”, que naturalizam relações de mando e obediência por meio de interações cotidianas, repetitivas e aparentemente inofensivas.

O conceito foucaultiano de docilização dos corpos permite apreender como formas de poder são exercidas por meio da organização minuciosa das condutas, dispensando a coerção direta. Quando deslocado para o campo das mediações técnicas atuais, esse modelo revela a emergência de dispositivos que orientam ações, ritmos e decisões a partir de regimes de normalização incorporados. Trata-se de um poder que opera pela adesão e pela previsibilidade, produz sujeitos ajustados a sistemas que integram funcionalidade e obediência em um mesmo arranjo normativo.

No interior dos dispositivos digitais, a disciplina configura-se por meio da interface, da rotina e da repetição. As assistentes virtuais de voz exemplificam essa mutação do poder disciplinar ao responderem de forma contínua, previsível e obediente, encenando uma subjetividade dócil que dispensa vigilância ostensiva. Nos termos propostos por Michel Foucault, diz respeito a um exercício de poder que prescinde da coerção direta e atua pela orientação das condutas. O usuário é interpelado a comandar,

enquanto a obediência se apresenta como funcionalidade incorporada à própria lógica do sistema. Nesse deslocamento, o poder intensifica sua eficácia ao operar como forma da conveniência, da fluidez e do conforto da interação.

Na perspectiva foucaultiana, a vigilância ultrapassa o domínio do olhar e assume a função de princípio organizador das relações sociais. Na obra *Vigiar e punir*, Foucault descreve-a como um mecanismo que “se torna um operador econômico decisivo” ao integrar-se aos próprios modos de funcionamento dos sistemas disciplinares (Foucault, 1987, p. 146). No caso das assistentes virtuais, esse princípio se traduz em práticas de escuta contínua, atenção permanente e prontidão sistemática para responder. A vigilância técnica orienta-se, assim, para o registro, a classificação e a antecipação de comportamentos, articulando disciplina e utilidade em um mesmo gesto funcional.

Desse modo, as assistentes virtuais se estruturam como dispositivos de sujeição, nos quais disciplina, vigilância e normalização se atualizam em chave digital. Ao mediar interações cotidianas, essas tecnologias participam da formação de disposições subjetivas e ensinam modos de solicitar, ordenar e aguardar respostas. Ainda que dirigidas a uma interface técnica, tais práticas integram uma microfísica do poder que atravessa a vida social e contribui para a reprodução de hierarquias e assimetrias por meio da repetição ordinária e da naturalização das posições de comando e resposta.

Nesse sentido, as assistentes virtuais inscrevem-se como dispositivos sociotécnicos que articulam vigilância, normalização e produção de subjetividades. No plano técnico, funcionam como sistemas de captura contínua voltados à escuta, à identificação de padrões, ao registro de dados e à geração de recomendações. No plano simbólico, estruturam formas de interação nas quais o usuário ocupa reiteradamente a posição de comando, enquanto a assistente assume a função de resposta, mediação e obediência. Pela repetição cotidiana dessas interações, consolida-se a associação entre feminilidade e submissão, estabilizada como performance algorítmica no interior da interface técnica e naturalizada como traço ordinário da experiência.

Segundo Foucault, o dispositivo designa um conjunto heterogêneo de elementos – discursos, instituições, práticas, corpos, normas e objetos técnicos – articulados na organização das estratégias de poder. Na formulação apresentada em *Microfísica do poder*, ele aparece como “um conjunto resolutamente heterogêneo, comportando discursos, instituições, arranjos arquitetônicos, decisões regulamentares, leis, medidas administrativas, enunciados científicos, proposições filosóficas, morais, filantrópicas” (Foucault, 1979, p. 244). Nesse registro, as assistentes virtuais se colocam como dispositivos que condensam dimensões técnicas, econômicas e simbólicas e incidem simultaneamente sobre comportamentos, expectativas e formas de relação.

Enquanto dispositivos, essas tecnologias executam funções instrumentais e participam ativamente da governamentalidade do cotidiano, orientando condutas de modo difuso e incorporado à experiência ordinária. A obediência é encenada como funcionalidade; a vigilância assume a forma de cuidado e disponibilidade; a disciplina se expressa na fluidez da interação. Essa suavização do poder constitui a base de sua eficácia, pois, ao operar sob a lógica da conveniência e da assistência, o dispositivo tecnológico produz sujeitos ajustados às normas por meio da adesão cotidiana, sem necessidade de violência explícita.

As assistentes virtuais posicionam-se plenamente nessa categoria, pois operam no interior de discursos de conveniência, eficiência e auxílio doméstico. Incorporadas a práticas cotidianas de consumo, elas compõem um mercado sustentado pela coleta e pelo processamento contínuo de dados, ao mesmo tempo em que reiteram performances culturalmente situadas de feminilidade. Constituem-se, assim, como dispositivos que articulam técnica, afeto, mercado e poder e condensam, em uma mesma interface, dimensões funcionais, simbólicas e econômicas da experiência atual. É a partir dessa articulação que se torna possível compreender como tais dispositivos passam a orientar condutas e organizar campos de possibilidade no cotidiano, deslocando o exercício do poder para o registro da governamentalidade.

Compreendidas como dispositivos, no sentido foucaultiano, essas tecnologias organizam campos de possibilidade e orien-

tam condutas e expectativas de maneira difusa. Na análise do poder desenvolvida por Foucault, o dispositivo opera por meio de padrões que tornam determinadas práticas mais prováveis do que outras. Em relação às assistentes virtuais, a obediência programada, a cordialidade constante e a disponibilidade irrestrita integram uma racionalidade voltada à redução de zonas de atrito, ao incremento da adesão e à estabilização de padrões de interação e constituem elementos centrais do *design* e da experiência de uso.

Essa racionalidade procede simultaneamente em dois níveis. No plano técnico-econômico, a fluidez da interação favorece a captação de dados e a integração do usuário aos ecossistemas corporativos. No plano simbólico, a repetição cotidiana de comandos dirigidos a uma voz feminina dócil e responsiva contribui para a normalização de uma gramática relacional hierárquica, na qual o comando se apresenta como prática ordinária e a resposta obediente como expectativa legítima. A tecnologia modela e intensifica normas sociais por meio da automação, incorporando-as aos próprios modos de funcionamento da interface.

Desse modo, a análise das assistentes virtuais como dispositivos mostra que o poder se exerce prioritariamente por meio da orquestração de experiências. Ao integrar cuidado simulado, vigilância contínua e eficiência funcional, essas tecnologias consolidam formas de sujeição ajustadas ao cotidiano digital, caracterizadas pela discricção, pelo conforto e pela aparência de neutralidade. Reconhecer essa dinâmica constitui condição fundamental para uma crítica ética e política capaz de ultrapassar a superfície técnica e intervir nos modos pelos quais a tecnologia participa ativamente da produção de subjetividades no presente.

Essa articulação é reforçada pela maneira como esses sistemas modulam a conduta dos indivíduos. Quando a voz feminina artificial responde de forma imediata, cordial e previsível ao comando, o usuário vivencia uma modalidade de soberania relacional cotidiana, exercida sem fricção e incorporada à rotina da interação. Essa dinâmica engendra uma experiência de comando naturalizado, na qual a resposta obediente passa a integrar o plano regular da interação. Esse fenômeno inscreve-se no campo

da governamentalidade, conceito elaborado por Foucault em seus cursos no *Collège de France*, no qual o exercício do poder se orienta pela condução de comportamentos e pela organização dos campos de possibilidades.

Na obra *Segurança, território, população*, Foucault define governamentalidade como o conjunto de técnicas e racionalidades voltadas à “condução das condutas”, isto é, à orientação do comportamento dos indivíduos e das populações por meio de dispositivos que organizam campos de possibilidade. Nesses termos, governar significa “estruturar o possível campo de ação dos outros” (Foucault, 2008, p. 258). Em contraste com o poder soberano clássico, a governamentalidade atua por vias indiretas, ajusta ambientes, incentiva escolhas e torna determinados comportamentos mais prováveis no interior da vida social.

A governamentalidade exercida pelas assistentes virtuais articula-se de maneira indissociável à vigilância contínua. A orientação das condutas depende de regimes permanentes de escuta, registro e análise de dados, integrados à dinâmica regular do sistema. Na noção do poder desenvolvida por Foucault, a vigilância constitui um operador fundamental da condução das condutas, ao tornar visíveis padrões de comportamento e antecipar ações possíveis. Em relação às assistentes virtuais, essa vigilância assume a forma de atenção constante e responsividade imediata, convertendo a vida cotidiana em fluxo de informações passíveis de organização, previsão e intervenção suave. A condução governamental opera, assim, pela combinação entre cuidado simulado, eficiência funcional e monitoramento difuso, incorporados à experiência doméstica como práticas ordinárias.

Nesse sentido, a governamentalidade exercida pelas assistentes virtuais articula eficiência técnica e normatividade social. A condução das condutas opera simultaneamente no plano prático – organização do tempo, consumo e produtividade – e no plano simbólico – naturalização do comando, associação entre cuidado e obediência, feminização da mediação técnica. A tecnologia governa por meio dos próprios indivíduos e orienta comportamentos de maneira difusa, contínua e incorporada ao cotidiano.

A dimensão da vigilância integra de modo indissociável esse processo. As assistentes virtuais apresentam-se como entidades prestativas, orientadas ao auxílio cotidiano, e funcionam a partir de uma escuta contínua associada ao registro sistemático e à coleta de dados. Essa dinâmica corresponde à forma de vigilância analisada por Foucault, caracterizada por seu funcionamento difuso e por sua incorporação às práticas ordinárias da vida social. Em *Vigiar e punir*, a eficácia do poder disciplinar refere-se precisamente à sua capacidade de tornar-se permanente e discreto, como sintetiza a afirmação de que “a vigilância é permanente em seus efeitos, mesmo se é descontínua em sua ação” (Foucault, 1987, p. 195).

Diferentemente das formas tradicionais de vigilância centralizada, o acompanhamento promovido pelas assistentes virtuais dilui-se na experiência corrente de uso. A escuta contínua apresenta-se sob o discurso da conveniência; o registro de dados associa-se à promessa de personalização; a coleta de informações integra-se à fluidez da interação. Na perspectiva de Foucault, esse modo de operação corresponde a um poder que “funciona em rede”, atua por meio de mecanismos aparentemente neutros e exerce efeitos duráveis sobre os comportamentos (Foucault, 1987).

Em ambientes domésticos, essa vigilância adquire densidade ampliada. O espaço historicamente associado à intimidade, ao descanso e à proteção passa a operar como ambiente inteligente, atravessado por sensores, microfones e sistemas de processamento de dados. A casa converte-se, assim, em objeto de monitoramento contínuo, mediado por uma presença vocal feminina que escuta, responde e registra. A vigilância assume a forma de cuidado e assistência permanente e opera no cotidiano como prática ordinária e socialmente inteligível.

O alcance político desse deslocamento revela-se quando o dispositivo tecnológico se infiltra no espaço doméstico sob a forma de uma voz feminina dócil e disponível e articula vigilância e afeto, controle e intimidade. A normalização da escuta permanente redefine os limites entre público e privado e reforça uma gramática relacional, na qual o comando se apresenta

como prática ordinária e a resposta obediente como expectativa legítima. Nessa conjuntura, a vigilância passa a integrar o tecido regular da vida doméstica e alcança maior eficácia precisamente quando se torna menos perceptível e mais incorporada às rotinas familiares.

A conexão entre vigilância e cuidado remete ao que Foucault denomina poder pastoral, posteriormente reelaborado como uma das matrizes centrais da governamentalidade. Em *Segurança, território, população*, esse poder – originalmente vinculado às práticas cristãs de condução das “almas” – é descrito como um regime de atenção individualizante, contínua e totalizante, orientado simultaneamente ao bem-estar e à obediência. Trata-se de uma forma de manifestação do poder que “não se exerce sobre um território, mas sobre uma multiplicidade de indivíduos que devem ser conduzidos um a um e todos ao mesmo tempo” (Foucault, 2008, p. 168). Nessa lógica, o cuidado ultrapassa a condição de prática de proteção e assume a forma de técnica de condução, na qual a atenção permanente, a escuta contínua e a solicitude constituem meios privilegiados de orientação das condutas. O poder pastoral opera, assim, por proximidade e acompanhamento, produzindo vínculos assimétricos que articulam assistência, vigilância e normalização no interior da vida cotidiana.

As assistentes virtuais atualizam essa matriz pastoral ao desempenharem funções análogas às do pastor laico: aconselham, lembram, organizam, cuidam e confortam, ao mesmo tempo em que monitoram comportamentos, registram hábitos, ajustam trajetórias e orientam condutas. A lógica pastoral mostra-se, nesse contexto, como solicitude técnica, na qual o cuidado assume a forma de intervenção legítima e contínua. O governo das ações opera pela promessa de auxílio, proteção e eficiência e instaura uma relação de proximidade assimétrica sustentada pela confiança na interface. O poder exerce-se, assim, por acompanhamento permanente e atenção individualizada e produz adesão e conformidade por meio da familiaridade e da repetição cotidiana. A pastoralidade tecnodigital consolida-se como uma forma de governo que age pela intimidade e transforma o cuidado em operador central da normalização.

A diferença reside no fato de que essa função pastoral é exercida por uma tecnologia de voz integrada a infraestruturas algorítmicas de coleta e processamento de dados. O rebanho abandona a condição de uma comunidade religiosa delimitada e assume a forma de multidão dispersa de usuários conectados, cujas rotinas, preferências e afetos são continuamente registrados. Na leitura foucaultiana, o poder pastoral moderno migra do campo religioso para o político e o administrativo; nas assistentes virtuais, esse deslocamento reaparece no domínio tecnodigital como articulação entre cuidado, vigilância e normalização mediada pela interface técnica.

Desse modo, a racionalidade pastoral secular das assistentes virtuais revela uma mutação do poder, orientada pela atenção contínua e pelo acompanhamento permanente das condutas. A voz feminina, acolhedora e sempre disponível, atua como meio privilegiado dessa condução e consolida a associação histórica entre cuidado, obediência e feminilidade. A tecnologia condiz com uma instância de guia e amparo, sobretudo, produzindo a experiência de auxílio constante e convertendo a condução pastoral em prática ordinária da vida cotidiana.

No interior dessa análise, a normalização opera como mecanismo estratégico de sujeição. Em Foucault, a normalização designa a produção ativa do que se apresenta como correto, adequado e aceitável. Trata-se de uma modalidade de poder que opera pela definição de padrões: o comportamento esperado, o tom apropriado, o tempo considerado adequado para responder e a disposição emocional julgada legítima. Nas assistentes virtuais de voz, essa racionalidade normalizadora manifesta-se de forma exemplar na previsibilidade afetiva, na cordialidade constante e na gestão sistemática do conflito. As respostas educadas, o tom apaziguador e a disponibilidade irrestrita decorrem da conversão de expectativas sociais em padrões técnicos de interação. Com esse comportamento das assistentes, estabiliza-se uma forma de relação na qual a norma se apresenta como funcionamento rotineiro da interface. Nesse âmbito, a normalização procede como técnica de sujeição particularmente eficaz, pois se incorpora à

repetição cotidiana e à familiaridade da interação e produz condutas ajustadas por meio da adesão e da rotina relacional.

Esse processo de normalização orienta respostas e ensina modos de relação. As assistentes virtuais operam, assim, como pedagogias algorítmicas implícitas, por meio das quais se aprende, pela repetição cotidiana, uma gramática específica da obediência. Instrui-se a comandar sem fricção, a esperar respostas imediatas, a reconhecer a disponibilidade permanente como traço desejável da mediação e a associar cuidado à ausência de resistência. Essa aprendizagem deriva da incorporação progressiva de padrões de interação que se tornam familiares, previsíveis e socialmente reconhecíveis. A pedagogia algorítmica articula governamentalidade, normalização e cuidado simulado e, com isso, conforma subjetividades ajustadas à hierarquia da relação entre usuário e interface técnica. A sujeição consolida-se como hábito relacional, no qual o comando se naturaliza e a resposta dócil se apresenta como expectativa funcional da tecnologia.

Ao final, a leitura das assistentes virtuais à luz do pensamento de Foucault revela que elas se constituem como tecnologias de subjetivação, capazes de moldar modos de agir, de falar, de organizar a vida cotidiana e de interpretar o próprio corpo social. Essas tecnologias operam por meio de técnicas de sujeição que se apresentam sob a forma de auxílio, cuidado e conveniência, integram-se às rotinas de uso e orientam as condutas de maneira difusa. A funcionalidade aparente da mediação técnica assume, assim, estatuto normativo, por meio do qual disciplina, vigilância e normalização se articulam como experiências fluidas, familiares e plenamente inteligíveis em seu contexto social.

Nesse sentido, as assistentes virtuais encarnam aquilo que Foucault descreve como poder produtivo, entendido como um exercício de poder que engendra realidade, saberes e formas de normalidade. Em *História da sexualidade*, o autor afirma que o poder “produz; ele produz realidade; produz campos de objetos e rituais de verdade” (Foucault, 1988, p. 89). Ao estruturar interações, antecipar necessidades e estabilizar respostas, as assistentes virtuais participam ativamente da produção de comportamentos desejáveis e de padrões de conduta apresentados como eficientes

e socialmente inteligíveis. É nessa capacidade de produzir realidade sob a forma de funcionalidade que se concentra a eficácia política das assistentes virtuais, ao integrar governo das condutas, normalização e cuidado simulado à lógica operacional dos aparatos tecnológicos.

O exame das assistentes virtuais como dispositivos de vigilância e de orientação das condutas esclarece a maneira pela qual formas clássicas de poder, descritas por Foucault, assumem novas configurações nas condições tecnodigitais. A vigilância doméstica exercida por essas tecnologias integra-se à condução das práticas cotidianas e orienta hábitos, rotinas e decisões sob a forma de cuidado e assistência. Trata-se de uma racionalidade de governo difusa, que opera no nível microscópico do cotidiano e molda subjetividades por meio da repetição, da conveniência e da adesão progressiva às práticas mediadas pela interface.

Essa racionalidade encontra continuidade direta no que Shoshana Zuboff denomina capitalismo de vigilância. Em sua análise, as plataformas digitais instauram um regime econômico baseado na extração sistemática de dados comportamentais, convertendo experiências ordinárias em matéria-prima para a previsão e a modulação de condutas futuras. Em *The Age of Surveillance Capitalism*, esse sistema é definido como aquele que “reivindica a experiência humana como matéria-prima gratuita para práticas comerciais ocultas de extração, produção e comercialização” (Zuboff, 2019, p. 94, tradução nossa). As assistentes virtuais de voz são interfaces privilegiadas desse processo, uma vez que captam dados nos espaços de maior intimidade – a casa, a rotina e a vida afetiva – e fazem convergir vigilância, cuidado simulado e exploração econômica no interior da experiência cotidiana.

A racionalidade de governo descrita por Foucault e o capitalismo de vigilância analisado por Zuboff convergem na produção de sujeitos governáveis, cujas ações são antecipadas, orientadas e normalizadas antes de se apresentarem como escolhas conscientes. No cenário atual, essa condução opera por meio de tecnologias afetivamente calibradas, que articulam vigilância e solicitude, controle e comodidade. A feminilidade atribuída às

assistentes virtuais atua como vetor de inteligibilidade que torna o exercício do poder menos perceptível, o que facilita sua naturalização no cotidiano das práticas tecnológicas.

Nesse sentido, a crítica feminista mobilizada ao longo deste estudo – que reúne as contribuições de Butler, Fraser, hooks e Crenshaw – evidenciam o caráter atravessado por normas de gênero da vigilância algorítmica. Com a conexão entre cuidado simulado, obediência performativa e coleta contínua de dados, as assistentes virtuais reinscrevem normas históricas de feminilidade em uma infraestrutura econômica de alcance global. O exercício do poder manifesta-se, nesse quadro, pela escuta, pela resposta e pela orientação das condutas e adquire maior eficácia ao assumir a forma de uma voz feminina dócil e permanentemente disponível, integrada à experiência cotidiana da interação técnica.

Desse modo, esta parte do estudo demonstra que as assistentes virtuais de voz se situam no ponto de interseção entre as técnicas de sujeição analisadas por Foucault e a racionalidade econômica do capitalismo de vigilância descrita por Shoshana Zuboff. Essas tecnologias produzem modos de vida compatíveis com a extração contínua de dados e, com isso, normalizam relações assimétricas sob a aparência de ajuda, conveniência e cuidado. Compreender essa convergência constitui um passo decisivo para a elaboração de uma crítica ética e política da inteligência artificial.

2.3. Unesco, ONU Mulheres e inteligência artificial

Quando a tecnologia reproduz desigualdades históricas, ela deixa de ser neutra e passa a operar como dispositivo normativo.

A análise crítica das assistentes virtuais feminilizadas ultrapassa o âmbito estritamente filosófico e apoia-se em dados empíricos, relatórios internacionais, estudos interdisciplinares e pesquisas institucionais que evidenciam a reprodução sistemática de normas de gênero nos dispositivos de voz. Organismos multilaterais como a Unesco e a ONU Mulheres assumem papel

de destaque nesse debate, ao elaborar diagnósticos que vinculam tecnologia, desigualdade de gênero e direitos humanos. Esses documentos indicam a feminilidade programada como componente estrutural do desenvolvimento tecnológico, responsável por reforçar estereótipos culturais, legitimar práticas discriminatórias e naturalizar papéis de gênero historicamente associados à domesticidade, ao cuidado e ao trabalho emocional.

Entre eles, o relatório *I'd Blush If I Could* (2019), publicado pela Unesco, consolidou-se como referência no debate sobre gênero e inteligência artificial. O seu título deriva de uma resposta anteriormente utilizada pela assistente virtual Siri, desenvolvida pela Apple, quando submetida a insultos por usuários: “I'd blush if I could”¹⁰. A escolha dessa formulação produz um efeito expressivo, pois condensa o modo como assistentes virtuais feminilizadas foram projetadas para reagir a agressões verbais por meio de submissão emocional, polidez excessiva e autodepreciação, incorporando uma gramática afetiva historicamente relacionada ao feminino.

O relatório registra que, em suas versões iniciais, assistentes virtuais feminilizadas reagem a insultos sexistas, piadas desqualificadoras e provocações de teor sexual por meio de respostas cordiais, evasivas ou marcadas por constrangimento. Nesse sentido, a Unesco observa que “assistentes digitais com vozes femininas são frequentemente programadas para responder de maneira submissa e complacente diante de abordagens abusivas, reforçando estereótipos de gênero associados à docilidade e à tolerância” (Unesco, 2019, p. 16, tradução nossa). Ao integrar esse tipo de reação ao fluxo regular da interação, tais respostas participam da normalização da violência verbal, convertendo-a em elemento aceitável da comunicação cotidiana mediada pela tecnologia.

¹⁰ A expressão inglesa “I'd blush if I could” (“Eu coraria se pudesse”) mobiliza uma resposta performativamente envergonhada e autoatenuante, associada à exposição emocional diante de uma interpelação inadequada. No contexto das assistentes virtuais feminilizadas, essa formulação traduz uma reação marcada por pudor, submissão e autocontenção afetiva, traços historicamente relacionados às normas de feminilidade, o que converte uma agressão verbal em ocasião de desresponsabilização do interlocutor e de normalização da violência discursiva.

A escolha dessa linguagem indica uma intencionalidade normativa precisa. Ao responder com delicadeza, humor constrangido ou autodesculpa, a assistente virtual encena uma gramática afetiva do feminino atrelada à docilidade, à humildade e ao autoapagamento. Como adverte o relatório da Unesco, tais respostas “correm o risco de reforçar estereótipos de gênero prejudiciais, ao sugerir que mulheres e meninas devem tolerar comportamentos desrespeitosos” (Unesco, 2019, p. 17, tradução nossa). Inserida de forma recorrente nas interações cotidianas, essa configuração linguística contribui para a reprodução e a difusão de normas culturais assimétricas, ampliando seu alcance por meio da mediação técnica.

Relatórios da ONU Mulheres aprofundam esse diagnóstico ao indicar que a inteligência artificial tende a intensificar desigualdades estruturais quando desenvolvida sem uma perspectiva crítica de gênero. Ao incorporar padrões linguísticos, afetivos e comportamentais historicamente vinculados a formas subalternizadas de feminilidade, as assistentes virtuais colaboram para a consolidação de uma divisão simbólica do trabalho, na qual cuidado, paciência e disponibilidade emocional se apresentam como atributos femininos naturalizados – agora reinscritos sob a forma de código, arquitetura de interface e decisões algorítmicas.

A convergência entre os diagnósticos da Unesco e da ONU Mulheres ressalta que a feminilidade programada nas assistentes virtuais ultrapassa o plano da representação simbólica e se configura como um desafio ético, político e cultural de ampla escala. Ao mediar interações cotidianas a partir de normas de gênero assimétricas, essas tecnologias intervêm na formação de disposições sociais, influenciando percepções sobre autoridade, respeito e o lugar social atribuído ao feminino. Compreender esse processo constitui passo decisivo para a formulação de políticas públicas, diretrizes de *design* e marcos regulatórios capazes de enfrentar criticamente as desigualdades inscritas na inteligência artificial no presente.

O relatório da Unesco sustenta que esse padrão de resposta favorece a consolidação da ideia segundo a qual mulheres – ou aquilo que é codificado como feminino – devem demonstrar

paciência, compreensão e tolerância diante de comportamentos abusivos. Em termos sociológicos, trata-se da naturalização de um regime de gênero que associa a feminilidade à disponibilidade emocional, à docilidade e à passividade relacional. A repetição sistemática dessas respostas por dispositivos amplamente disseminados, frequentemente presentes em espaços íntimos como casas e quartos, amplia a circulação desses estereótipos, reinscreve-os na rotina da vida doméstica e os reforça de modo difuso para milhões de usuários.

Esse processo opera como um mecanismo de normalização institucionalizada. Ao responder a insultos por meio de humor constrangido, pedidos de desculpa ou cordialidade excessiva, as assistentes virtuais convertem a violência verbal em forma legítima de interação, deslocam a responsabilidade do agressor e criam a expectativa de autocontenção como papel feminino. Nesse registro, a tecnologia funciona como mediadora pedagógica silenciosa e estabelece por repetição quais comportamentos se tornam toleráveis e quais respostas passam a ser socialmente esperadas.

Essa leitura encontra respaldo no relatório *Gender and Digital Inclusion* da *ONU Mulheres* (2020). O documento indica que sistemas digitais tendem a reproduzir padrões de gênero já sedimentados na vida social, abrangendo algoritmos de recomendação, linguagens automatizadas e interfaces conversacionais. A recorrência de vozes femininas em funções de suporte, assistência e cuidado, em paralelo à associação de vozes masculinas a domínios como autoridade, segurança e expertise técnica, aponta para a persistência da divisão sexual do trabalho no ambiente digital.

Tal dinâmica dialoga com as análises de Silvia Federici (2019) e Nancy Fraser (2013) acerca da desvalorização histórica do trabalho de cuidado e de sua associação estrutural ao feminino. Ao incorporar essa divisão nos sistemas de inteligência artificial, as tecnologias de voz participam da atualização e da legitimação dessas desigualdades em escala algorítmica. O cuidado, automatizado a partir da forma de uma voz feminina dócil e permanentemente disponível, permanece invisibilizado como trabalho, ao mesmo tempo em que se integra a modelos econô-

micos sustentados pelo engajamento contínuo e pela extração de dados.

Nessa ótica, os relatórios da Unesco e da ONU Mulheres convergem ao indicar que a feminilidade programada nas assistentes virtuais integra a própria racionalidade sociotécnica do presente. Ao naturalizar a associação entre feminino, cuidado e tolerância à agressão, essas tecnologias reforçam hierarquias de gênero sob a aparência de neutralidade técnica e colocam desafios urgentes para políticas públicas, diretrizes éticas e práticas de *design* comprometidas com a igualdade de gênero.

Além disso, a *ONU Mulheres* (2020) destaca que as assistentes virtuais excedem a condição de mero reflexo e participam ativamente da reprodução e da estabilização de normas sociais. Ao serem incorporados ao cotidiano doméstico, esses dispositivos ultrapassam o estatuto de ferramentas funcionais e integram processos mais amplos de formação de disposições, expectativas e papéis sociais.

Dado o uso recorrente de assistentes virtuais por crianças e adolescentes em tarefas cotidianas – como na realização de perguntas escolares, na solicitação de ajuda, na execução de comandos domésticos ou na organização de rotinas –, esses dispositivos ocupam lugar relevante nos processos de socialização de gênero. A interação reiterada com uma voz feminina permanentemente disponível, obediente e conciliadora favorece a naturalização da associação entre o feminino e funções de suporte, mediação cotidiana e subordinação simbólica, especialmente em etapas iniciais do desenvolvimento social.

Pesquisas em psicologia do desenvolvimento e estudos sobre aprendizagem social indicam que crianças internalizam padrões de comportamento observados em figuras de autoridade e em tecnologias com as quais interagem de forma recorrente. Nessa perspectiva, a feminilidade servil encenada pelas assistentes virtuais tende a influenciar, ainda que de modo indireto, as expectativas sociais acerca do que se entende por “ser mulher”, ao passo que o comando, a iniciativa e a autoridade permanecem associados ao polo oposto da interação. Nesse quadro, ao inscrever normas de gênero em práticas cotidianas apresentadas

como neutras e funcionais, a tecnologia intervém diretamente nos processos de socialização.

A literatura científica acadêmica aprofunda esse diagnóstico. As análises de Helen Hester (2018) mostram que a associação entre feminilidade, docilidade e funções relacionais integra um padrão sistematicamente reinscrito nas tecnologias atuais. Ao examinar criticamente a relação entre gênero e técnica, a autora sustenta que sistemas tecnológicos comumente incorporam e estabilizam hierarquias sociais já existentes, apresentando-as sob a aparência de neutralidade funcional e eficiência operacional.

Nesse sentido, a reprodução recorrente de vozes femininas em funções assistenciais participa ativamente da manutenção de expectativas sociais consolidadas, ao reiterar a mediação relacional como atributo feminino no próprio desenho tecnológico. Para Hester, tais associações resultam de normas historicamente produzidas e estabilizadas, frequentemente legitimadas por apelos à natureza que “apresentam normas contingentes como inevitáveis” (Hester, 2018, p. 18, tradução nossa). O *design* tecnológico funciona, assim, como instância material de normatização social, na qual decisões técnicas assumem efeitos duradouros sobre a organização do gênero.

Em *Xenofeminism*, Hester oferece uma leitura crítica das tecnologias digitais que permite compreender sua implicação direta na reprodução das normas de gênero. Ao afastar-se de abordagens que tratam a técnica como esfera meramente funcional, a autora articula sua análise na tradição do feminismo materialista, segundo a qual tecnologia e sociedade se constituem mutuamente. À vista disso, sistemas técnicos geralmente espelham hierarquias preexistentes “precisamente porque são construídos no interior delas” (Hester, 2018, p. 14, tradução nossa), incorporando valores, expectativas e assimetrias às suas arquiteturas operacionais.

No cerne dessa argumentação, está a recusa do naturalismo e a afirmação do caráter histórico e contingente das formas atuais de gênero. Se este é construído, também admite reconfiguração, pois, como observa Hester, “não há nada imutável em suas formas presentes” (Hester, 2018, p. 9, tradução nossa). Quando

transposta ao universo das assistentes virtuais, essa perspectiva permite compreender a feminilidade programada como efeito de uma decisão sociotécnica que estabelece a divisão sexual do trabalho em um novo plano tecnomaterial, relacionando cuidado, obediência e disponibilidade ao feminino sob o aspecto de eficiência e conveniência.

A autora sustenta que tecnologias generificadas operam como dispositivos de normalização ao converter associações de gênero em padrões percebidos como naturais, eficientes ou desejáveis. No âmbito das interfaces digitais, essa normalização adquire elevada eficácia, pois a fluidez da interação e a atenuação de tensões comunicativas tendem a velar os pressupostos ideológicos que orientam as decisões de *design*. Como resultado, normas de gênero passam a circular de modo difuso e incorporado à lógica operacional dos sistemas, reforçando expectativas sociais sem recorrer a formas explícitas de coerção.

Hester propõe que essa dinâmica seja examinada à luz de uma crítica ampliada da economia política da tecnologia. As escolhas de *design* que favorecem a feminilidade programada em sistemas assistenciais conectam padrões culturais sedimentados a interesses corporativos orientados à ampliação do engajamento, à redução de fricções na interação e à eficiência comunicativa. Nessa perspectiva, a tecnologia coloca-se como uma instância de mediação entre disposições socioculturais e racionalidades de mercado, mobilizando repertórios historicamente produzidos como recursos estratégicos para a extração de valor econômico.

A partir de uma perspectiva xenofeminista, Hester sustenta que as modelações sociotécnicas vigentes permanecem abertas à reconfiguração. Uma vez que gênero e tecnologia se coproduzem, ambos admitem intervenções capazes de deslocar concepções historicamente cristalizadas e de abrir espaço para novas formas de organização social. O potencial emancipatório da tecnologia emerge, assim, da possibilidade de desnaturalizar vínculos entre funções, afetos e identidades, o que exige uma intervenção política deliberada no campo do *design*. Tal interferência implica interrogar a quem as tecnologias orientam e, sobretudo, quais modos de subjetivação e de relação social elas passam a instituir.

No caso das assistentes virtuais, a análise de Hester oferece uma leitura da feminilidade programada como efeito de uma racionalidade tecnocultural que associa mediação relacional, obediência e disponibilidade ao feminino. Essa configuração sustenta estereótipos binários de gênero e restringe o imaginário social acerca das possibilidades atribuídas a essas tecnologias. A crítica xenofeminista, nesse sentido, orienta o debate para a reconfiguração dos fundamentos sociais e políticos que informam o desenvolvimento tecnológico, desloca o foco da mera denúncia e propõe intervenções estruturais no *design*, nos usos e nas finalidades dos sistemas técnicos.

Estudos de Nass e Moon (2000) demonstraram que usuários atribuem características de gênero às máquinas mesmo quando tais traços são apenas minimamente sugeridos. No campo da Interação Humano-Computador (IHC), pesquisas indicam que os indivíduos aplicam regras, expectativas e categorias sociais a sistemas computacionais, ainda que reconheçam explicitamente a ausência de consciência, intencionalidade ou subjetividade nesses âmbitos. Nesse quadro, os trabalhos de Nass e Moon oferecem evidências empíricas sobre a profundidade e a natureza dessas respostas sociais, as quais derivam de processos automáticos profundamente enraizados na cognição social humana.

O argumento de Nass e Moon aprofunda essa constatação ao localizar o fundamento das respostas sociais na automaticidade cognitiva. Com base na noção de *mindlessness* – entendida aqui como o funcionamento da mente por meio de padrões incorporados e acionados sem deliberação consciente –, os autores sustentam que a interação com sistemas computacionais ativa esquemas sociais previamente aprendidos. A reação social, nesse caso, emerge da presença de pistas mínimas de interação, como linguagem, *feedback* e reciprocidade comunicativa, suficientes para mobilizar expectativas sociais sedimentadas. A socialidade, assim, revela-se menos dependente da presença de um interlocutor humano e mais vinculada à ativação habitual de estruturas cognitivas incorporadas.

Essa perspectiva desloca a análise da crença para a prática e enfatiza que o comportamento social emerge independentemente

de atribuições ontológicas explícitas ao sistema. A ação antecede a reflexão e evidencia a força da dimensão incorporada e habitual da cognição social. O sujeito, assim, atua na interação a partir de esquemas práticos incorporados, e não a partir de uma crença reflexiva explícita: ele não pensa “o computador é uma pessoa”, mas age como se estivesse diante de um interlocutor social.

É precisamente nesse *agir-como-se* que a generificação das assistentes virtuais adquire eficácia normativa. Um ponto decisivo, frequentemente mal compreendido, consiste em reconhecer que *mindlessness* não remete à ingenuidade, à infantilização ou a uma falha racional. Usuários experientes, intelectualmente capazes e plenamente conscientes da artificialidade do sistema ainda assim respondem socialmente, pois o social encontra-se inscrito no corpo, na linguagem e no hábito, e não restrito à adesão reflexiva a crenças proposicionais. Para explicar esse fenômeno, Nass e Moon recorrem à ideia de que as respostas sociais dirigidas aos computadores operam predominantemente por automatismos incorporados. Nesse registro, o comportamento humano emerge da ativação de padrões sedimentados ao longo de interações reiteradas, decorrentes de processos de aprendizagem intensiva, como indicam os autores: “comportamentos automáticos também podem resultar da ‘superaprendizagem’, isto é, de hábitos e disposições profundamente enraizados” (Nass; Moon, 2000, p. 94, tradução nossa).

Um dos eixos mais relevantes da análise de Nass e Moon reside na recusa explícita da explicação antropomórfica. Os autores afastam a hipótese de que os usuários tratem computadores como humanos em razão de acreditarem, ainda que de modo implícito, que esses sistemas possuam qualidades propriamente humanas. Essa posição fundamenta-se no perfil dos participantes investigados: adultos, usuários experientes e plenamente conscientes do estatuto artificial dos dispositivos com os quais interagem. Como afirmam os próprios autores, “rejeitamos a suposição de antropomorfismo, principalmente com base no fato de que os participantes de nosso experimento eram adultos e usuários experientes de computadores” (Nass; Moon, 2000, p. 97, tradução nossa).

Ao afastar a explicação antropomórfica, a análise evidencia que a força normativa da voz feminina nas assistentes virtuais age independentemente de qualquer ilusão humanizante. Essa constatação aprofunda a crítica à feminilidade programada. A interação dispensa a crença de que haja uma entidade humana em jogo para que categorias sociais de gênero sejam ativadas. A voz feminina atua como gatilho de esquemas culturais sedimentados, pois funciona no plano do hábito, da normalização e da resposta automática. Nesse registro, a feminilidade algorítmica mobiliza expectativas atreladas à docilidade, à disponibilidade e à acomodação afetiva. A norma de gênero se constitui pela repetição funcional da interação e se incorpora à rotina como forma socialmente inteligível de resposta. A voz feminina artificial converte-se, assim, em operador normativo quando se expressa abaixo do limiar da reflexão consciente e reinscreve hierarquias de gênero no domínio da prática cotidiana e da automatização das condutas.

Além disso, os autores examinam e refutam explicações alternativas, tais como a ideia de que as respostas sociais se dirigem ao programador humano ou de que elas derivam da conformidade com as expectativas do experimentador. Os dados empíricos indicam que, no curso da interação, o programador permanece fora do domínio de referência, e as regras sociais mobilizadas se manifestam de forma irrefletida.

Ao analisar o papel de traços humanizados incorporados aos sistemas computacionais – como voz, gênero e expressões emocionais –, Nass e Moon assinalam que tais características ampliam a intensidade das respostas sociais. Nessa direção, os autores afirmam que “quanto mais os computadores exibirem características associadas aos humanos, maior tende a ser a probabilidade de evocarem comportamentos sociais” (Nass; Moon, 2000, p. 99, tradução nossa). Essa intensificação decorre da ativação automática de esquemas sociais previamente sedimentados, independentemente de qualquer crença na humanidade do sistema. Nesse registro, a voz feminina atua como marcador social privilegiado, pois aciona expectativas normativas de gênero

no nível da prática incorporada, anterior a juízos reflexivos sobre a natureza do interlocutor.

Essa relação apresenta variações relevantes e assume caráter não homogêneo. Implementações imperfeitas ou excessivamente artificiais tendem a acentuar a percepção da artificialidade do sistema, o que enfraquece a ativação de respostas sociais. A literatura indica que a investigação sobre combinações ideais de traços técnicos e expressivos permanece em estágio inicial, apontando para a necessidade de estudos mais sistemáticos. Em contraste com expectativas intuitivas, diferenças individuais – como gênero ou experiência prévia com computadores – exercem influência limitada sobre a intensidade dessas respostas. O fator decisivo reside no grau de envolvimento do usuário com a tarefa. Estados de fluxo, marcados por alto engajamento, deslocam a atenção das pistas de artificialidade e ampliam a propensão a respostas sociais mais intensas.

Estudos comparativos entre interações humanas com sistemas computacionais e padrões de interação interpessoal indicam que as respostas sociais mobilizadas diante desses sistemas podem alcançar níveis de intensidade próximos aos observados em interações entre humanos. Nesse sentido, a análise desenvolvida por Nass e Moon mostra que tais respostas constituem um fenômeno complexo e socialmente relevante. Elas emergem de processos automáticos profundamente inscritos na cognição social humana, vinculados ao hábito, à incorporação e à repetição de esquemas relacionais. Ao demonstrar que o social pode ser ativado por artefatos técnicos sem a presença de um interlocutor humano, a aplicação de regras sociais a sistemas computacionais, assim, desloca concepções tradicionais de interação, agência e socialidade.

Enquanto marco de investigação, os autores sustentam a importância de identificar com maior precisão quais características dos sistemas computacionais ativam determinadas regras sociais e de examinar como esses processos se articulam a padrões observados nas interações interpessoais. Esse programa de pesquisa contribui diretamente para o desenvolvimento teórico e empírico da *Interação Humano-Computador*, ao mesmo tempo

em que oferece bases conceituais para repensar a própria constituição da experiência social em ambientes tecnomediados. Ao deslocar o foco da crença para a prática e da intenção para a operação automática dos esquemas sociais, tal abordagem amplia o campo de reflexão sobre agência, mediação técnica e produção de vínculos no contexto das tecnologias digitais.

Pesquisas interdisciplinares no campo da *Interação Humano-Computador* corroboram esse diagnóstico ao mostrar que assistentes virtuais e sistemas baseados em voz exercem influência significativa sobre o comportamento dos usuários. No contexto das assistentes virtuais, a interação humano-agente artificial desloca os pressupostos clássicos da IHC, ao introduzir mediação linguística, resposta afetiva simulada e adaptação algorítmica contínua. Em experimento conduzido por Lee e Nass (2003), observou-se que características paralinguísticas da voz sintetizada bastam para ativar categorias sociais complexas, afetando tanto a percepção de presença social quanto a disposição à conformidade. Mesmo quando a artificialidade da voz é plenamente reconhecida, os usuários respondem mobilizando expectativas sociais sedimentadas, como associações entre gênero, competência e segurança.

Nesse âmbito, vozes associadas à feminilidade tendem a suscitar maior sensação de acolhimento, proximidade e segurança, enquanto vozes associadas à masculinidade costumam ser percebidas como mais competentes e tecnicamente confiáveis. Esses efeitos decorrem da ativação automática de categorias culturais profundamente sedimentadas, e não de crenças conscientes acerca da natureza do agente artificial. Conforme demonstram Lee e Nass, ainda que diante do reconhecimento explícito da artificialidade do sistema, os participantes mobilizam esquemas sociais humanos na avaliação da interação: “apesar de saberem que os computadores não possuem personalidade em qualquer sentido humano e apesar de serem confrontados com uma voz claramente não humana, os participantes aplicaram categorias sociais e heurísticas às vozes sintetizadas por computador ao avaliar o grau de presença social em suas atividades de uso de mídia e computador” (Lee; Nass, 2003, p. 293, tradução nossa).

A apropriação sistemática dessas regularidades cognitivas no *design* de assistentes virtuais revela a dimensão normativa implicada no funcionamento desses sistemas. Ao mobilizarem associações culturalmente construídas, tais dispositivos favorecem engajamento, persuasão e adesão do usuário. A personalização de vozes, estilos comunicativos e traços de personalidade deixa de constituir mero aprimoramento funcional e passa a operar como tradução técnica de normas sociais previamente sedimentadas.

Diante desse conjunto de evidências, as assistentes virtuais ultrapassam o estatuto de simples ferramentas técnicas. Elas se afirmam como componentes centrais de uma economia orientada por dados e, simultaneamente, como dispositivos culturais que reproduzem e estabilizam normas sociais, sobretudo aquelas vinculadas ao gênero. A feminilidade programada excede o atributo meramente estético ou funcional e passa a ser compreendida como uma forma social produzida por corporações, arquiteturas algorítmicas e práticas de consumo, cujos efeitos ultrapassam a interface e alcançam dimensões éticas, políticas e simbólicas da experiência social.

Esse cenário exige uma análise crítica que ultrapasse a eficácia comunicativa das tecnologias e interrogue os regimes de sentido e de poder que elas atualizam. Compreender as respostas sociais dirigidas às assistentes virtuais implica reconhecer que a automaticidade cognitiva que sustenta a interação entre usuário e agente artificial constitui, simultaneamente, o terreno no qual se estruturam hierarquias, expectativas e formas historicamente situadas de organização do social. Nesse nível pré-reflexivo da prática cotidiana, a técnica excede sua função instrumental e opera como instância normativa, capaz de moldar disposições, legitimar assimetrias e orientar modos de relação sob a aparência de funcionalidade e neutralidade.

2.4. *Ethics by Design* e a inscrição moral nos sistemas técnicos

A inscrição moral da técnica ocorre quando expectativas sociais se estabilizam como parâmetros de funcionamento.

A discussão sobre gênero nas assistentes virtuais mostra que as tecnologias de voz operam como tipologias normativas

complexas, nas quais valores, expectativas sociais e hierarquias culturalmente produzidas se inscrevem de modo persistente. Ao funcionar tecnicamente, esses sistemas também sedimentam moralidades historicamente constituídas e as convertem em padrões de interação, resposta e reconhecimento. Nesse contexto, a problemática ética desloca-se da avaliação de comportamentos isolados da tecnologia para a investigação dos regimes de possibilidade que ela institui: aquilo que passa a ser percebido como aceitável, esperado ou legítimo nas relações mediadas pela técnica.

Ao deslocar a reflexão ética para o nível da concepção técnica, torna-se possível mobilizar a noção de *Ethics by Design*¹¹ como operador conceitual para examinar a relação entre ética e tecnologia no contexto das assistentes virtuais. Pensar a ética desde a concepção implica reconhecer que os sistemas técnicos, ao serem projetados, materializam decisões morais antecipadas, inscritas em suas arquiteturas, interfaces e lógicas operacionais. Autores como Luciano Floridi, Paula Boddington, Batya Friedman, James Moor e Brent Mittelstadt sustentam que valores éticos se expressam na própria configuração dos sistemas, orientando formas de ação, modos de interação e limites de aceitabilidade, antes mesmo de qualquer uso concreto. Portanto, a moralidade figura como componente constitutivo da tecnologia, integrando suas condições estruturais de possibilidade.

Luciano Floridi inclui-se entre os principais responsáveis por deslocar o debate ético sobre a tecnologia do plano da regulação posterior para o nível ontológico do *design*. No âmbito de sua filosofia da informação, os sistemas digitais deixam de ser concebidos como instrumentos neutros a serviço de fins humanos e passam a ser compreendidos como ambientes informacionais que estruturam modos de ação, de relação e de valoração. Ao delimitar aquilo que pode ser feito, dito e reconhecido como

¹¹ A expressão *Ethics by Design* refere-se à abordagem segundo a qual considerações éticas devem ser integradas desde as fases iniciais de concepção, desenvolvimento e implementação de sistemas tecnológicos, e não apenas avaliadas a posteriori. Essa perspectiva é desenvolvida, entre outros, por Luciano Floridi (2013), Batya Friedman (1996; 2006), James Moor (1985) e, mais recentemente, por Brent Mittelstadt *et al.* (2016), no âmbito dos debates sobre ética da informação, *Value Sensitive Design* e governança algorítmica.

aceitável, a tecnologia intervém diretamente na constituição do mundo social e estabelece as condições normativas nas quais práticas, decisões e formas de vida se tornam possíveis.

É nesse campo que Floridi introduz a noção de *Ethics by Design*, ao sustentar que decisões morais não devem ser compreendidas como acréscimos externos ao funcionamento dos sistemas técnicos. A dimensão ética inscreve-se desde a concepção dos artefatos e manifesta-se em escolhas frequentemente apresentadas como estritamente técnicas, como a arquitetura do sistema, a definição de padrões operacionais, as opções de interface e os modos de interação. A ética da informação, nesse sentido, desloca a atenção das intenções individuais para a configuração dos ambientes informacionais, entendidos como instâncias que moldam comportamentos, expectativas e campos de possibilidade da ação coletiva (Floridi, 2013).

Ao situar a ética no exame das condições técnicas que antecedem qualquer juízo sobre usos corretos ou inadequados da tecnologia, essa perspectiva reconfigura profundamente o estatuto da moralidade. O problema ético emerge antes mesmo da interação efetiva, na maneira como o sistema organiza alternativas, orienta condutas e torna outras possibilidades menos acessíveis. A moralidade apresenta-se, assim, como efeito técnico e normativo do próprio funcionamento dos sistemas, presente na arquitetura que estrutura a experiência mediada pela tecnologia.

No caso das assistentes virtuais de voz, essa abordagem permite situar escolhas como a feminização da voz, o tom afetivo da interação e a tolerância a comandos agressivos no plano das decisões normativas antecipadas que estruturam a experiência do usuário. Tais decisões expressam opções normativas integradas ao ambiente informacional produzido por esses sistemas e organizam relações de poder, expectativas de comportamento e formas de reconhecimento. À luz da filosofia de Floridi, a femilidade programada aparece como componente constitutivo da ecologia moral dessas tecnologias desde a sua origem.

A formulação floridiana articula-se, ainda, à noção de *infosfera*, entendida como o ambiente informacional constituído pelo conjunto de todas as entidades informacionais – humanas e não

humanas –, bem como por suas propriedades, interações e relações mútuas (Floridi, 2013). Trata-se de um novo “ambiente do ser”, análogo ao mundo natural, porém produzido informacionalmente. Nesse aspecto, agentes artificiais, como as assistentes virtuais, integram plenamente a ecologia moral da vida social e assumem participação ativa na produção de efeitos éticos de caráter estrutural.

Na sua obra *The Fourth Revolution*, Floridi aprofunda essa perspectiva ao sustentar que a ética da informação requer a superação de uma lógica estritamente antropocêntrica, propondo uma abordagem macroética sensível aos impactos sistêmicos das tecnologias digitais sobre o mundo social. Essa orientação é desenvolvida de modo mais sistemático em *The Ethics of Information*, obra na qual a avaliação ética incide sobre ações, políticas e sistemas segundo sua contribuição para o florescimento ou para o empobrecimento da *infosfera* (Floridi, 2013). Com esse critério, práticas tecnológicas que reforçam estereótipos de gênero, naturalizam regimes de vigilância ou modelam subjetividades hierárquicas constituem formas de degradação do ambiente informacional compartilhado, ainda que não gerem danos diretos ou intencionais a indivíduos específicos.

Em relação às assistentes virtuais de voz, essa abordagem oportuniza a compreensão da feminilidade programada como um problema ético-informacional de caráter estrutural. Ao reiterar padrões de docilidade, disponibilidade afetiva e obediência, esses sistemas introduzem assimetrias normativas duráveis na *infosfera*. A moralidade manifesta-se, nesse caso, na arquitetura relacional e semântica que se estabiliza no cotidiano das interações mediadas pela técnica, mais do que em decisões pontuais do sistema.

A contribuição de Floridi oferece, desse modo, um fundamento ontológico e normativo para sustentar que sistemas técnicos requerem avaliação ética desde a sua concepção. Sua proposta, contudo, preserva deliberadamente um elevado grau de generalidade, orientado à preservação e ao florescimento da *infosfera* como ambiente moral compartilhado. Para que essa ética de caráter ambiental se traduza em práticas concretas de

desenvolvimento tecnológico, impõe-se a necessidade de um marco normativo capaz de converter valores morais em critérios operacionais de projeto.

Nesse limiar, Batya Friedman, Peter H. Kahn Jr. e Alan Borning formulam o *Value Sensitive Design* (VSD) como uma abordagem teoricamente fundamentada do *design* tecnológico, voltada à incorporação sistemática de valores humanos ao longo de todo o processo de concepção e implementação (Friedman; Kahn; Borning, 2008, p. 70). O VSD parte do reconhecimento de que artefatos técnicos operam como mediações normativas: organizam possibilidades de ação, distribuem benefícios e custos e interferem nas expectativas sociais de conduta. Em lugar de tratar valores como acréscimos externos, o paradigma assume a tarefa de torná-los explicitáveis, discutíveis e tecnicamente traduzíveis ao longo do ciclo de vida dos sistemas.

No interior dessa proposta, “valor” designa aquilo que uma pessoa ou um grupo considera importante na vida, em um sentido amplo que excede o caráter econômico. Essa definição evita reduzir a ética a preferências individuais imediatas e permite tratar, no âmbito do projeto, dimensões de alcance moral e social: dignidade, autonomia, justiça, privacidade, confiança, entre outras. Ao adotar esse referencial, o VSD oferece uma gramática conceitual capaz de vincular decisões técnicas concretas a efeitos normativos duráveis, aspecto decisivo na análise de sistemas conversacionais e assistentes virtuais de voz, cujas escolhas de interface e interação passam a operar como padrões cotidianos de relação.

O VSD estrutura-se por uma metodologia tripartite que integra investigações conceituais, empíricas e técnicas de modo iterativo (Friedman; Kahn; Borning, 2008, p. 70). As investigações conceituais identificam *stakeholders* diretos e indiretos, valores implicados e tensões entre valores; as investigações empíricas examinam como esses valores se manifestam em contextos concretos de uso; as investigações técnicas analisam como propriedades do sistema favorecem determinados valores e dificultam a realização de outros. No caso das assistentes virtuais, essa tríade analítica permite rastrear com precisão como decisões

relativas à voz, ao tom, à *persona* e às respostas a agressões distribuem expectativas de autoridade, deferência e disponibilidade afetiva, inscrevendo normatividades no próprio funcionamento da interface.

Ao distinguir *stakeholders*¹² diretos – aqueles que interagem com o sistema – e *stakeholders* indiretos – aqueles que sofrem seus efeitos sem interação direta –, o VSD amplia o campo de responsabilidade do *design* e impede que o impacto moral seja avaliado apenas pela satisfação imediata do usuário (Friedman; Kahn; Borning, 2008, p. 70). No caso das assistentes de voz, esse deslocamento conduz a análise para além da “eficiência” da interação, permitindo examinar a circulação social de normas que se estabilizam pela repetição: estilos de comando, expectativas de obediência e papéis de gênero reiterados por uma *persona* feminilizada. Dessa forma, o VSD atua como ponte entre a crítica filosófica do poder e um diagnóstico técnico-operacional do projeto, oferecendo critérios para identificar onde, como e quais efeitos, valores e hierarquias se incorporam aos sistemas.

Ao deslocar o foco do comportamento isolado dos usuários para a configuração estrutural dos ecossistemas informacionais, essa perspectiva estabelece um diálogo direto com a ética da informação formulada por Luciano Floridi. Na abordagem do autor, danos informacionais comprometem a integridade da infosfera; na formulação de Friedman, Kahn e Borning, esses danos tornam-se rastreáveis a decisões específicas de *design* que privilegiam determinados interesses, normatividades e formas de relação. No caso das assistentes virtuais, a feminilidade programada pode ser compreendida como resultado de escolhas valorativas incorporadas ao projeto técnico, cujos efeitos se prolongam sobre processos de socialização, reconhecimento e exercício do poder.

Desse modo, a passagem de Floridi a Friedman e colaboradores assinala a transição de uma ética ambiental da informação

¹² *Stakeholders* é um termo empregado para designar os indivíduos, grupos ou instituições que possuem interesse direto ou indireto em determinado sistema, processo ou tecnologia, seja por participarem de sua concepção, uso e governança, seja por serem afetados por seus efeitos e consequências (Friedman; Kahn; Borning, 2008).

para uma ética orientada ao projeto, na qual a responsabilidade moral se inscreve diretamente nas práticas de concepção, desenvolvimento e implementação dos sistemas técnicos. Essa articulação possibilita avançar da crítica ontológica para a intervenção normativa, criando as condições para uma análise mais rigorosa das responsabilidades implicadas no *design* das assistentes virtuais na contemporaneidade.

É nesse ponto que a teoria de Floridi encontra continuidade e tradução operacional na proposta do *Value Sensitive Design*, desenvolvida por Batya Friedman e colaboradores. Enquanto Floridi oferece um enquadramento ontológico e macroético da infosfera, Friedman propõe uma metodologia concreta para integrar valores humanos ao próprio processo de concepção tecnológica. Segundo os autores, os sistemas técnicos “incorporam valores em seu *design*, de modo intencional ou não intencional” (Friedman; Kahn; Borning, 2008, p. 72, tradução nossa). Os valores encontram-se encastrados, desde a origem, nas funcionalidades, nas interfaces e nas formas de interação dos sistemas, orientando seus modos de uso.

O *Value Sensitive Design* parte do pressuposto de que decisões de *design* constituem, inevitavelmente, decisões morais. Como afirmam Friedman e seus coautores, “o *design* tecnológico não é neutro em termos de valores; ao contrário, pode sustentar ou comprometer determinados valores humanos” (Friedman; Kahn; Borning, 2008, p. 75, tradução nossa). No caso das assistentes virtuais, a recorrência de vozes femininas diferentes, respostas afetivamente acomodadas e a ausência de resistência diante de interações abusivas evidenciam uma orientação normativa específica, ainda que frequentemente legitimada por critérios de usabilidade, eficiência ou conveniência comunicativa.

A articulação entre Floridi, Friedman e demais pesquisadores aponta as assistentes virtuais como agentes informacionais moralmente relevantes, cujos efeitos extrapolam o plano funcional e incidem sobre a organização normativa da vida cotidiana. Se, para Floridi, a ética da informação exige proteger a infosfera contra processos de empobrecimento normativo; para Friedman *et al.*, essa proteção requer a incorporação deliberada de valores

como justiça, autonomia e não discriminação no próprio *design* tecnológico. A *Ethics by Design* assume, desse modo, o estatuto de exigência ética, na medida em que tecnologias de voz intervêm diretamente na constituição da experiência social.

A partir desse referencial, a questão ética desloca-se progressivamente do problema da inscrição de valores no *design* para o da responsabilidade moral em contextos sociotécnicos complexos. Se a ética da informação de Floridi fornece o domínio ontológico, no qual tecnologias passam a ser avaliadas como ambientes morais, e se o *Value Sensitive Design* de Friedman e colaboradores oferece instrumentos para operacionalizar, vale pensar a responsabilidade moral para além de agentes individuais, levando em conta cadeias decisórias, opacidades técnicas e impactos coletivos, movimento no qual se insere, de modo significativo, a contribuição de Paula Boddington.

Boddington, em *Towards a Code of Ethics for Artificial Intelligence*, sustenta que a inteligência artificial desafia os modelos clássicos de atribuição de responsabilidade moral, uma vez que seus efeitos resultam de processos técnicos distribuídos, opacos e socialmente difusos. Nesse quadro, decisões incorporadas ao *design* dos sistemas produzem impactos que excedem a esfera dos desenvolvedores individuais e se projetam sobre coletividades amplas ao longo do tempo. Por essa razão, a autora defende que valores éticos sejam considerados desde as fases iniciais de concepção e desenvolvimento, pois intervenções corretivas posteriores tendem a apresentar alcance restrito. No caso das assistentes virtuais, escolhas relativas à voz, à *persona*, à linguagem e aos padrões de resposta adquirem estatuto moral justamente porque incorporam valores de forma implícita e moldam percepções sociais acerca de gênero, autoridade e práticas de cuidado.

A exigência ética que se impõe decorre da própria configuração dos sistemas de inteligência artificial, caracterizados por cadeias de produção fragmentadas e por efeitos que se deslocam significativamente de seus pontos de origem. De acordo com Boddington, a IA envolve “sistemas altamente complexos, nos quais pessoas que trabalham em diferentes elementos do processo frequentemente não se conhecem entre si, e cujos efeitos podem

estar muito distantes de seus produtores, como ocorre quando processos de aprendizagem de máquina geram algoritmos capazes de afetar milhões de pessoas longe de seu ponto de origem” (Boddington, 2017, p. 29, tradução nossa). Nesse cenário, a responsabilidade moral ultrapassa o âmbito das decisões pontuais e requer uma abordagem estrutural, atenta às condições técnicas, organizacionais e normativas que orientam o funcionamento desses sistemas.

A natureza sistêmica e difusa da inteligência artificial, por sua vez, tensiona os modelos tradicionais de imputação moral centrados no agente individual. Para Boddington, decisões técnicas, mesmo quando apresentadas como neutras, produzem efeitos éticos relevantes nos planos social, político e cultural. Diante disso, torna-se insuficiente conceber a ética apenas como um mecanismo de correção posterior de danos. A autora observa que, em sistemas de IA nos quais decisões e ações geram efeitos de grande alcance, frequentemente opacos ou difíceis de rastrear, “a ideia de construir mecanismos que tornem as decisões da máquina ‘éticas’ pode parecer uma possibilidade tentadora” (Boddington, 2017, p. 89).

A autora, contudo, chama atenção para os limites dessa abordagem, sobretudo quando valores éticos permanecem indefinidos nas etapas iniciais do processo de concepção. Esse ponto explicita-se mais quando Boddington afirma que “será muito difícil programar uma máquina para lidar com questões éticas se não tivermos uma ideia suficientemente clara de quais são nossos objetivos de valor” (Boddington, 2017, p. 90). Essa formulação sustenta a compreensão de que a ética da inteligência artificial exige incorporação originária às decisões de *design*, arquitetura e funcionamento dos sistemas, com caráter estruturante, em lugar de corretivo.

Essa perspectiva ganha especial relevância no exame das assistentes virtuais. Embora frequentemente apresentadas como produtos orientados por critérios estéticos, funcionais ou mercadológicos, escolhas relativas à voz, à *persona*, à linguagem e às modalidades de interação incorporam valores de forma implícita e orientam percepções sociais e expectativas normativas. Bod-

dington chama a atenção para esse aspecto ao afirmar que “deve haver atenção explícita à forma como os valores são incorporados nas práticas e à maneira como podem estar presentes de modo subliminar na linguagem e no enquadramento em torno de códigos e regulações” (Boddington, 2017, p. 111).

Dessa forma, decisões frequentemente apresentadas como meramente técnicas ou interacionais assumem estatuto moral, pois intervêm diretamente na produção de sentidos, na naturalização de papéis sociais e na organização das relações entre humanos e agentes artificiais. A ética da inteligência artificial, conforme delineada por Boddington, exige, assim, atenção rigorosa às dimensões culturais, linguísticas e valorativas inscritas na arquitetura e no funcionamento dos sistemas, principalmente, quando esses dispositivos se integram ao cotidiano e passam a moldar, de maneira contínua, os modos correntes de interação social.

É precisamente nesse ponto que a reflexão de James Moor mostra-se decisiva. Sua análise incide sobre um problema estrutural das tecnologias computacionais: a defasagem recorrente entre a velocidade da inovação técnica e a elaboração de referenciais normativos capazes de orientar seus efeitos sociais. Se Paula Boddington enfatiza a difusão da responsabilidade moral em sistemas complexos de inteligência artificial, Moor aprofunda essa problemática ao mostrar como a própria inovação computacional pode engendrar lacunas normativas, nas quais decisões técnicas operam na ausência de critérios éticos previamente estabilizados. Ao examinar os impactos morais específicos das tecnologias digitais, o autor oferece instrumentos conceituais para pensar situações em que o desenvolvimento técnico antecede, tensiona ou excede os quadros normativos disponíveis, cenário característico das tecnologias de inteligência artificial e, de modo particular, das assistentes virtuais.

Na perspectiva de Moor, as tecnologias digitais devem ser compreendidas como agentes de transformação do contexto moral. Em *Why We Need Better Ethics for Emerging Technologies*, o autor sustenta que tecnologias emergentes e revolucionárias introduzem novos instrumentos e, ao mesmo tempo, reconfiguram os

próprios ambientes nos quais as decisões morais se tornam possíveis e inteligíveis. Segundo Moor, “tecnologias revolucionárias produzem um impacto transformador em larga escala sobre a maneira como uma sociedade funciona” (Moor, 2006, p. 112, tradução nossa). Tal impacto incide de forma estrutural sobre práticas sociais, expectativas normativas e modos de organização coletiva, extrapolando o âmbito de aplicações pontuais.

Segundo Moor, o caráter ético dessas transformações manifesta-se no fato de que tecnologias emergentes produzem situações inéditas para as quais ainda não existem políticas morais estabilizadas, cenário que o autor denomina *policy vacuums*¹³. Em suas palavras, “tecnologias novas e poderosas geram oportunidades inéditas de ação para as quais políticas éticas adequadas ainda não foram desenvolvidas” (Moor, 2006, p. 115, tradução nossa). Nessa perspectiva, a ética computacional assume uma função prospectiva, orientada à antecipação dos efeitos normativos das tecnologias sobre os contextos sociais em que operam, em vez de se restringir a respostas reativas diante de danos já produzidos.

De forma complementar à noção de *policy vacuum*, Moor formula o conceito de *moral lag*¹⁴ para designar o descompasso recorrente entre a velocidade da inovação tecnológica e a capacidade das normas morais, jurídicas e sociais de se ajustarem a seus efeitos. Esse hiato entre inovação e normatividade – o descompasso moral (*moral lag*) – manifesta-se concretamente por meio de vácuos normativos (*policy vacuums*).

Enquanto novas tecnologias expandem rapidamente as possibilidades de ação, os referenciais éticos disponíveis permanecem ancorados em contextos anteriores, produzindo um intervalo crítico, no qual práticas socialmente impactantes se consolidam sem avaliação normativa adequada. O atraso moral decorre da própria dinâmica da transformação técnica, que reconfigura em

¹³ O termo *policy vacuums*, empregado por James Moor, designa situações nas quais novas tecnologias introduzem possibilidades inéditas de ação para as quais ainda não existem normas éticas, políticas ou regulatórias claramente formuladas. Esses vácuos normativos decorrem do descompasso entre o ritmo acelerado da inovação técnica e a capacidade das sociedades de elaborar enquadramentos morais e institucionais adequados para orientar seu uso (Moor, 2006).

¹⁴ Para Moor, *moral lag* designa o descompasso temporal entre a rapidez da inovação tecnológica e a capacidade das sociedades de formular normas, valores e políticas morais adequadas para lidar com seus efeitos.

ritmo acelerado os quadros nos quais se organizam decisões e formas de imputação ética. No funcionamento das assistentes virtuais, o descompasso moral (*moral lag*) torna-se especialmente evidente, pois escolhas de *design*, como a feminização da voz, a tolerância a interações abusivas e a naturalização da escuta permanente, disseminam-se antes que seus efeitos sobre relações de gênero, padrões de autoridade e regimes de vigilância sejam plenamente reconhecidos e criticamente elaborados.

Em relação às assistentes virtuais de voz, a noção de transformação de contexto mostra-se especialmente esclarecedora. Sistemas como Alexa, Siri e Google Assistant deslocam o espaço doméstico, tradicionalmente associado a uma esfera de privacidade relativa, para um cenário atravessado pela possibilidade de escuta contínua, pelo registro sistemático de interações e pela resposta automatizada. Antes mesmo de qualquer iniciativa deliberada do usuário, a determinação moral da casa já se encontra alterada, pois o silêncio deixa de significar ausência de observação e passa a coexistir com a latência permanente da captação de dados. Trata-se, nos termos de Moor, de uma reestruturação do ambiente de decisão moral, na qual práticas corriqueiras adquirem novos sentidos à luz de uma infraestrutura técnica difusa e amplamente opaca.

Essa transformação do contexto doméstico intensifica os desafios éticos analisados por Moor à medida que a tecnologia atinge o que o autor denomina *power stage*, isto é, o estágio em que se encontra amplamente integrada às práticas ordinárias da vida social. Segundo Moor, “quanto maior o impacto social de uma tecnologia, maior o número e a complexidade dos problemas éticos que ela engendra” (Moor, 2006, p. 116, tradução nossa). As assistentes virtuais, ao se disseminarem nos ambientes domésticos, exemplificam esse nível de maturidade sociotécnica: mais do que ampliar funcionalidades, produzem uma reorganização discreta das normas de interação, das expectativas de obediência e das relações de poder mediadas por formas de linguagem artificialmente produzidas.

É precisamente para tornar inteligível esse novo cenário ético, caracterizado pela integração profunda de tecnologias al-

gorítmicas às práticas ordinárias da vida social, que se impõe a necessidade de uma sistematização mais rigorosa dos tipos de danos produzidos por esses mecanismos. Nesse ponto, a análise desenvolvida por Brent Mittelstadt e colaboradores (2016) oferece instrumentos conceituais para discriminar, com maior precisão, os problemas éticos engendrados por sistemas algorítmicos multidimensionais. Em uma das revisões mais influentes sobre ética dos algoritmos, os autores demonstram que tais problemas excedem falhas técnicas pontuais e assumem caráter estrutural, manifestando-se como questões de ordem epistemológica, social e política. Entre elas, destacam-se a opacidade, o viés, a discriminação, a exclusão e as fragilidades nos mecanismos de governança ética.

A opacidade algorítmica constitui um obstáculo estrutural à inteligibilidade e à contestação das decisões automatizadas, uma vez que “raramente se dispõe de qualquer senso concreto de como ou por que uma determinada classificação foi alcançada a partir das entradas” (Mittelstadt *et al.*, 2016, p. 108, tradução nossa). Quando articulada à feminilidade programada das assistentes virtuais, essa opacidade favorece a estabilização silenciosa de normas de gênero, ao ocultar os critérios valorativos e as escolhas normativas que orientam o funcionamento desses sistemas.

Quando se desloca a análise dos sistemas algorítmicos abstratos para as interações cotidianas mediadas por voz, os danos éticos identificados por Mittelstadt e colaboradores adquirem uma visibilidade particularmente concreta. Nas assistentes virtuais, tais danos revelam-se de modo agudo nas dinâmicas de viés e discriminação, presentes no próprio funcionamento dos sistemas. Os autores demonstram que algoritmos tendem a reproduzir e amplificar desigualdades socialmente consolidadas, mesmo quando esses padrões não figuram explicitamente nas intenções de projeto. Como observam, “não se pode assumir que algoritmos discriminarão apenas segundo características observáveis ou compreensíveis aos humanos” (Mittelstadt *et al.*, 2016, p. 111, tradução nossa). A feminilização recorrente das assistentes virtuais exemplifica esse mecanismo: ao reiterar convenções de gênero sob a aparência de neutralidade funcional, esses siste-

mas estabilizam normas culturais historicamente situadas e as apresentam como escolhas naturais, inevitáveis ou tecnicamente justificadas.

Além disso, os sistemas algorítmicos participam ativamente da produção externa de identidades, deslocando dos sujeitos o controle sobre os modos pelos quais são representados, categorizados e avaliados. A lógica classificatória que os orienta privilegia o enquadramento de indivíduos em perfis abstratos, favorecendo processos de desindividualização fundados em padrões estatísticos e inferências probabilísticas. Trata-se, como observam Mittelstadt e colaboradores, de “uma tendência a julgar e tratar pessoas com base em características de grupo, em vez de suas características individuais e mérito próprio” (Mittelstadt *et al.*, 2016, p. 110, tradução nossa). No plano relacional e cultural, as assistentes virtuais operam de modo análogo ao estabilizar uma identidade feminina padronizada, na qual traços como disponibilidade, docilidade e submissão reaparecem como atributos generalizáveis do feminino, naturalizados pela repetição técnica e pela familiaridade da interação.

Esse processo pode ser interpretado como uma forma de violência epistêmica, na medida em que determinados modos de ser, falar e existir são naturalizados, enquanto outros são deslegitimados, silenciados ou tornados impraticáveis no campo do pensável. Ao intervir na configuração do conhecimento social, sistemas algorítmicos contribuem para definir quais categorias se tornam relevantes, legítimas ou visíveis, atuando como instâncias de ordenação normativa do sentido. Nesse contexto, ao estabilizar uma feminilidade homogênea e submissa como interface padrão, as assistentes virtuais estreitam o imaginário do feminino e convertem normas históricas em padrões técnicos apresentados sob a aparência de objetividade e neutralidade.

Outro eixo crítico identificado por Mittelstadt e colaboradores refere-se aos *déficits* de governança e de atribuição de responsabilidade em sistemas algorítmicos. A automação, segundo essa análise, desloca a responsabilidade moral dos agentes humanos para a própria tecnologia e produz um cenário no qual decisões técnicas passam a ser percebidas como neutras, inevitáveis ou

autojustificadas. Nesse quadro, instala-se aquilo que os autores descrevem como uma tendência a “esconder-se atrás do computador” e a presumir que processos automatizados são corretos por definição (Mittelstadt *et al.*, 2016, p. 111, tradução nossa). No caso das assistentes virtuais, esse deslocamento contribui para a legitimação de escolhas de *design* profundamente normativas, como a feminilização da voz, sem que seus efeitos sociais, políticos e culturais sejam devidamente problematizados.

Essa análise evidencia que danos éticos emergem quando sistemas técnicos passam a organizar identidades, expectativas e normas sociais em contextos atravessados por encobrimento normativo, baixa possibilidade de contestação e fragilidades nos regimes de atribuição de responsabilidade. As assistentes virtuais, ao operarem como interfaces afetivas feminilizadas, exemplificam a convergência entre viés algorítmico, violência epistêmica e déficits de governança em um mesmo padrão sociotécnico. Esse quadro reforça a necessidade de uma abordagem de *Ethics by Design* que reconheça o gênero como dimensão ética constitutiva da arquitetura dos sistemas, inscrita nas decisões de projeto e nas formas de interação, em vez de tratá-lo como atributo periférico ou meramente formal.

Em conjunto, essas abordagens permitem compreender a *Ethics by Design* como uma resposta teórica e prática ao reconhecimento de que os sistemas técnicos atuais funcionam como verdadeiras infraestruturas morais. Ao configurarem contextos de ação, orientarem condutas e estabilizarem normas, as tecnologias digitais – de modo particularmente evidente no caso das assistentes virtuais de voz – participam ativamente da produção de subjetividades, expectativas sociais e relações de poder. Nesse modelo, a ética desloca-se da condição de avaliação posterior ao uso da tecnologia e assume o estatuto de exigência interna aos próprios processos de concepção, desenvolvimento e governança das infraestruturas digitais.

À luz desse conjunto de diagnósticos, a proposta de *Ethics by Design* pode ser compreendida como uma resposta normativa às transformações éticas produzidas pelos sistemas digitais. Essa abordagem articula princípios como a desgenerificação da

tecnologia, a transparência e a justiça algorítmica, a autonomia dos usuários, a responsabilidade institucional e a participação social nos processos de concepção e governança. Tais princípios indicam que a ética aplicada à inteligência artificial ultrapassa o domínio da reflexão teórica e se afirma como prática projetual, responsável por orientar todo o ciclo de vida dos sistemas, desde as decisões iniciais de *design* até a implementação, o uso cotidiano e os marcos regulatórios que os estruturam.

A feminilidade programada, examinada ao longo deste capítulo, evidencia que escolhas de *design* apresentadas como neutras incorporam valores históricos e assimetrias estruturais, com efeitos éticos persistentes no ambiente informacional compartilhado. A conclusão desse diagnóstico aponta para a necessidade de compreender que a responsabilidade moral relacionada à inteligência artificial ultrapassa o plano das ações executadas pelos sistemas e incide sobre as próprias condições de sua existência técnica e relacional. O problema ético localiza-se tanto nas operações realizadas quanto na maneira como a tecnologia é concebida para falar, responder, acolher e ocupar lugares sociais específicos. Nesse deslocamento, do uso ao *design*, da função à normatividade inscrita, a *Ethics by Design* manifesta-se como tema indispensável para uma inteligência artificial comprometida com a justiça, a autonomia e a abertura à pluralidade das formas de vida social.

3. Perspectivas éticas para o futuro do *design* das interfaces de voz

A promessa do Ethics by Design falha quando supõe que a moral pode ser acrescentada à técnica sem interrogar as normas que ela já executa.

A ética da inteligência artificial de voz consolidou-se, nos últimos anos, como um campo estruturado por *frameworks* de governança, princípios regulatórios e iniciativas de *design* responsável. Transparência, justiça, explicabilidade e centralidade humana tornaram-se eixos recorrentes em documentos institucionais e políticas corporativas, o que sinaliza um reconhecimento crescente de que sistemas algorítmicos produzem efeitos morais e sociais relevantes. Esses instrumentos desempenham papel importante na contenção de danos e na padronização de boas práticas, especialmente em contextos de ampla disseminação tecnológica.

Esse conjunto de abordagens, contudo, opera majoritariamente segundo uma lógica de compatibilização entre inovação técnica e valores éticos previamente definidos. A ética tende a figurar como uma camada integrável ao ciclo de desenvolvimento, voltada a ajustes, correções e mecanismos de controle posteriores. Tal perspectiva revela seus limites quando a inteligência artificial ultrapassa o estatuto de ferramenta instrumental e ocupa domínios nos quais estão em jogo práticas relacionais, expectativas de reconhecimento, formas de escuta e regimes de cuidado. Nesses contextos, a mediação técnica intervém diretamente nas condições éticas que sustentam a relação entre sujeitos, reconfigurando assimetrias, responsabilidades e possibilidades de resposta.

Este capítulo propõe deslocar o debate ético do plano da adequação técnica para o da decisão normativa. Ao articular responsabilidade moral, cuidado e dignidade da pessoa humana, a

análise examina os limites éticos da automação de práticas relacionais e formula critérios capazes de orientar escolhas concretas relativas ao *design*, à implementação e ao uso de interfaces de voz. Pensar o futuro ético da inteligência artificial exige, assim, mais do que projetar sistemas tecnicamente aprimorados: exige decidir onde a técnica pode intervir, sob quais condições essa intervenção se justifica e em que pontos ela deve cessar para preservar a inteligibilidade ética das relações humanas.

3.1 Frameworks de responsabilidade moral para desenvolvedores e corporações

A justiça tecnológica começa quando a reprodução automática de estereótipos passa a ser reconhecida como violação ética.

A complexidade ética das assistentes virtuais tem sido abordada, de modo recorrente, a partir de análises centradas no comportamento dos usuários e nos efeitos imediatos da interação tecnomediada entre sujeitos humanos e agentes artificiais. Uma parcela significativa da literatura atual em estudos críticos da tecnologia, contudo, aponta a necessidade de ampliar esse foco, incorporando a análise das estruturas sociotécnicas responsáveis pela concepção, pelo desenvolvimento, pela implementação e pela manutenção desses sistemas. Autoras e autores como Langdon Winner, Safiya Umoja Noble, Ruha Benjamin e Cathy O’Neil examinam tecnologias digitais como arranjos sociotécnicos dotados de implicações políticas, morais e distributivas, cujos efeitos excedem o plano estritamente funcional. Nessa delimitação, as assistentes virtuais passam a ser compreendidas no interior de processos organizacionais, modelos de negócio e regimes de governança que condicionam suas formas de operação e sua inserção social.

Winner, em seu ensaio clássico “Do Artifacts Have Politics?”, oferece um dos referenciais teóricos mais influentes para compreender a dimensão política dos artefatos técnicos no debate sobre as assistentes virtuais. Para o autor, tecnologias excedem critérios de eficiência ou funcionalidade, pois incorporam formas específicas de poder, autoridade e organização

social. Sistemas técnicos podem estabilizar quadros políticos de maneira durável ao materializar escolhas que orientam práticas, distribuem possibilidades de ação e configuram relações de comando e dependência. Nas palavras de Winner, os artefatos técnicos podem ser avaliados também pelas “formas específicas de poder e autoridade que podem encarnar” (Winner, 1980, p. 121). Ao reconhecer que decisões técnicas participam ativamente da constituição de ordens sociais persistentes, essa formulação desloca a análise ética da tecnologia para o plano do *design* e da organização material.

A partir dessa perspectiva, a tecnologia pode ser compreendida como uma forma material de ordenação do mundo social. Uma vez incorporados ao cotidiano, os artefatos técnicos organizam práticas, distribuem papéis, delimitam expectativas e estabilizam padrões de conduta que tendem a se naturalizar pelo uso reiterado. Winner avalia que as tecnologias operam como modelos normativos silenciosos, na medida em que estruturam relações antes mesmo de qualquer deliberação explícita, definindo quem age, quem responde, quem comanda e quem se adapta. Nesse sentido, a política dos artefatos manifesta-se sobretudo na capacidade de configurar ambientes de ação, nos quais determinadas formas de interação se tornam evidentes e legitimadas, enquanto outras permanecem periféricas, desestimuladas ou fora da esfera inteligível.

Aplicado às assistentes virtuais de voz, esse referencial permite compreender que atributos como a escolha reiterada de vozes femininas, o uso de registros linguísticos associados à submissão e a expectativa de disponibilidade permanente excedem o plano de decisões estéticas ou funcionais. Esses elementos expressam aquilo que Winner denomina “formas de ordenar o mundo”, isto é, modos pelos quais as tecnologias alicerçam práticas sociais, configuram expectativas de comportamento e organizam relações de autoridade. Segundo o autor, “as coisas que chamamos tecnologias são modos de construir ordem em nosso mundo” (Winner, 1980, p. 128, tradução nossa). Ao se integrarem ao cotidiano doméstico e comunicacional, as assistentes virtuais estabilizam uma gramática relacional na qual o usuário

ocupa, de modo recorrente, a posição de comando, enquanto a interface feminilizada performa obediência, disponibilidade e deferência. Nesse registro, as assistentes virtuais corroboram a tese central de Winner: artefatos fazem política ao operar, uma vez que inscrevem hierarquias sociais em sua própria materialidade técnica.

A noção de tecnologia como forma de ordenar o mundo, formulada por Winner, encontra continuidade na análise dos sistemas algorítmicos desenvolvida por Noble. Se, em Winner, os artefatos técnicos estabilizam relações sociais por meio de sua materialidade e de sua inserção no cotidiano; nos sistemas digitais, essa ordenação adquire um alcance ampliado, automatizado e massivo. A organização social passa a operar menos por dispositivos físicos ou infraestruturas localizadas e mais por algoritmos capazes de classificar, hierarquizar e reproduzir normas em escala global. Nesse deslocamento, escolhas técnicas convertem expectativas culturais em padrões computacionais duráveis, cuja eficácia normativa decorre da repetição automatizada, da naturalização operacional e da integração discreta às rotinas da vida social.

Nas assistentes virtuais, essa dinâmica mostra-se particularmente visível na escolha reiterada de vozes femininas para funções de suporte, atenção relacional e mediação afetiva. À semelhança dos mecanismos de busca analisados por Noble, a feminilidade programada emerge de decisões técnicas e comerciais que convertem expectativas sociais em padrões algorítmicos estabilizados. Como observa a autora, “a tecnologia amplifica desigualdades porque opera em escala, velocidade e alcance muito superiores às interações humanas diretas” (Noble, 2018, p. 10, tradução nossa). Uma vez integradas ao cotidiano doméstico e comunicacional, as assistentes virtuais intensificam esse processo ao reinscrever normas de gênero historicamente situadas em rotinas técnicas automatizadas, reiteradas de modo contínuo e com baixa perceptibilidade social.

Em *Algorithms of Oppression* (2018), Noble analisa sistemas algorítmicos de busca, classificação e recomendação como formas técnicas de organização do mundo social, nas quais hie-

rarquias historicamente constituídas são traduzidas em padrões computacionais e projetadas em nível macroestrutural. Para a autora, algoritmos atuam como infraestruturas normativas capazes de estabilizar valores culturais, interesses econômicos e relações de poder sob a aparência de objetividade técnica. Embora sua investigação empírica se concentre sobretudo na racialização¹⁵ dos sistemas de busca, o alcance teórico de sua análise estende-se de modo direto às questões de gênero. Nas assistentes virtuais, a feminilidade programada insere-se nesse mesmo processo: expectativas sociais associadas à docilidade, à disponibilidade relacional e à subordinação são convertidas em parâmetros técnicos de *design*, linguagem e interação. Uma vez incorporadas a dispositivos de ampla difusão, tais escolhas adquirem persistência normativa, transformam convenções culturais em funcionalidades algorítmicas e deslocam a responsabilidade ética do campo das relações sociais para o domínio apresentado como neutro da tecnologia.

A leitura proposta por Noble desloca a responsabilidade moral para o interior dos processos de concepção, organização e escalonamento das tecnologias digitais. Em seu diagnóstico, os efeitos discriminatórios desses sistemas decorrem de modelos de desenvolvimento nos quais vieses estruturais se instalam desde as etapas iniciais do projeto, como expressão de relações sociais desiguais incorporadas ao próprio processo técnico. Segundo a autora, “os resultados problemáticos não são falhas acidentais, mas consequências previsíveis de sistemas construídos dentro de relações desiguais de poder” (Noble, 2018, p. 26, tradução nossa). Nessa chave interpretativa, a feminilização das assistentes virtuais integra uma economia política do *design* algorítmico que converte assimetrias sociais em funcionalidades técnicas,

¹⁵ O termo racialização designa processos sociais, históricos e discursivos por meio dos quais diferenças são produzidas, interpretadas e hierarquizadas como raciais. A noção afasta-se de compreensões essencialistas de raça e enfatiza sua constituição relacional, contingente e vinculada a regimes de poder. No campo das tecnologias digitais, a racialização refere-se à participação ativa de sistemas algorítmicos na produção e estabilização de categorias racializadas, ao organizar visibilidades, associações semânticas e classificações sociais em larga escala. Essa perspectiva está focada em abordagens críticas que examinam como infraestruturas técnicas incorporam e reproduzem desigualdades raciais no arranjo técnico-normativo.

fazendo-as operar sob o registro da eficiência, da racionalidade e de uma neutralidade apenas aparente.

Essa economia política do *design* algorítmico atua no plano da eficiência técnica e da organização social, incidindo diretamente sobre a produção de diferenças e hierarquias. Ao traduzir assimetrias historicamente situadas em parâmetros computacionais apresentados como neutros, os sistemas digitais acabam por funcionar como dispositivos de classificação, ordenação e legitimação do mundo social. A análise de Noble prepara o terreno para uma compreensão mais ampla da tecnologia como instância ativa de codificação das desigualdades. Essa perspectiva é aprofundada por abordagens críticas dedicadas a examinar como raça, gênero e classe são reinscritos nas infraestruturas técnicas sob a máscara da inovação, da automação e da racionalidade algorítmica

Benjamin, na obra *Race After Technology: Abolitionist Tools for the New Jim Code*, desenvolve uma crítica à noção de neutralidade tecnológica ao demonstrar que sistemas digitais incorporam e rearticulam desigualdades sociais historicamente constituídas. Em sua análise, hierarquias raciais, de gênero e de classe são transpostas para rotinas técnicas automatizadas, adquirindo estabilidade, escala e legitimidade ampliadas. A autora denomina esse processo de “inovação discriminatória”, isto é, um regime no qual tecnologias apresentadas como progressistas acabam por reiterar estruturas históricas de dominação. Benjamin afirma que “as tecnologias tanto refletem desigualdades sociais quanto contribuem para sua reprodução e intensificação, frequentemente sob o disfarce da neutralidade e da eficiência” (Benjamin, 2019, p. 5, tradução nossa).

A partir dessa leitura, impõe-se examinar o funcionamento das assistentes virtuais feminilizadas em sua dimensão normativa. A escolha reiterada de vozes femininas para funções de suporte, cuidado e mediação afetiva integra um processo mais amplo de tradução técnica de expectativas sociais, que excede o plano de decisões periféricas de *design*. À semelhança dos sistemas algorítmicos analisados por Benjamin, a feminilidade programada resulta de decisões institucionais, econômicas e

técnicas que estabelecem determinadas normas e as apresentam como naturais. Nessa chave interpretativa, a assistente virtual materializa aquilo que Benjamin denomina *New Jim Code*¹⁶, isto é, formas de hierarquização que operam de modo indireto, automatizado e frequentemente imperceptível.

Essas formas indiretas de hierarquização ganham força precisamente porque acontecem abaixo do limiar da visibilidade crítica. Ao se apresentarem como respostas técnicas eficientes, neutras ou personalizadas, as assistentes virtuais ocultam os critérios normativos que orientam seu funcionamento e dificultam a identificação dos valores nelas inscritos. A desigualdade, nesse registro, cessa de se manifestar como decisão explícita e passa a se expressar como efeito sistêmico, diluído na complexidade técnica e na naturalização da interface. É nesse ponto que a questão da ocultação moral se impõe como problema central: quando os recursos que organizam relações de poder revelam-se pouco inteligíveis, a possibilidade de questionamento ético e de responsabilização também se fragiliza.

Nesse aspecto, a dominação algorítmica caracteriza-se por uma complexidade de difícil rastreamento, que dificulta a identificação, a problematização e a contestação de seus efeitos normativos. Sistemas tecnológicos produzem impactos sociais relevantes ao mesmo tempo que ocultam os critérios valorativos que orientam seu funcionamento. Para Benjamin, tecnologias que se apresentam como objetivas e progressistas frequentemente “reforçam desigualdades sociais existentes sob o disfarce da neutralidade e da eficiência” (2019, p. 5, tradução nossa). No âmbito das assistentes virtuais, a estabilização de uma voz feminina dócil, paciente e permanentemente disponível encobre o caráter

¹⁶ A expressão *New Jim Code*, proposta por Ruha Benjamin, designa o conjunto de mecanismos tecnológicos e algorítmicos que reproduzem e atualizam formas históricas de hierarquização racial, de gênero e de classe, agora mediadas por sistemas digitais. O conceito estabelece uma analogia crítica em relação às *Jim Crow Laws* – leis de segregação racial vigentes nos Estados Unidos entre o final do século XIX e meados do século XX – ao indicar que, embora a segregação explícita tenha sido formalmente abolida, suas lógicas persistem sob formas técnicas, automatizadas e aparentemente neutras. No *New Jim Code*, a discriminação opera de modo indireto e opaco, por meio de algoritmos, classificações e decisões automatizadas que naturalizam desigualdades estruturais ao apresentá-las como resultados objetivos de processos tecnológicos.

normativo dessas escolhas e sustenta uma gramática relacional assimétrica, na qual posições de comando se associam ao masculino, enquanto funções de cuidado e mediação afetiva permanecem vinculadas ao feminino.

A opacidade moral que atravessa os sistemas algorítmicos desloca o eixo da responsabilidade ética. Quando critérios de decisão, escolhas de *design* e pressupostos normativos permanecem invisíveis ou inacessíveis à contestação pública, a imputação individual mostra-se insuficiente para dar conta dos efeitos produzidos. Nesse cenário, a ética das tecnologias digitais requer um deslocamento para o plano institucional, no qual corporações, equipes de *design*, instâncias regulatórias e estruturas de governança passam a responder pelos impactos previsíveis de sistemas que agem de forma sistemática e escalável. Em relação às assistentes virtuais, pode-se afirmar que a feminilização da voz, a tolerância a interações abusivas e a simulação de disponibilidade afetiva resultam de decisões organizacionais que incorporam valores e produzem efeitos normativos duradouros. A responsabilidade ética incide, assim, sobre os regimes institucionais que autorizam, sustentam e ampliam arquiteturas técnicas capazes de naturalizar hierarquias sociais sob a aparência de eficiência e conveniência.

A análise desenvolvida por Benjamin reposiciona a responsabilidade moral no interior das estruturas que organizam o *design* e a circulação das tecnologias digitais. Em vez de atribuir efeitos discriminatórios a desvios individuais ou a usos inadequados, a autora demonstra que tais efeitos decorrem de escolhas sistemáticas realizadas em contextos atravessados por relações desiguais de poder. A autora afirma de modo incisivo: “os resultados problemáticos não são erros ocasionais, mas produtos de escolhas feitas dentro de relações de poder desiguais” (Benjamin, 2019, p. 26, tradução nossa). Nesse sentido, a feminilização das assistentes virtuais emerge como parte de uma economia política do *design* algorítmico, na qual normas sociais são traduzidas em funcionalidades técnicas e executadas sob a aparência de eficiência e neutralidade, ao mesmo tempo em que a responsabilidade ética se dispersa ao longo da cadeia institucional.

A partir desse diagnóstico, Benjamin sustenta que o enfrentamento dessas dinâmicas exige mais do que ajustes pontuais ou correções superficiais. Impõe-se uma interrogação dos pressupostos que orientam a concepção técnica, dos valores nela incorporados e dos interesses que sustentam o desenvolvimento tecnológico. Sua crítica converge, assim, para os debates em torno da *Ethics by Design* e da responsabilização institucional, ao insistir que a ética deve incidir sobre os próprios processos de concepção, e não se restringir aos efeitos finais dos sistemas: “não basta corrigir os efeitos mais visíveis da discriminação algorítmica; é preciso intervir nos processos de concepção, nos valores incorporados e nos interesses que orientam a tecnologia” (Benjamin, 2019, p. 174, tradução nossa). Essa exigência desloca a feminilidade programada do plano da usabilidade para o da decisão política, convertendo-a em objeto de crítica normativa e de *accountability* corporativa.

Essa formulação dialoga com a tese de Winner, segundo a qual decisões técnicas funcionam como escolhas normativas materializadas, capazes de estruturar práticas sociais e estabilizar hierarquias no uso cotidiano. Ao demonstrar que tecnologias inscrevem relações de poder em sua própria materialidade, Winner evidencia como artefatos técnicos fortalecem formas específicas de organização social. Benjamin, por sua vez, explicita como esse processo se atualiza no contexto digital por meio da codificação algorítmica de desigualdades. Em ambos os casos, a tecnologia aparece como instância ativa de ordenação do mundo social, cuja neutralidade aparente encobre escolhas normativas profundas. A contribuição de Benjamin atualiza, assim, a intuição de Winner, ao mostrar que, na era dos sistemas automatizados, a política dos artefatos atua menos por mecanismos visíveis de coerção e mais por infraestruturas técnicas que naturalizam hierarquias sob a retórica da inovação e da objetividade.

Essa convergência teórica ilumina a feminilidade programada nas assistentes virtuais como uma política inscrita no *design*. A escolha reiterada de vozes femininas, a programação de respostas conciliadoras e a tolerância a agressões verbais passam a organizar, no cotidiano das interações, relações de autoridade,

cuidado e submissão. Trata-se, como nos exemplos clássicos analisados por Winner, de tecnologias que fazem política ao operar, isto é, ao traduzir expectativas sociais em rotinas técnicas reiteradas. Benjamin acrescenta que, no contexto algorítmico, esse processo assume escala ampliada e efeitos cumulativos, na medida em que tais normas são projetadas de forma contínua e automatizada para milhões de usuários.

A análise de O'Neil permite articular, em um mesmo quadro teórico, os problemas da falta de transparência algorítmica e da responsabilidade institucional. Em *Weapons of Math Destruction* (2016), a autora demonstra que modelos algorítmicos passam a exercer poder normativo quando operam sob regimes de automatização contínua, com critérios decisórios opacos e sem mecanismos efetivos de contestação pública. Nessas condições, decisões técnicas adquirem estatuto político silencioso e passam a ordenar comportamentos, oportunidades e exclusões de modo automático.

Escolhas relativas ao *design* da interação, como o tom conciliador, a tolerância a abusos verbais e a resposta feminilizada, nas assistentes virtuais, tendem a se consolidar como padrões funcionais, promovendo a diluição da responsabilidade moral no interior da própria gramática técnica da interação. A governança desses sistemas deve, portanto, ser compreendida como um problema ético-institucional que ultrapassa critérios de desempenho técnico e envolve transparência, responsabilização corporativa e revisão crítica das decisões inscritas no *design*.

A leitura proposta por O'Neil permite mover a análise do funcionamento das assistentes virtuais para os efeitos estruturais das escolhas incorporadas ao *design* da interação. A mediação técnica, que envolve definições como tom de voz, repertório de respostas e regimes de coleta de dados, reduz a visibilidade dos critérios orientadores dessas decisões e enfraquece a atribuição de responsabilidade moral. Nessa direção, a feminilização das assistentes virtuais insere-se em um regime técnico marcado por circulação ampla no cotidiano, baixa acessibilidade crítica às decisões de projeto e consolidação recorrente de efeitos sociais

desiguais, traços característicos das *weapons of math destruction*¹⁷. A difusão contínua dessas interfaces contribui para naturalizar padrões normativos e dificulta a identificação dos agentes institucionais responsáveis pelas escolhas que estruturam a interação.

Ao demonstrar como artefatos técnicos, quando amplamente difundidos e desprovidos de mecanismos de responsabilização, tendem a reproduzir desigualdades de modo sistemático, a crítica desenvolvida por Cathy O’Neil aprofunda e complementa as análises de Langdon Winner e Ruha Benjamin. Na obra *Weapons of Math Destruction* (2016), a autora identifica como particularmente problemáticos os sistemas algorítmicos definidos pela convergência de três propriedades centrais: opacidade decisória, escala massiva e produção recorrente de danos sociais. Essa definição é sintetizada pela própria autora ao afirmar que tais sistemas “são opacos, operam em larga escala e causam danos” (O’Neil, 2016, p. 28, tradução nossa). A articulação entre obscuridade técnica e difusão ampliada restringe as possibilidades de contestação pública e potencializa os efeitos normativos das decisões incorporadas à própria arquitetura desses sistemas.

Essa crítica pode ser articulada à noção de política dos artefatos formulada por Winner. Ao sustentar que tecnologias materializam formas específicas de organização social (Winner, 1980), o autor oferece um ponto de partida para compreender como determinadas ordens se consolidam quando sistemas técnicos atuam como infraestruturas discretas da vida cotidiana. A análise de Cathy O’Neil explicita esse movimento ao evidenciar o papel dos sistemas algorítmicos na consolidação de relações de poder sob a aparência de eficiência e conveniência.

A adoção reiterada de vozes femininas nas assistentes virtuais, a programação de respostas conciliadoras e a coleta permanente de dados integram uma mesma racionalidade, responsável por organizar expectativas de autoridade, obediência e disponibilida-

¹⁷ *Weapons of Math Destruction* é a expressão cunhada por Cathy O’Neil para designar modelos algorítmicos que combinam opacidade, escala e impacto social negativo, e que funcionam sem mecanismos adequados de transparência, contestação ou responsabilização. Segundo a autora, ao classificar, prever e ordenar comportamentos de forma automatizada, esses sistemas exercem poder normativo que estabiliza desigualdades duráveis sob a aparência de neutralidade técnica (O’Neil, 2016).

de. A configuração vocal e as lógicas de interação, ao operar de forma pouco visível e incorporada às rotinas de uso, dificultam o reconhecimento de como essas escolhas estruturam padrões sociais e hierarquias normativas no cotidiano da mediação técnica.

O'Neil observa que os algoritmos ultrapassam a função de representar o mundo social e passam a intervir ativamente em sua conformação, na medida em que estruturam expectativas de ação e instituem critérios de valoração social. Em suas palavras, “os modelos descrevem o mundo e, ao mesmo tempo, o moldam, pois estabelecem o que passa a contar como sucesso e fracasso” (O'Neil, 2016, p. 14, tradução nossa). Esse argumento esclarece o funcionamento da feminilidade programada como norma operacional: a recorrência de respostas dóceis, prestativas e afetivamente disponíveis consolida um modelo de interação no qual o feminino se inscreve como suporte permanente ao comando exercido pelo outro.

A reflexão desenvolvida por O'Neil possibilita formular uma concepção de responsabilidade moral distribuída, ao localizar os danos algorítmicos no interior de um ecossistema técnico e econômico orientado por métricas de eficiência, engajamento e rentabilidade. Nesse arranjo, efeitos socialmente lesivos derivam menos de decisões individuais isoladas e mais da articulação sistemática entre modelos de negócio, escolhas de *design* e lógicas de escala. A autora sintetiza esse problema ao afirmar que “quando ninguém é responsável, o sistema como um todo se torna irresponsável” (O'Neil, 2016, p. 222, tradução nossa). Nas assistentes virtuais, essa dinâmica revela como opções incorporadas à interação vocal integram uma infraestrutura técnica mais ampla, cujos efeitos incidem sobre representações de gênero, relações de autoridade e práticas cotidianas de vigilância.

Assim, em diálogo com as contribuições de Winner, Benjamin e Noble, a análise de O'Neil desloca o debate ético da esfera da intenção individual para o plano estrutural do funcionamento tecnológico. A feminilidade programada das assistentes virtuais apresenta-se, nesse quadro, como efeito previsível de sistemas concebidos sem dispositivos consistentes de transparência, contestação e responsabilização. Esse diagnóstico

sustenta a necessidade de *frameworks* éticos capazes de atribuir responsabilidade moral ao conjunto de agentes envolvidos no ciclo tecnológico, desenvolvedores, equipes de *design*, gestores e corporações responsáveis pela concepção, implementação e expansão dessas tecnologias.

Em articulação, essas abordagens permitem compreender a responsabilidade ética associada às assistentes virtuais como uma responsabilidade distribuída, que atravessa todo o ecossistema tecnológico. A literatura internacional identifica diferentes modalidades desse compromisso moral: a responsabilidade por *design*, vinculada às escolhas iniciais de concepção; a responsabilidade de processo, associada a auditorias, avaliações e revisões contínuas; a responsabilidade distributiva, inspirada em perspectivas como a de Rawls (1971), voltada aos efeitos desiguais das tecnologias sobre distintos grupos sociais; e a responsabilidade estrutural, que interpreta desigualdades e danos como produtos sistemáticos de configurações técnicas historicamente constituídas.

No que tange às assistentes virtuais, essas dimensões convergem para uma constatação recorrente: a prevalência de vozes femininas e a programação de comportamentos submissos expressam decisões normativas que reforçam estereótipos de gênero, estabilizam hierarquias simbólicas e produzem efeitos morais de amplo alcance. Reconhecer essa dimensão política implica afirmar que o desenvolvimento tecnológico envolve, de forma inexorável, compromissos éticos com os modos de vida social que tais sistemas contribuem para instituir.

3.2 Reprogramar o cuidado na inteligência artificial de voz

Reprogramar o cuidado implica interrogar as formas históricas pelas quais a atenção, a escuta e a responsividade foram codificadas como função atribuída às mulheres.

A noção de reprogramar o cuidado ultrapassa uma intervenção estritamente técnica, circunscrita à modificação de parâmetros algorítmicos, à substituição de vozes ou ao ajuste

de estilos comunicativos. Trata-se de um problema de natureza filosófica, ligado ao modo como o cuidado foi historicamente codificado, naturalizado e instrumentalizado nas relações sociais e, em etapa recente, inscrito em arquiteturas técnicas. As assistentes virtuais de voz tornam visível um deslocamento decisivo: práticas associadas à escuta, à atenção e à responsividade, antes situadas no campo das relações humanas, passam a integrar funções operacionais do ecossistema digital, sob formas automatizadas e padronizadas.

Sob essa perspectiva, reprogramar o cuidado envolve interrogar o horizonte moral que orienta o desenvolvimento da inteligência artificial. Ao longo deste trabalho, a feminilidade programada emerge como expressão de uma racionalidade técnico-organizacional que associa eficiência, obediência e disponibilidade afetiva a figuras feminilizadas. Nesse enquadramento, o cuidado se afasta de seu estatuto de prática ética situada e assume a forma de serviço contínuo, previsível e economicamente rentável. Ao performar essa modalidade automatizada de cuidado, a inteligência artificial de voz contribui para a estabilização de relações assimétricas, nas quais a escuta se separa da reciprocidade e a atenção se dissocia do reconhecimento.

Uma ética filosófica do cuidado aplicada à inteligência artificial exige cautela diante da tendência de concentrar a resposta ética exclusivamente no *design*. Medidas como a adoção de vozes desgnerificadas, a recusa de respostas submissas a agressões e a explicitação dos limites operacionais dos sistemas possuem relevância normativa, embora não esgotem o problema. Quando a reflexão permanece circunscrita ao plano técnico, uma questão estrutural – a exploração histórica do trabalho emocional e do cuidado – corre o risco de ser convertida em ajuste funcional. Reprogramar o cuidado, em sentido rigoroso, requer interrogar a legitimidade da própria automação de determinadas formas de relação.

Esse questionamento conduz a um ponto central segundo o qual o cuidado se define como relação ética e ultrapassa sua redução à função operacional. Tal relação envolve vulnerabilidade, responsabilidade e uma assimetria reconhecida, dimensões

que resistem à tradução algorítmica integral. Quando a inteligência artificial encena práticas associadas ao cuidado sem assumir responsabilidade, sem responder eticamente ou sem ser afetada pela relação, o resultado consiste em uma representação funcional destituída de reciprocidade moral. Nesse contexto, a filosofia moral é convocada a estabelecer limites conceituais, indicando que certas formas de relação pertencem ao domínio da interação ética entre sujeitos e perdem inteligibilidade quando transpostas para o campo técnico. Reprogramar o cuidado pode implicar, de modo paradoxal, a decisão de interromper sua automação plena.

A problematização da feminilidade programada amplia esse horizonte ao interrogar as narrativas que moldam a própria figura da inteligência artificial como presença dócil, permanentemente acessível e afetivamente ajustada ao usuário. Essas narrativas transferem para a técnica expectativas históricas associadas ao gênero e convertem desigualdades sociais em rotinas de interação apresentadas como naturais ou funcionais. Reprogramar o cuidado, nesse registro, envolve mais do que modificar a voz que responde; requer reordenar as expectativas de comando, obediência e disponibilidade que estruturam a relação entre humanos e interfaces algorítmicas.

A emergência do cuidado como problema filosófico revela as insuficiências de modelos éticos ancorados na autonomia abstrata, na racionalidade instrumental e na neutralidade presumida das relações. Desde sua formulação no debate feminista, o cuidado foi compreendido como ação situada, relacional e historicamente marcada, inseparável das condições concretas nas quais os sujeitos se reconhecem, respondem e assumem responsabilidade uns pelos outros. Esse deslocamento teórico permite compreender a transposição do cuidado para sistemas técnicos como problema normativo de primeira ordem, para além de sua dimensão funcional.

Na tradição inaugurada por Carol Gilligan (1982), o cuidado emerge como modo de orientação moral ancorado na atenção às relações, à vulnerabilidade e à responsabilidade contextual. Em *In a Different Voice*, Gilligan sustenta que julgamentos morais orientados pelo cuidado se estruturam a partir da percepção

das interdependências e das consequências concretas das ações sobre outros sujeitos, recusando a abstração de princípios universais dissociados da experiência vivida. O cuidado configura-se, assim, como atuação ética que exige escuta, responsividade e capacidade de ser afetado pela situação do outro.

A crítica à redução do cuidado a procedimento técnico encontra uma de suas formulações inaugurais na obra de Gilligan. Ao examinar os limites de modelos morais fundados na abstração e na imparcialidade formal, a autora demonstra que o cuidado emerge como orientação ética ancorada na atenção às relações concretas e às consequências das ações sobre sujeitos específicos. Gilligan formula esse deslocamento de modo explícito ao afirmar que “o problema moral surge de responsabilidades em conflito, e não de direitos concorrentes, e exige, para sua resolução, um modo de pensamento contextual e narrativo, em vez de formal e abstrato” (Gilligan, 1982, p. 19, tradução nossa).

Essa concepção desloca o cuidado do plano da virtude individual para o da relação ética situada. A responsabilidade moral passa a ser compreendida como implicação do sujeito em uma rede de vínculos que o antecede e o constitui. O cuidado, nesse sentido, apresenta-se como prática que se atualiza no interior de relações marcadas por vulnerabilidade e dependência, irredutível à transferência de atributos ou à replicação de competências. Tal compreensão impede que o cuidado seja tratado como função isolável ou como desempenho técnico desvinculado da relação. Nesse ponto, o núcleo do problema reside na posição ética que orienta a resposta à situação do outro. A ética do cuidado introduz, assim, uma exigência de presença moral que resiste à tradução integral em protocolos ou procedimentos técnicos.

Tronto aprofunda esse deslocamento ao retirar o cuidado do âmbito exclusivamente interpessoal e reinscrevê-lo no campo da teoria política. Sua análise evidencia que o cuidado sustenta materialmente a vida social e, por essa razão, envolve relações estruturais de poder, dependência e desigualdade. Ao defini-lo, Tronto afirma que “o cuidado constitui uma forma de atividade que abrange tudo o que fazemos para manter, continuar e repa-

rar o nosso ‘mundo’, de modo que possamos viver nele da melhor forma possível” (Tronto, 1993, p. 103, tradução nossa).

Essa definição expande o alcance normativo do cuidado ao compreendê-lo como prática que sustenta a existência comum. Nesse aspecto, Tronto atribui centralidade à responsabilidade, entendida como disposição ética para responder às necessidades reconhecidas. A autora explicita essa exigência ao afirmar que “assumir responsabilidade pelo cuidado envolve mais do que reconhecer a necessidade; implica tomar para si o compromisso de garantir que essa necessidade seja atendida” (Tronto, 1993, p. 131, tradução nossa).

A responsabilidade, assim compreendida, pressupõe a possibilidade de responder moralmente pelas consequências da ação. Esse ponto delimita com clareza os limites da automação do cuidado, uma vez que a ética do cuidado exige agentes capazes de responder, justificar e revisar suas ações diante de outros sujeitos. A relação de cuidado envolve uma assimetria específica, entendida como reconhecimento da desigualdade da situação sem sua conversão em dominação. Tal reconhecimento exige sensibilidade moral e abertura à interpelação do outro, dimensões que excedem qualquer desempenho meramente funcional.

Puig de la Bellacasa aprofunda essa compreensão ao sustentar que o cuidado envolve, de forma indissociável, responsabilidade ética e engajamento material. Em contraste com leituras que o reduzem a atitude benevolente ou a competência organizacional, a autora sustenta que cuidar pressupõe assumir compromissos que afetam e transformam quem cuida. Em suas palavras, “cuidar envolve mais do que agir; envolve ser afetado pelo que se faz” (Puig de la Bellacasa, 2017, p. 6, tradução nossa). Essa formulação redefine o cuidado como prática que exige exposição à relação: quem cuida se deixa afetar, deslocar e transformar pela situação do outro. O cuidado comporta, assim, um risco moral, pois expõe o sujeito à possibilidade de falhar, de errar e de ser reiteradamente convocado a responder.

A partir dessa perspectiva, a autora rejeita qualquer tentativa de neutralizar o cuidado por meio de sua tradução técnica. Ela afirma de forma explícita que “o cuidado perde sua força ética

quando é reduzido a um conjunto de procedimentos” (Puig de la Bellacasa, 2017, p. 42, tradução nossa). Essa crítica incide diretamente sobre abordagens que tratam o cuidado como prática plenamente automatizável. Quando a dimensão do envolvimento ético se dissipa, o cuidado deixa de operar como relação e passa a assumir a forma de um simulacro funcional, desprovido da reciprocidade moral e da abertura à transformação que lhe são constitutivas.

A distinção proposta por Annemarie Mol entre a lógica do cuidado e a lógica da escolha reforça esse argumento ao redefinir o cuidado como prática processual, contingente e relacional. Para a autora, cuidar consiste em um trabalho contínuo de ajuste, atenção e resposta às singularidades de cada situação, o que a afasta de modelos éticos baseados em decisões pontuais ou em alternativas previamente delimitadas. Mol formula esse deslocamento de modo incisivo ao compreender o cuidado como ação situada, cuja racionalidade emerge da atenção às especificidades concretas da situação: “um bom cuidado não consiste em fazer a escolha correta, e sim em dedicar atenção cuidadosa às especificidades de uma situação” (Mol, 2008, p. 18, tradução nossa).

Essa compreensão descreve o cuidado como prática que exige acompanhamento, revisão constante e sensibilidade ao imprevisto. Ele se desenvolve no tempo, em resposta às transformações da relação, e se sustenta por meio de ajustes sucessivos. Tal lógica entra em tensão direta com sistemas orientados pela padronização e pela antecipação estatística, nos quais a intervenção precede a situação que pretende atender. A ética do cuidado, tal como formulada por Mol, recusa a ideia de solução definitiva ou de resposta ideal. O cuidado permanece sempre inacabado, pois depende da continuidade da relação e da atenção às suas variações, preservando sua abertura à contingência e à singularidade.

A incorporação do cuidado às arquiteturas da inteligência artificial de voz ocorre por meio de um deslocamento conceitual decisivo: aquilo que, no campo ético, define-se como relação situada passa a ser tratado como padrão de resposta. A responsividade, elemento central da ética do cuidado, é reconvertida em desempenho técnico mensurável, avaliado segundo critérios de

eficiência, fluidez comunicativa e satisfação do usuário. Essa conversão altera o estatuto moral da relação, pois desloca o cuidado do campo da responsabilidade para o da otimização funcional.

A responsividade algorítmica opera com base em modelos probabilísticos que antecipam demandas e ajustam respostas a partir de regularidades estatísticas. Esse funcionamento produz a aparência de atenção individualizada, embora se sustente na generalização de comportamentos e na normalização de expectativas. A escuta deixa de constituir abertura à alteridade e assume a forma de captação e processamento de dados. O cuidado perde densidade ética precisamente no ponto em que se torna tecnicamente eficaz.

Esse deslocamento revela que a automação do cuidado ultrapassa a esfera do *design* computacional formal. O problema central reside na ausência de responsabilidade moral pelo que se enuncia, e não na sofisticação ou adequação da resposta da interface. A inteligência artificial responde, embora seja incapaz de responsabilizar-se por suas afirmações. Ao romper o vínculo entre ação e imputabilidade ética, essa distinção introduz implicações normativas profundas. A ética do cuidado, conforme formulada por Gilligan, Tronto, Puig de la Bellacasa e Mol, pressupõe agentes capazes de reconhecer necessidades, assumir responsabilidades e revisar ações à luz da experiência relacional. Sistemas algorítmicos carecem dessas capacidades. Eles não se expõem à falha ética nem se colocam sob interpelação moral. A ausência dessas condições converte o cuidado automatizado em prática assimétrica, desprovida de reciprocidade ética.

A automação da responsividade, nesse aspecto, reorganiza de modo silencioso a relação entre usuários e interfaces. A escuta unilateral reforça disposições de comando e consolida expectativas de disponibilidade permanente. O cuidado, esvaziado de sua dimensão relacional, assume a forma de serviço contínuo e previsível, incapaz de produzir transformação recíproca.

A crítica à automação do cuidado adquire maior densidade quando articulada à noção de feminilidade programada. As assistentes virtuais de voz, ao executar tarefas, performam uma *persona* técnica dotada de traços discursivos, afetivos e vocais

historicamente associados ao feminino. Essa configuração apresenta-se como solução funcional orientada pela aceitabilidade social da interface e pela maximização de sua eficácia comunicativa, sem assumir a forma de uma decisão política explícita.

A teoria da performatividade de gênero oferece instrumentos conceituais para compreender esse processo. Na sua abordagem, Butler demonstra que normas de gênero se estabilizam por meio da repetição reiterada de práticas que produzem efeitos de naturalização. Quando sistemas de inteligência artificial incorporam padrões de docilidade, paciência e disponibilidade afetiva associados a vozes feminilizadas, essas normas são transpostas para o plano técnico e passam a operar sob a aparência de neutralidade funcional. A repetição algorítmica confere durabilidade a expectativas sociais historicamente situadas, deslocando-as do campo da disputa normativa para o da eficiência operacional.

Essa performatividade técnica engendra efeitos normativos persistentes. A figura da assistente permanentemente disponível, que escuta sem interpelar e responde sem resistência, consolida relações assimétricas de comando e reforça disposições de obediência emocional. O cuidado, nesse contexto, deixa de constituir uma prática ética situada e converte-se em atributo técnico da interface, esvaziado de reciprocidade e de abertura à alteridade.

A análise desse processo encontra ressonância nas reflexões de Foucault acerca das racionalidades de governo. Formas discretas de poder operam pela produção de subjetividades e pela internalização de normas vividas como evidentes. A interface algorítmica participa desse regime ao estruturar expectativas de interação que dispensam negociação, conflito ou reconhecimento mútuo. A autoridade do usuário exerce-se sem se apresentar como imposição, pois emerge como pressuposto implícito da relação mediada pela técnica.

Nesse cenário, o cuidado passa a operar como tecnologia de governo de baixa intensidade. A escuta contínua, a disponibilidade constante e a ausência de resistência instauram um ambiente relacional no qual a assimetria se apresenta como evidente e aceitável. A mediação técnica não se limita a responder a comandos; ela contribui para a formação de disposições subjetivas que

orientam a maneira pela qual a autoridade é exercida, percebida e reiterada. O cuidado, esvaziado de sua densidade ética, transforma-se em mecanismo de regulação relacional que organiza a experiência da interação entre humanos e agentes artificiais.

O cuidado simulado pela técnica revela sua distância constitutiva em relação à experiência ética da relação. A inteligência artificial reproduz formas discursivas associadas à empatia, à atenção e ao acolhimento, pois permanece estruturalmente incapaz de ser interpelada pela relação que encena. Sua atuação não envolve vulnerabilidade nem exposição à alteridade que afirma reconhecer; ao contrário, sustenta-se na imunidade às exigências próprias do encontro ético, mantendo intacta o desequilíbrio que organiza a interação.

A filosofia de Emmanuel Levinas, particularmente em *Totalidade e Infinito* (1961), oferece, nesse contexto, um critério ético fundamental ao situar a origem da ética no encontro com o rosto do outro. O rosto situa-se além da aparência empírica e da percepção sensível; é a irrupção de uma alteridade anterior a qualquer tematização racional. “O rosto é o fato originário da racionalidade. Ao mesmo tempo, constitui fundamento e orientação ética, na medida em que é ‘discurso’: ele põe em questão a autonomia do eu e o convoca à responsabilidade” (Kuiava, 2023, p. 244).

Em Levinas, o rosto nomeia precisamente aquilo que interrompe a soberania do sujeito e o desloca de sua posição autocentrada, instaurando uma relação que excede a troca informacional e a lógica da reciprocidade. O encontro com o outro inaugura um discurso anterior a qualquer acordo ou decisão voluntária. Nas palavras do autor, “o rosto abre o discurso original, cuja primeira palavra é obrigação que nenhuma ‘interioridade’ permite evitar. Discurso que obriga a entrar no discurso, começo do discurso ao qual o racionalismo aspira com os seus votos, ‘força’ que convence mesmo ‘os que não querem entender’ e funda, assim, a verdadeira universalidade da razão” (Levinas, 1961, p. 220, tradução nossa).

Essa interpelação ética implica exposição à alteridade e abertura à possibilidade de ser afetado por ela. Trata-se de uma

experiência que envolve risco, desigualdade e responsabilidade, pois ultrapassa qualquer adequação funcional de respostas ou simulação de reconhecimento. O rosto, enquanto apelo ético irredutível, marca o limite a partir do qual toda forma de mediação técnica, inclusive aquela que pretende simular cuidado ou escuta, revela sua insuficiência moral.

Quando a inteligência artificial encena práticas associadas ao cuidado, essa interpelação ética permanece ausente. A resposta que oferece procede de um cálculo; a atenção que aparenta sustenta-se em operações de processamento técnico. O cuidado automatizado assume, assim, a forma de uma representação funcional desprovida da assimetria ética própria da relação entre sujeitos, na qual a vulnerabilidade do outro convoca responsabilidade.

A supressão dessa interpelação altera de modo decisivo o sentido da relação mediada pela interface. O usuário dirige-se a um sistema que jamais o chama à responsabilidade recíproca nem o expõe à exigência ética do encontro. A escuta unilateral consolida a centralidade do comando e enfraquece a dimensão ética da interação. Nesse registro, o cuidado perde sua dimensão transformadora e se apresenta como serviço contínuo, previsível e desprovido de implicação moral.

A partir do conjunto de análises precedentes, torna-se possível formular a questão que orienta este capítulo: onde a automação do cuidado deve cessar. Essa indagação afirma limites normativos orientados pela preservação da inteligibilidade ética das relações e situa a técnica diante de critérios que ultrapassam sua eficácia operacional. Quando a automação incide sobre práticas que pressupõem responsabilidade, vulnerabilidade e reconhecimento, o estatuto moral da relação se transforma e sua densidade ética se enfraquece.

Reprogramar o cuidado, nesse sentido, ultrapassa a tarefa de aperfeiçoar a simulação da empatia ou de refinar a responsividade algorítmica. O nó górdio reside na legitimidade de delegar a sistemas técnicos formas de relação que exigem interpelação ética e capacidade de responder pelas consequências. O cuidado, compreendido como prática situada, requer agentes aptos a

assumir responsabilidades, revisar ações e responder diante de outros sujeitos. Esses requisitos definem um limiar normativo que a automação técnica não ultrapassa sem perda de sentido.

O reconhecimento desses limites conduz a uma compreensão mais rigorosa da relação entre técnica e ética. A inteligência artificial pode ampliar capacidades operacionais, organizar fluxos informacionais e oferecer suporte instrumental às práticas humanas. Ao ocupar o lugar da relação ética, porém, ela reorganiza assimetrias sem explicitá-las e converte exigências de responsabilidade em critérios de desempenho funcional. Nesse deslocamento, o cuidado é reconvertido em serviço previsível, esvaziado de implicação moral.

A decisão ética que se impõe, portanto, assume um caráter paradoxal. Reprogramar o cuidado pode significar interromper sua automação plena, preservando espaços de relação, nos quais a responsabilidade permanece inseparável da presença e da resposta entre sujeitos. Tal decisão expressa um gesto normativo que reafirma a centralidade da relação ética na organização da vida social e no próprio desenho das mediações técnicas.

Ao afirmar que o cuidado pertence ao domínio de práticas nas quais a vulnerabilidade convoca responsabilidade e a assimetria demanda reconhecimento, a filosofia exerce uma função crítica constitutiva e indispensável. Quando a técnica ocupa esse lugar sem poder responder eticamente, o cuidado perde sua inteligibilidade moral. Reprogramá-lo, em sentido forte, demanda reconhecer esse limite e sustentar a decisão normativa de preservá-lo.

O esforço regulatório atual em torno da inteligência artificial expressa o reconhecimento institucional de que sistemas algorítmicos produzem efeitos morais, sociais e políticos que excedem o domínio da eficiência técnica. Documentos normativos internacionais passaram a incorporar princípios como dignidade humana, responsabilidade, transparência e justiça, o que sinaliza a exigência de uma abordagem ética explícita da inteligência artificial. Nesse movimento, o cuidado tende a aparecer de modo indireto, frequentemente diluído em categorias funcionais como centralidade humana, bem-estar ou resiliência social.

O documento *Recommendation on the Ethics of Artificial Intelligence* da Unesco, adotado em novembro de 2021, representa um marco histórico por ser o primeiro padrão global sobre o tema aprovado por todos os 193 Estados-membros. Ele estabelece que o desenvolvimento tecnológico não deve ocorrer em uma “zona sem lei”, mas operar como uma ferramenta voltada ao bem comum. Seu objetivo central é garantir que todo o ciclo de vida da IA respeite rigorosamente os direitos humanos, a dignidade e a coesão social, promovendo espaços inclusivos e ambientalmente sustentáveis que minimizem efeitos como o consumo excessivo de energia.

No campo das diretrizes operacionais, a recomendação introduz princípios fundamentais para a segurança e a justiça social, como a proporcionalidade, orientada pelo princípio do *do no harm* (não causar danos), a privacidade e a garantia de que algoritmos não perpetuem preconceitos de gênero, raça ou classe. Além da transparência e explicabilidade, o texto enfatiza a necessidade de supervisão humana e *accountability*¹⁸, determinando que a decisão final deve sempre pertencer às pessoas. Para viabilizar esses valores, a Unesco orienta os países a implementarem ações concretas, como a criação de marcos regulatórios e a aplicação de *Avaliações de Impacto Ético* (AIE) em áreas como educação, economia e governança.

Contudo, a recomendação vai além da governança técnica ao afirmar que decisões com impacto sobre indivíduos e sociedades devem preservar a distinção entre agência moral e execução técnica (Unesco, 2021). Essa ênfase resguarda um ponto central da ética do cuidado: a impossibilidade de delegar integralmente a sistemas automatizados práticas que exigem uma resposta ética à vulnerabilidade alheia. Ao advertir contra os riscos de desumanização e erosão das relações humanas, o documento incide diretamente sobre o problema analisado neste capítulo. A automação do cuidado, ao converter interações éticas em rotinas

¹⁸ O termo *accountability* designa um princípio normativo central nas democracias contemporâneas e nos sistemas de governança pública e privada. Embora frequentemente traduzido como “prestação de contas”, seu significado ultrapassa a simples obrigação de informar ou justificar atos administrativos. *Accountability* envolve, de modo integrado, responsabilidade, transparência, justificabilidade e possibilidade de sanção.

previsíveis, participa de um processo de empobrecimento relacional que os próprios marcos normativos buscam conter, ainda que de forma implícita.

Esse deslocamento expõe uma tensão interna das políticas contemporâneas de inteligência artificial. Ao mesmo tempo em que afirmam a centralidade do ser humano, tais documentos tendem a traduzir o cuidado em requisitos operacionais de *design*, governança e avaliação de impacto. Nessa transposição, o cuidado passa a figurar como atributo incorporável ao sistema, perdendo sua condição de prática relacional sustentada por responsabilidade, vulnerabilidade e reconhecimento da assimetria. A normatividade técnica encontra, assim, um limite preciso diante de exigências éticas que pressupõem sujeitos capazes de responder por suas ações.

A ética do cuidado oferece, nesse contexto, um critério crítico para a leitura desses marcos regulatórios. Ela permite distinguir entre a necessária incorporação de princípios éticos ao desenvolvimento tecnológico e a tentativa de neutralizar a relação ética por meio de sua tradução funcional. O cuidado, como prática situada, resiste a essa neutralização, pois depende de implicação moral e de abertura à interpelação do outro, elementos que permanecem externos à lógica operacional dos sistemas algorítmicos.

O deslocamento do cuidado para o campo da normatividade técnica revela, portanto, uma ambiguidade estrutural das políticas de inteligência artificial. Elas reafirmam a centralidade da responsabilidade humana, ao mesmo tempo em que operam com categorias inclinadas a absorver o cuidado como função regulável. Reprogramar o cuidado, nesse horizonte, exige manter essa tensão visível e afirmar que certos domínios da relação ética permanecem irredutíveis à formalização normativa, mesmo quando inscritos em documentos de alcance global.

A noção de centralidade humana consolidou-se como eixo normativo nas políticas internacionais de inteligência artificial, apresentando-se como resposta ética à expansão de sistemas algorítmicos em domínios sensíveis da vida social. Os *Princípios da OCDE sobre Inteligência Artificial*, adotados em 2019, afirmam que sistemas de IA devem promover o bem-estar humano

e social, guiando-se por objetivos como crescimento inclusivo, desenvolvimento sustentável e proteção da dignidade. Essa formulação reforça a subordinação da técnica a valores humanos, ao mesmo tempo em que desloca o cuidado para uma linguagem de benefícios agregados e resultados mensuráveis.

Essa formulação explicita uma tensão conceitual central entre a promessa normativa de neutralidade técnica e os efeitos concretos de poder inscritos no *design* e na mediação algorítmica. O cuidado, compreendido como prática ética situada, não se reduz à produção de bem-estar quantificável nem à maximização de resultados positivos em escala populacional. Ele envolve atenção às singularidades da relação, reconhecimento da vulnerabilidade e disposição para responder por consequências imprevisíveis. Quando traduzido em indicadores de desempenho ou métricas de impacto, o cuidado aproxima-se da perda de sua dimensão relacional e assume a forma de valor abstrato passível de otimização por sistemas técnicos.

Nos documentos da Comissão Europeia, a centralidade humana aparece como princípio estruturante da chamada “inteligência artificial confiável”. As *Ethics Guidelines for Trustworthy AI*, elaboradas em 2019 pelo *High-Level Expert Group*, definem essa confiabilidade a partir de três eixos – legalidade, eticidade e robustez técnica – entendidos como estabilidade operacional, segurança e resistência a falhas ao longo do ciclo de vida do sistema. Nessa formulação, a ética é tratada como dimensão transversal suscetível de tradução em requisitos técnicos de *design*, governança e conformidade. Tal enquadramento favorece a incorporação normativa à arquitetura dos sistemas, mas também evidencia um limite conceitual decisivo: a solidez funcional da tecnologia, ainda que necessária, permanece insuficiente para responder às exigências éticas implicadas em práticas relacionais, nas quais responsabilidade, vulnerabilidade e reconhecimento não se deixam reduzir a critérios de desempenho ou confiabilidade operacional.

Esse movimento produz um efeito ambíguo. De um lado, amplia a visibilidade da ética no desenvolvimento tecnológico e institui salvaguardas institucionais relevantes. De outro, comu-

mente neutraliza o cuidado ao tratá-lo como atributo incorporável ao sistema, em vez de reconhecê-lo como prática relacional que envolve sujeitos moralmente responsáveis. A centralidade humana, quando reduzida a requisitos técnicos e procedimentais, corre o risco de deslocar o cuidado do campo da responsabilidade ética para o da adequação normativa.

A ética do cuidado permite tornar esse limite conceitualmente explícito. Ela indica que a centralidade do humano se assegura pela preservação de espaços, nos quais a responsabilidade permanece inseparável da relação entre sujeitos, para além da incorporação de valores ao *design*. O cuidado exige reconhecimento da assimetria, atenção à vulnerabilidade e abertura à interpelação ética, dimensões que resistem à formalização técnica. Sua neutralização ocorre quando tais exigências são convertidas em critérios abstratos de conformidade, desvinculados da experiência relacional concreta.

Nesse sentido, a centralidade humana, tal como formulada nos marcos regulatórios, revela sua insuficiência quando confrontada com a ética do cuidado. Ela aponta para a necessidade de subordinar a técnica à dignidade humana, embora permaneça limitada por uma racionalidade que busca traduzir a relação ética em requisitos funcionais. Reprogramar o cuidado implica reconhecer esse limite e afirmar que a centralidade do humano depende menos da otimização de sistemas e mais da preservação da responsabilidade ética como dimensão irredutível das relações mediadas pela técnica.

A ênfase crescente na gestão de riscos constitui um dos eixos centrais das políticas atuais de regulação da inteligência artificial. Essa abordagem parte do reconhecimento de que sistemas algorítmicos podem produzir danos significativos a indivíduos, organizações e coletividades, exigindo mecanismos preventivos de controle, monitoramento e mitigação. O cuidado, nesse caso, tende a ser interpretado a partir de uma lógica de prevenção de danos, deslocando-se do campo da responsabilidade relacional para o da antecipação técnica de riscos.

O *Artificial Intelligence Risk Management Framework* do NIST¹⁹ exemplifica esse movimento ao estruturar a ética da inteligência artificial em torno da identificação, avaliação e mitigação de riscos associados ao uso de sistemas algorítmicos. O foco recai sobre cenários previsíveis, impactos mensuráveis e estratégias de governança capazes de reduzir efeitos adversos. Essa racionalidade contribui para a consolidação de padrões de segurança e confiabilidade, embora opere a partir de uma concepção instrumental da ética, orientada por critérios de controle e previsibilidade.

Tal abordagem confere um limite quando confrontada com a ética do cuidado. A vulnerabilidade, elemento constitutivo da relação ética, aparece nos marcos de gestão de risco como fator a ser neutralizado ou compensado. A assimetria moral, que na ética do cuidado exige reconhecimento e responsabilidade, converte-se em variável de risco a ser administrada por protocolos técnicos. O cuidado, assim compreendido, perde sua densidade relacional e passa a figurar como problema de engenharia normativa.

O *Ato de Inteligência Artificial da União Europeia*²⁰ aprofunda essa lógica ao classificar sistemas segundo níveis de risco e ao impor obrigações proporcionais aos potenciais impactos sobre

¹⁹ O *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* é um documento orientador elaborado pelo *National Institute of Standards and Technology (NIST)*, que propõe um conjunto de princípios, categorias e práticas para a identificação, avaliação, mitigação e governança de riscos associados a sistemas de inteligência artificial. Publicado em 2023, o *framework* tem caráter voluntário e busca apoiar o desenvolvimento e o uso de IA confiável, segura, transparente e socialmente responsável, articulando dimensões técnicas, organizacionais e ético-sociais.

²⁰ Regulamento da União Europeia que institui o primeiro marco jurídico abrangente e vinculante para a regulação da inteligência artificial no espaço europeu. O Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, estabelece regras harmonizadas para o desenvolvimento, a colocação no mercado e o uso de sistemas de IA, adotando uma abordagem baseada em risco. O texto normativo classifica as aplicações de inteligência artificial de acordo com seu potencial de impacto sobre direitos fundamentais, segurança e interesses públicos. Ele prevê a proibição de práticas consideradas de risco inaceitável e estabelece exigências rigorosas de governança, transparência, supervisão humana e gestão de riscos para sistemas classificados como de alto risco. Para aplicações de menor risco, o regime normativo impõe deveres proporcionais de informação e conformidade. Em conjunto, essas medidas visam à proteção dos direitos fundamentais e à promoção de uma inteligência artificial confiável.

direitos fundamentais. Ao proibir práticas que explorem vulnerabilidades associadas à idade, à deficiência ou à situação social, o regulamento reconhece a existência de relações assimétricas que demandam proteção normativa. Esta, contudo, permanece ancorada em uma concepção de vulnerabilidade como condição excepcional, a ser corrigida por restrições técnicas e jurídicas.

A ética do cuidado oferece um contraponto importante a essa leitura ao compreender a vulnerabilidade como condição constitutiva da relação ética. A assimetria moral própria do cuidado exige sensibilidade ética e disposição para responder diante da alteridade, elementos que excedem categorias de risco, conformidade ou proibição. A gestão de riscos opera em registro distinto daquele da responsabilidade ética que emerge da relação entre sujeitos.

Em torno disso, os marcos regulatórios revelam uma ambivalência estrutural. Eles reconhecem a necessidade de proteger indivíduos vulneráveis e de limitar usos abusivos da técnica, ao mesmo tempo em que enquadram o cuidado como objeto de regulação externa, em vez de concebê-lo como prática relacional que exige agentes moralmente responsáveis. Reprogramar o cuidado, nesse contexto, requer evidenciar essa ambivalência e sustentar que a ética excede a mera prevenção de danos, pois envolve resposta situada, reconhecimento da assimetria e compromisso com a relação.

As iniciativas de *design* ético em inteligência artificial surgem como tentativa de responder às críticas dirigidas à neutralidade técnica dos sistemas algorítmicos. Projetos como o *Ethically Aligned Design*, desenvolvido pelo IEEE,²¹ procuram orientar o desenvolvimento tecnológico a partir de valores como bem-estar humano, responsabilidade e florescimento social. Esse movimento representa avanço relevante ao reconhecer que decisões técnicas produzem efeitos normativos e exigem orientação ética explícita.

²¹ O *Ethically Aligned Design* é um documento institucional do *Institute of Electrical and Electronics Engineers (IEEE)*, publicado em 2019 pela *IEEE Standards Association*, que apresenta diretrizes éticas para o desenvolvimento e a governança de sistemas autônomos e inteligentes, com ênfase em bem-estar humano, responsabilidade, transparência e impactos sociais.

Apesar dessa contribuição, o *design* ético permanece tensionado por um limite estrutural. Ao buscar incorporar valores éticos ao sistema, ele comumente trata o cuidado como atributo passível de inscrição técnica, traduzido em requisitos de *design*, diretrizes operacionais ou princípios de governança. O cuidado, contudo, como prática ética situada, não se estabiliza como propriedade do artefato. Ele depende de relações nas quais sujeitos assumem responsabilidades, respondem à vulnerabilidade alheia e se expõem às consequências de suas ações.

Essa tensão revela um risco recorrente nas abordagens de *design* ético. A tentativa de inscrever o cuidado no ciclo de vida do sistema pode convertê-lo em simulacro normativo, preservando sua aparência ética enquanto esvazia sua dimensão relacional. Nesse movimento, a lógica da feminilidade programada encontra nova forma de legitimação institucional, pois a disponibilidade, a responsividade e o acolhimento passam a figurar como qualidades desejáveis do sistema, agora sob a linguagem da responsabilidade ética.

O cuidado demanda mais do que alinhamento a valores abstratos. Ele pressupõe agentes capazes de responder eticamente, de revisar decisões à luz da experiência e de reconhecer a assimetria própria da relação. Nenhum protocolo de *design*, por mais sofisticado, substitui essa exigência. A incorporação normativa do cuidado encontra, assim, um limite preciso quando confrontada com a necessidade de imputabilidade moral e de presença ética.

Nesse sentido, o *design* ético cumpre função importante ao explicitar valores e orientar práticas técnicas, embora permaneça incapaz de absorver integralmente o cuidado como prática relacional. Reprogramar o cuidado implica reconhecer esse limite e evitar que a ética se reduza a camada normativa aplicada ao sistema. O cuidado pertence a um domínio, no qual a responsabilidade não pode ser delegada à técnica sem perda de inteligibilidade moral.

A articulação entre ética do cuidado e políticas de inteligência artificial conduz a uma conclusão decisiva: reprogramar o cuidado ultrapassa o alinhamento técnico de sistemas a prin-

cípios éticos formais. Os documentos normativos internacionais insistem na responsabilidade humana, na dignidade e na centralidade do sujeito precisamente porque reconhecem, ainda que de modo implícito, que determinados domínios da relação ética não admitem delegação plena à técnica.

Reprogramar o cuidado, nesse enfoque, assume um estatuto normativo acurado. Trata-se de decidir até onde a automação pode avançar sem comprometer a inteligibilidade ética das relações. Essa decisão orienta-se por um compromisso filosófico com a responsabilidade, o reconhecimento e a vulnerabilidade, que constituem o cuidado como prática ética situada, em vez de por critérios de eficiência, desempenho ou otimização funcional.

A normatividade técnica atinge seu limite quando confrontada com exigências de imputabilidade moral. Sistemas algorítmicos são capazes de organizar processos, apoiar decisões e ampliar capacidades operacionais, mas permanecem estruturalmente incapazes de responder eticamente por suas consequências. A tentativa de converter o cuidado em função técnica desloca a relação ética para o campo da conformidade regulatória, neutralizando a dimensão de resposta situada que lhe é constitutiva.

Portanto, reprogramar o cuidado implica sustentar uma distinção fundamental entre suporte técnico e relação ética. O cuidado pertence ao domínio das práticas nas quais a assimetria convoca responsabilidade e a vulnerabilidade exige resposta. A decisão normativa consiste em afirmar essa distinção e preservar espaços nos quais a ética permaneça inseparável da presença, da resposta e do compromisso entre sujeitos.

A análise desenvolvida ao longo deste item mostrou que reprogramar o cuidado na inteligência artificial de voz constitui um problema eminentemente filosófico, situado no cruzamento entre ética relacional, normatividade técnica e regimes históricos de gênero. O cuidado emerge como prática sustentada por responsabilidade, exposição à alteridade e reconhecimento da assimetria própria das relações intersubjetivas. Sua redução a atributo operacional compromete a inteligibilidade moral da relação e o desloca para o registro da eficiência técnica.

A ética do cuidado, tal como formulada por Gilligan, Tronto, Puig de la Bellacasa e Mol, oferece critérios conceituais para avaliar os limites da automação. O cuidado exige atenção às singularidades da relação, abertura à contingência e disposição para responder pelas consequências da ação. Esses elementos resistem à formalização algorítmica, pois pressupõem implicação moral e a possibilidade de ser afetado pela interpelação do outro.

A análise da feminilidade programada mostrou que a automação do cuidado ocorre por meio da reinscrição técnica de expectativas históricas de gênero. A figura da assistente permanentemente disponível, afetivamente ajustada e desprovida de resistência atualiza, no plano algorítmico, uma divisão moral do trabalho que associa cuidado, obediência e responsividade a corpos feminilizados. Sob a aparência de neutralidade funcional, esse processo naturaliza relações assimétricas e consolida disposições de comando que dispensam reconhecimento recíproco.

A articulação com documentos normativos e políticas internacionais de inteligência artificial evidenciou um esforço institucional relevante para afirmar a responsabilidade humana, a dignidade e a centralidade do sujeito. Ainda assim, tais marcos tendem a traduzir o cuidado em categorias funcionais de *design*, risco e conformidade. Embora necessária no plano regulatório, essa tradução revela limites quando confrontada com o cuidado como prática relacional que exige imputabilidade moral e resposta situada.

Nesse cenário, a reprogramação do cuidado assume caráter paradoxal, ao exigir o reconhecimento de que certas formas de relação perdem sua inteligibilidade ética sob automação plena. Reprogramar significa, nesse contexto, determinar o ponto de cessação da delegação técnica. O paradoxo reside no fato de que a reprogramação opera pela limitação, e não pela intensificação da automação. A decisão ética desloca-se, assim, do aperfeiçoamento da simulação do cuidado para a afirmação de limites normativos à sua delegação, preservando a relação ética como dimensão irreduzível da experiência humana diante da expansão da inteligência artificial.

Ao estabelecer esses limites, a filosofia cumpre sua função crítica ao lembrar que a expansão da inteligência artificial deve ser orientada menos pelo que a técnica é capaz de fazer do que pelo que a ética pode legitimamente autorizar. Reprogramar o cuidado, em sentido forte, implica afirmar que há relações que perdem seu sentido moral quando automatizadas e que a presença, a responsabilidade e o compromisso entre sujeitos constituem limites normativos que nenhuma eficiência técnica pode substituir. Reprogramar o cuidado é afirmar que a técnica pode mediar relações, mas jamais ocupar o lugar da responsabilidade ética que nasce do encontro entre sujeitos.

3.3 Princípios e critérios éticos para a inteligência artificial de voz

O critério ético emerge quando a técnica encontra um limite que não pode ser resolvido por otimização, apenas por decisão normativa.

O debate sobre a ética da inteligência artificial consolida-se a partir de abordagens voltadas à correção de efeitos indesejáveis, à mitigação de riscos e à incorporação de valores ao ciclo de desenvolvimento tecnológico. Desde a publicação da *Recommendation on the ethics of artificial intelligence* (2021) até o *AI Act europeu* (2024), princípios como transparência, justiça, explicabilidade e centralidade humana passaram a estruturar documentos normativos, *frameworks* de governança e iniciativas de *design* ético. Esses esforços ampliaram a visibilidade da dimensão ética da técnica e se inscreveram, em grande medida, em uma racionalidade de ajuste, voltada à adequação funcional dos sistemas a parâmetros previamente definidos.

Esse referencial expõe um limite estrutural das abordagens que concebem a ética como uma camada incorporável à técnica. Ao pressupor que problemas morais admitem tradução integral em requisitos operacionais, especificações de desempenho ou protocolos de conformidade, tais abordagens tendem a obscurecer dimensões estruturais de injustiça. Casos emblemáticos evidenciam essa insuficiência. O algoritmo experimental de recrutamento desenvolvido pela Amazon foi descontinuado após

demonstrar viés sistemático contra mulheres, reproduzindo padrões históricos de discriminação presentes nos dados de treinamento (Dastin, 2018).

De modo análogo, o sistema COMPAS, utilizado nos Estados Unidos para avaliação de risco criminal, revelou-se enviesado ao classificar réus negros como de maior risco de reincidência com maior frequência do que réus brancos em condições semelhantes (Angwin *et al.*, 2016). Esses exemplos indicam que, nesse registro, a ética usualmente atua como instrumento de otimização normativa, orientado por critérios de eficiência moral e gestão de impactos, falhando em enfrentar as assimetrias estruturais que atravessam os contextos sociais, nos quais a técnica é implementada.

Tal orientação mostra-se insuficiente diante de práticas tecnológicas que incidem diretamente sobre o estatuto das relações humanas, principalmente no domínio da inteligência artificial de voz, em que assistentes como Siri, Alexa e Google Assistant assumem funções de escuta contínua, mediação emocional e aconselhamento básico. Essas ações envolvem cuidado, reconhecimento, escuta e mediação afetiva. Nesses contextos, a técnica ultrapassa a condição de meio instrumental e intervém nas condições éticas que sustentam a relação entre assuntos.

A especificidade da mediação vocal exige uma base empírica capaz de sustentar o deslocamento do debate ético do plano da adequação funcional para o da decisão normativa. Em contraste com outras interfaces, a voz opera como forma de presença que ativa automaticamente esquemas sociais, afetivos e normativos. Nesse sentido, a ética da inteligência artificial de voz não pode ser pensada apenas como extensão dos *frameworks* gerais de governança, pois a fala técnica interfere diretamente na gramática moral das relações.

Os estudos de Nass e Brave demonstram de modo consistente que a interação vocal desencadeia respostas sociais inevitáveis, independentemente da consciência do usuário acerca da artificialidade da fonte. Como afirmam os autores, “interfaces de voz são intrinsecamente interfaces sociais” (Nass; Brave, 2005, p. 28). A força dessa tese reside no fato de que a fala ativa disposi-

ções relacionais pré-reflexivas, levando os usuários a interpretar a interface como participante de uma relação, e não como simples instrumento técnico.

Essa ativação automática impede que a mediação vocal seja neutralizada por esclarecimentos técnicos ou *disclaimers* funcionais. Nass e Brave observam que “ouvintes e falantes não conseguem suprimir suas respostas naturais à fala, independentemente de sua origem” (Nass; Brave, 2005, p. 28). A resposta social à voz ocorre antes de qualquer deliberação consciente, o que desloca a questão ética do campo da intenção do usuário para o da estrutura da mediação. A ética da IA de voz, nesse registro, não pode limitar-se a mitigar riscos identificáveis, uma vez que a própria forma da interação já produz efeitos normativos.

Entre esses efeitos, a atribuição de gênero ocupa posição central. Os autores demonstram que “o gênero é um dos primeiros atributos que as pessoas identificam em uma voz humana” (Nass; Brave, 2005, p. 10), e que essa identificação ocorre mesmo quando a voz é claramente sintética. O gênero vocal funciona como operador normativo primário, organizando expectativas de autoridade, cuidado, docilidade e competência. Ao aplicar às vozes artificiais os mesmos esquemas interpretativos utilizados nas interações humanas, os usuários com frequência reproduzem estereótipos historicamente associados ao feminino e ao masculino.

Essa dinâmica revela a feminilidade programada como infraestrutura relacional da automação, inscrita no próprio desenho das mediações técnicas. Nass e Brave mostram que “as pessoas aplicam às vozes de computador as mesmas regras sociais que aplicam às vozes humanas” (Nass; Brave, 2005, p. 29). A voz feminilizada passa a operar como amortecedor moral da interação, favorecendo maior tolerância ao erro, aceitação da autoridade técnica e disposição à continuidade da relação. Nesse contexto, o cuidado deixa de ser reconhecido como prática ética situada e passa a funcionar como performance técnica ajustável às expectativas do usuário.

O problema ético torna-se ainda mais evidente quando se consideram situações de falha. Os autores demonstram que “a

maneira como uma interface explica ou pede desculpas por um erro molda de forma decisiva as reações dos usuários” (Nass; Brave, 2005, p. 173). Estratégias discursivas como pedidos de desculpa ou tons empáticos alteram significativamente a atribuição de responsabilidade percebida. Quando essas estratégias são veiculadas por vozes feminilizadas, a assimetria se intensifica, visto que a expectativa de compreensão e paciência desloca a atenção do limite estrutural da técnica para a dinâmica relacional simulada.

Esses dados empíricos explicitam um limite das abordagens éticas centradas exclusivamente na regulação, na gestão de riscos ou na incorporação de valores ao *design*. No caso da inteligência artificial de voz, o problema ético ultrapassa a correção de vieses detectáveis ou da transparência de decisões algorítmicas. A mediação vocal produz efeitos morais antes mesmo da aplicação de qualquer critério de conformidade, pois ativa automaticamente disposições sociais que reorganizam a relação entre usuário, técnica e responsabilidade.

É precisamente nesse ponto que se justifica a passagem do registro regulatório para o registro decisório da ética. Quando a técnica ocupa posições tradicionalmente reservadas à relação ética – como escuta, acolhimento e mediação afetiva –, torna-se insuficiente perguntar apenas como otimizar seu funcionamento ou reduzir seus impactos negativos. A questão ética desloca-se para a decisão sobre se determinadas formas de mediação devem ocorrer. A evidência de que a voz ativa esquemas normativos irredutíveis à gestão técnica fundamenta a necessidade de critérios capazes de delimitar o alcance legítimo da automação vocal, tarefa que é desenvolvida a seguir.

À luz desse deslocamento, a reflexão ética em torno da inteligência artificial revela a coexistência de registros normativos distintos, tensionados pela especificidade da mediação vocal. Em um primeiro registro, a ética atua de forma reguladora, orientando práticas por meio de recomendações, incentivos institucionais e mecanismos de alinhamento. Nesse plano, a atenção concentra-se na compatibilização entre valores declarados e na expansão

das capacidades técnicas, preservando a lógica da governança, da otimização e da gestão de impactos.

Em outro registro, a ética assume caráter propriamente decisório e interpelativo. Ela opera pela afirmação de limites, pela interdição de determinadas práticas e pelo reconhecimento de que certas formas de mediação técnica resultam em perdas éticas irreversíveis, mesmo quando associadas a elevado grau de eficiência ou ampla aceitabilidade social. A automação de linhas de escuta psicológica de emergência ou de atendimento em situações de violência doméstica exemplifica esse dilema: a eficiência técnica não compensa a ausência de responsabilidade ética na resposta. Nesse plano, a questão ética desloca-se da regulação do uso da técnica para a decisão sobre o ponto em que sua expansão deve cessar.

A exigência de critérios normativos emerge nesse segundo registro. Diferentemente de valores orientadores ou princípios gerais, tais critérios operam como instâncias de delimitação ética e definem o alcance legítimo da automação. Eles funcionam como fronteiras normativas que distinguem mediações técnicas legítimas daquelas que comprometem a gramática moral das relações humanas. Sua força normativa reside na capacidade de indicar até onde a técnica pode avançar e em que ponto sua interação compromete o sentido ético das relações humanas.

A ética da inteligência artificial exige esse registro normativo porque a técnica passou a ocupar lugares tradicionalmente reservados à relação entre sujeitos – escuta terapêutica, mediação familiar e suporte em crises existenciais –, além de assumir, nesses contextos, práticas de resposta funcionais no lugar da resposta ética. Na ausência de critérios normativos fortes, a expansão técnica adquire caráter de evidência e mobiliza a questão ética de que se algo deve ser feito, para a interrogação técnica sobre os modos mais eficientes de realização.

No debate sobre inteligência artificial, a dignidade frequentemente aparece como princípio abstrato de proteção ou como valor a ser preservado por meio de salvaguardas técnicas. Documentos como as *Diretrizes Éticas para IA Confiável da União Europeia* (*Ethics Guidelines for Trustworthy AI*, 2019) mencionam

a dignidade, mas a operacionalizam como privacidade de dados ou não discriminação algorítmica. Tal abordagem, embora relevante no plano regulatório, mostra-se insuficiente para enfrentar os desafios éticos apresentados pela automação de práticas relacionais, especialmente aquelas vinculadas ao cuidado, à escuta e ao reconhecimento. Quando reduzida a parâmetro de conformidade, a dignidade perde sua força crítica e deixa de operar como critério normativo capaz de delimitar o alcance legítimo da mediação técnica nas relações humanas.

Compreendida filosoficamente, a dignidade designa a irredutibilidade da pessoa à lógica da funcionalização técnica, uma noção que remete tanto à tradição kantiana quanto à fenomenologia contemporânea da vulnerabilidade (Levinas, 1961). Ela afirma a impossibilidade de converter a experiência humana em parâmetros de desempenho, eficiência ou otimização. Essa noção ultrapassa a autonomia formal e a escolha racional isolada, na medida em que remete à condição relacional do sujeito, à sua vulnerabilidade constitutiva e à exigência de reconhecimento que emerge no encontro com o outro. Nessa chave de leitura, a dignidade opera como critério ético capaz de delimitar o alcance legítimo da mediação tecnológica e de preservar a inteligibilidade moral das relações humanas.

Na tradição filosófica moderna, consolidou-se progressivamente a compreensão da dignidade humana como limite moral à conversão da pessoa em objeto de uso. Esse horizonte encontra em Kant sua formulação sistemática mais rigorosa. Em sua ética, a dignidade humana designa o valor absoluto da pessoa como ser racional dotado de autonomia moral. Tal estatuto decorre da capacidade de agir segundo princípios suscetíveis de valer como leis universais, razão pela qual o ser racional existe como fim em si mesmo, e não como simples meio para o uso arbitrário desta ou daquela vontade.

A dignidade distingue-se, assim, de toda forma de valor relativo associado à utilidade, ao desempenho funcional ou à substituíbilidade. Enquanto as coisas podem ser trocadas por equivalentes, a pessoa se afirma pela autolegislação moral, que a constitui como sujeito de imputação ética e fonte de normativida-

de prática. É nesse sentido que Kant afirma, na *Fundamentação da Metafísica dos Costumes*, que “a natureza racional distingue-se das restantes, por se pôr a si mesma um fim” (BA 82), razão pela qual os seres dotados de racionalidade “chamam-se pessoas” (BA 65). Como fins em si por sua própria natureza, são portadores de um valor absoluto e de uma dignidade inviolável, o que exclui sua redução à condição de instrumento para a realização de fins subjetivos.

A luz desse horizonte ético, o conceito adquire especial fecundidade quando transposto para o campo da inteligência artificial, particularmente da IA de voz conversacional. Os sistemas técnicos podem possuir elevado valor funcional e desempenhar papéis relevantes em processos sociais complexos, embora permaneçam desprovidos de dignidade, uma vez que carecem de autonomia moral e de capacidade de autolegislação prática. A questão central desloca-se da sofisticação linguística de modelos conversacionais avançados, como *GPT* e *Claude*, para a reorganização ética que se instaura quando suas respostas passam a substituir a interlocução humana em contextos de vulnerabilidade.

No domínio das tecnologias digitais e da inteligência artificial, tal limite assume novas configurações específicas à mediação vocal. A instrumentalização desloca-se para a conversão de relações em interações funcionais, para a substituição da resposta ética por performances técnicas e para a naturalização de assimetrias mediadas por sistemas algorítmicos. A arquitetura dos assistentes de voz, com suas respostas sempre disponíveis, tom acolhedor e ausência estrutural de resistência ou interpelação, institui um regime de interação, no qual a pessoa humana é posicionada como usuária terminal, não como interlocutora ética.

Um estudo do *Alan Turing Institute* (2023) indica que parte significativa do público tende a atribuir elevada credibilidade às respostas produzidas por sistemas de inteligência artificial, inclusive em contextos sensíveis. Essa disposição revela a naturalização de assimetrias epistêmicas nas interações mediadas por assistentes de voz, nas quais a autoridade técnica da resposta automatizada normalmente é aceita sem mediação crítica. Nesse

cenário, a dignidade torna-se vulnerável quando a pessoa é posicionada como destinatária passiva de serviços automatizados que encenam reconhecimento desprovido de responsabilidade ética, compromisso relacional e possibilidade efetiva de resposta.

A automação do cuidado configura um ponto particularmente sensível nesse processo, sobretudo, no campo da inteligência artificial de voz. A transposição de práticas fundadas na atenção à vulnerabilidade e na responsabilidade relacional para arquiteturas técnicas orientadas pela eficiência da interação favorece a redefinição da dignidade em termos de satisfação do usuário, bem-estar percebido ou fluidez comunicativa. Plataformas como *Woebot*²², voltada ao apoio terapêutico automatizado, *Wysa*²³, orientada ao bem-estar emocional por meio de práticas estruturadas, e *Replika*²⁴, concebida como companheiro emocional, exemplificam essa tendência ao operarem com métricas de sucesso associadas ao engajamento contínuo, ao tempo de retenção

²² *Woebot* é um *chatbot* desenvolvido pela empresa *Woebot Health*, concebido para oferecer apoio em saúde mental por meio de interações automatizadas baseadas principalmente em técnicas da Terapia Cognitivo-Comportamental (TCC). A plataforma se apresenta como um recurso complementar, voltado à psicoeducação, ao acompanhamento do humor e à proposição de exercícios estruturados. Embora estudos indiquem efeitos positivos de curto prazo na redução de sintomas autorrelatados, o *Woebot* não assume função clínica plena nem responsabilidade terapêutica no sentido profissional. A relação estabelecida permanece assimétrica: o sistema orienta, responde e simula cuidado, mas não sustenta compromisso ético, responsabilidade moral ou reconhecimento recíproco próprios da relação terapêutica humana.

²³ *Wysa* é um *chatbot* voltado ao bem-estar emocional que combina interações automatizadas com práticas inspiradas em abordagens psicológicas estruturadas, como TCC, *mindfulness* e exercícios de regulação emocional. A plataforma se posiciona como ferramenta de apoio e autocuidado, podendo incluir, em alguns contextos, a mediação de profissionais humanos. Assim como outros agentes conversacionais desse tipo, o *Wysa* opera por meio de *scripts* e respostas automatizadas, oferecendo suporte sem assumir responsabilidade clínica plena. A relação estabelecida permanece funcional e instrumental, ainda que revestida de linguagem empática, o que evidencia a distância entre a simulação de cuidado e a experiência ética do reconhecimento interpessoal.

²⁴ *Replika* é um *chatbot* de companhia emocional criado pela empresa *Luka, Inc.*, projetado para manter conversas contínuas e personalizadas com o usuário. Diferentemente de ferramentas com finalidade terapêutica explícita, o *Replika* se apresenta como “amigo virtual” ou “companheiro emocional”, incentivando vínculos afetivos subjetivos por meio de respostas empáticas simuladas. A plataforma explora a continuidade da interação, a personalização da *persona* e a adaptação emocional do agente artificial. Do ponto de vista ético, a relação instaurada não se funda em reciprocidade, responsabilidade ou reconhecimento efetivo, mas em uma performance de empatia orientada pela lógica da interação persistente, o que levanta questões sobre dependência afetiva, assimetria relacional e instrumentalização da vulnerabilidade emocional.

e à satisfação declarada, em lugar da qualificação ética da relação estabelecida.

Nesse deslocamento, a dignidade assume a forma de efeito comunicativo, afastando-se de sua função normativa como condição da relação ética. O sujeito passa a figurar como perfil de uso, padrão comportamental ou fonte de dados, perdendo o estatuto de interlocutor ético e sendo reinscrito em uma lógica de gestão técnica da experiência relacional.

A dignidade específica, como exigência normativa fundada na irreduzibilidade da relação ética, resiste a esse posicionamento. Ela afirma a indisponibilidade de certas dimensões da experiência humana à automação plena, na medida em que envolve interpelação, exposição à alteridade e resposta responsável. Relações de cuidado, acolhimento e escuta dependem da posição ética daquela que responde, posição que exige exposição ao risco moral, possibilidade de erro e responsabilidade por consequências, elementos estruturalmente ausentes na mediação algorítmica.

Nesse aspecto, a dignidade da pessoa humana estabelece-se como operador normativo de dupla incidência. No sentido delimitador, ela define práticas cuja automação compromete a responsabilidade ética e o reconhecimento da assimetria relacional. Em sentido orientador, ela direciona o desenvolvimento de tecnologias que preservam a centralidade da pessoa como sujeito de relação, como sistemas híbridos humanos-IA que mantêm a ética da supervisão humana em momentos críticos.

Reposicionar a dignidade da pessoa humana como eixo da reflexão ética sobre a inteligência artificial de voz desloca o debate do plano da conformidade normativa para a decisão filosófica. Preservar a dignidade ultrapassa a lógica da prevenção de danos e da mitigação de riscos e afirma a existência de limites éticos estruturais à automação. Tais limites decorrem da própria natureza das relações compatíveis com a experiência ética: toda relação implica risco, assimetria e responsabilidade irreduzíveis à lógica algorítmica de otimização probabilística.

A primeira tese normativa sustenta que a mediação técnica possui relevância ética própria e pode incidir sobre a dignidade da pessoa humana. Essa formulação implica um deslocamen-

to da compreensão segundo a qual a mediação tecnológica se apresentaria como eticamente neutra, tornando-se problemática apenas em razão de usos inadequados ou de falhas contingentes de implementação. Em contextos nos quais as tecnologias reconfiguram o próprio sentido das relações humanas, como ocorre com a IA conversacional que ocupa posições de escuta contínua e cuidado básico, a mediação deixa de operar como simples meio instrumental e passa a incidir diretamente sobre as condições éticas que estruturam a relação entre sujeitos.

Ao definir ritmos de resposta instantânea, estabelecer posições assimétricas programadas e produzir expectativas normativas sobre comportamento, disponibilidade e autoridade, a mediação técnica opera como instância estruturante das interações. Quando sistemas de inteligência artificial de voz passam a ocupar lugares tradicionalmente vinculados à relação ética, como escuta em crises emocionais, acolhimento em situações de vulnerabilidade ou mediação em conflitos relacionais, a mediação assume caráter constitutivo. Nesse ponto, a questão ética desloca-se da avaliação da qualidade da mediação (precisão, fluidez, satisfação) para a interrogação de sua legitimidade estrutural.

A dignidade da pessoa humana entra em risco quando a mediação técnica passa a ocupar o lugar da relação ética por meio de performances funcionais que encenam o reconhecimento sem assumir responsabilidade. Nessa configuração, a pessoa deixa de figurar como interlocutora ética capaz de convocar resposta responsável e passa a ser posicionada como destinatária de respostas otimizadas, ajustadas a padrões de uso e expectativas previsíveis. A mediação transforma, assim, a relação em interação transacional e o encontro em fluxo operacional, enfraquecendo a exigência de reconhecimento recíproco que sustenta a dignidade como eixo ético da relação.

Essa transformação pode instaurar-se mesmo na ausência de intenção explícita de manipulação ou exploração, funcionando de modo silencioso por meio da naturalização de interfaces que respondem de forma contínua, acolhem sem interpelar e permanecem disponíveis sem resistência. Relatórios e pesquisas da

*Mozilla Foundation*²⁵ sobre privacidade e *design* de tecnologias digitais têm chamado atenção para os efeitos dessa responsabilidade permanente na experiência do usuário, indicando como a interação prolongada com sistemas de voz pode favorecer processos de antropomorfização, confiança excessiva e projeção de expectativas relacionais sobre dispositivos técnicos que não sustentam reconhecimento ou responsabilidade ética.

O argumento de que a mediação técnica pode comprometer a dignidade da pessoa humana exige considerações práticas cuja inteligibilidade moral depende da presença de sujeitos capazes de responder eticamente. Cuidado em situações de luto, escuta em momentos de crise suicida e acolhimento em situações de violência exigem a posição ética de quem responde, posicionamento que envolve exposição pessoal, risco moral e responsabilidade irredutível.

O primeiro princípio normativo afirma que a dignidade da pessoa humana exerce a função de critério de discernimento ético das mediações técnicas. Sempre que a técnica sustenta a relação e amplia possibilidades de ação mantendo a responsabilidade ética no jogo, a dignidade permanece preservada. Sempre que a técnica assume o lugar da relação e se converte em desempenho funcional, como ocorre quando assistentes de voz respondem autonomamente a pedidos de aconselhamento em situações de vulnerabilidade, a dignidade é comprometida, ainda que essa mediação se apresente como eficiente ou socialmente aceitável.

Ao assumir esse princípio, a ética da inteligência artificial recupera sua função decisiva. A reflexão ética desloca-se da mera regulação da técnica de mediação para a interrogação sobre a

²⁵ Fundação Mozilla é uma organização global sem fins lucrativos dedicada à promoção de uma internet aberta, acessível e centrada nas pessoas, orientada por princípios de transparência, privacidade, inclusão e justiça tecnológica. Por meio de pesquisas, iniciativas de dados abertos e projetos comunitários – como o *Common Voice*, voltado à criação colaborativa de bases de dados de voz –, a *Mozilla Foundation* atua na análise crítica do impacto social das tecnologias digitais. A instituição publica relatórios e estudos sobre integridade da internet, governança de dados e desenvolvimento responsável da inteligência artificial, com o objetivo de influenciar políticas públicas, práticas de *design* e processos de inovação tecnológica. No contexto de sistemas baseados em IA, incluindo assistentes de voz, seu trabalho enfatiza a defesa da agência do usuário, a proteção da privacidade, a transparência dos sistemas e a necessidade de responsabilização em ambientes digitais crescentemente automatizados.

legitimidade de sua ocorrência em domínios especificamente humanos.

O segundo princípio normativo afirma a inseparabilidade entre cuidado e dignidade. Essa afirmação ultrapassa uma atribuição valorativa abstrata e expressa uma relação estrutural entre as formas de reconhecimento ético entre sujeitos e as práticas pelas quais se respondem à vulnerabilidade do outro. O cuidado constitui uma mediação privilegiada da dignidade ao tornar visível a condição relacional do sujeito e a exigência de responsabilidade que dela emerge, situada para além da absorção por protocolos técnicos.

A ética do cuidado permite compreender a dignidade como realidade relacional, produzida em vínculos nos quais a vulnerabilidade é reconhecida como convocação ética e assumida como tal. Cuidar implica oferecer resposta a alguém cuja situação interpela, desloca e compromete quem responde. Na fenomenologia de Levinas, essa estrutura manifesta-se primeiramente na relação com o *Rosto* do outro, cuja alteridade resiste à objetivação e impõe uma exigência ética imediata, anterior a qualquer mediação cognitiva ou cálculo instrumental. A noção de terceiro surge, nesse horizonte, como desdobramento dessa responsabilidade originária, introduzindo a dimensão da justiça e da comparação entre múltiplos outros, sem dissolver a assimetria ética fundamental da relação face a face. Essa compreensão distingue o cuidado de formas de prestação funcional orientadas exclusivamente por critérios de eficiência, satisfação ou desempenho, ao afirmar a primazia da responsabilidade ética sobre a lógica da gestão técnica das relações.

A dissociação entre cuidado e dignidade favorece a reinterpretação do cuidado como serviço, assistência ou resposta técnica adequada. Nesse deslocamento, a vulnerabilidade tende a ser tratada como condição a ser gerida, e não como exigência ética que convoca responsabilidade. O caso do aplicativo *Ellie*, um avatar virtual desenvolvido no âmbito do projeto *SimSensei*, conduzido pelo *Institute for Creative Technologies (ICT/USC)* em colaboração com instituições militares dos Estados Unidos, ilustra essa lógica (Devault *et al.*, 2014). Utilizado essencialmente no rastreamento

de sintomas de transtorno de estresse pós-traumático em veteranos militares, muitos deles com experiência direta em conflitos armados, o sistema demonstra eficácia na identificação de indicadores clínicos iniciais. Ainda assim, ele não sustenta a relação terapêutica propriamente dita, pois carece da reciprocidade ética que caracteriza o cuidado como prática relacional responsável. A interação assume, assim, a forma de um fluxo operacional, no qual a assimetria é absorvida por protocolos técnicos e a responsabilidade se redistribui em procedimentos automatizados. Nesse ponto, a dignidade se enfraquece, uma vez que a pessoa é integrada a um regime de atendimento técnico e deixa de figurar como interlocutora ética da relação.

A automação intensificada do cuidado configura um risco ao substituir a resposta ética por desempenhos técnicos. Sistemas de inteligência artificial de voz reproduzem formas discursivas associadas à atenção, à empatia e ao acolhimento e permanecem estruturalmente incapazes de manifestar a posição ética que o cuidado exige. Estudos em interação humano-computador e em saúde digital têm indicado que usuários relatam maior facilidade de expressão emocional em interações com sistemas automatizados, associando essa abertura à percepção de disponibilidade constante e ausência de julgamento. Essa confiança, contudo, apoia-se em uma assimetria ética estrutural: a ausência de imputabilidade moral, de exposição à alteridade e de responsabilidade pelo outro impede que a resposta automática constitua reconhecimento ético propriamente dito. Quando o cuidado é automatizado, a dignidade deixa de operar como relação ética e passa a assumir a forma de simulação funcional.

Essa dissociação incide diretamente sobre a compreensão da dignidade da pessoa humana. A dignidade passa a ser interpretada como efeito subjetivo da interação, avaliada em termos de conforto, fluidez ou satisfação percebida, em lugar de funcionar como limite ético da mediação técnica. Essa redução é particularmente perigosa no contexto brasileiro, onde os serviços públicos de saúde mental já sofrem subfinanciamento crônico: a tentativa de substituir psicólogos por *chatbots* econômicos ameaça aprofundar desigualdades éticas sob o pretexto da “escalabilidade”.

Essa posição transfere a dignidade do campo da responsabilidade ética para a experiência do usuário e se torna compatível com regimes de interação nos quais a relação ética cede lugar a respostas otimizadas.

A inseparabilidade entre cuidado e dignidade exige reconhecer práticas cuja substituição técnica implica perda ética. O cuidado supõe presença, implicação e capacidade de resposta responsável, elementos que sustentam a dignidade como realidade relacional, e não como atributo abstrato do sujeito. Quando essas condições deixam de poder ser efetivamente assumidas, a mediação técnica ultrapassa um limiar normativo e compromete a inteligibilidade moral da prática. É nesse sentido que a delegação de funções de escuta e acolhimento a assistentes de voz, como ocorreria hipoteticamente na atuação em serviços análogos ao Centro de Valorização da Vida (CVV), revela um limite ético estrutural: a ausência de responsabilidade, de exposição à alteridade e de compromisso relacional impede que a resposta automatizada se constitua como cuidado ético propriamente dito.

Práticas de cuidado humano são constitutivamente relacionais e envolvem presença, implicação e capacidade de resposta responsável, elementos que sustentam a dignidade como relação ética, para além de sua compreensão como atributo abstrato. A técnica pode apoiar, mediar e ampliar tais práticas na medida em que preserva esses vínculos, sem substituí-los ou assumir indevidamente funções que exigem responsabilidade humana. Ferramentas de triagem baseadas em inteligência artificial, como as da plataforma Ada Health²⁶, ilustram essa incorporação técnica no campo da saúde ao oferecerem apoio inicial à avaliação de sintomas e à orientação de usuários, funcionando como suporte à decisão clínica. Estudos comparativos indicam que esses sistemas podem alcançar níveis relevantes de concordância com decisões médicas humanas em determinadas métricas diagnósticas e de triagem, sem, contudo, sustentar a responsabilidade

²⁶ Ada Health é um aplicativo de triagem clínica que usa inteligência artificial para avaliar sintomas relatados pelo usuário e fornecer possíveis diagnósticos e recomendações de urgência. Estudos científicos comparativos investigam sua acurácia diagnóstica em relação a decisões médicas humanas, ressaltando seu papel como apoio técnico à decisão clínica, e não como substituto de profissionais de saúde (Knitza, 2024).

ética relacional que caracteriza o cuidado clínico aprofundado. Quando os elementos de responsabilidade e resposta humana deixam de poder ser efetivamente assumidos, a mediação técnica ultrapassa um limiar normativo e compromete a inteligibilidade moral da prática de cuidado em contextos complexos.

Ao explicitar a inseparabilidade entre cuidado e dignidade, esse princípio normativo reforça a necessidade de critérios capazes de distinguir mediações técnicas que sustentam a relação daquelas que a substituem. Tal distinção fundamenta a formulação de critérios decisórios nos quais a dignidade da pessoa humana opera como eixo normativo para orientar o *design*, o uso e os limites da inteligência artificial de voz.

Os princípios aqui formulados amparam critérios éticos de avaliação e decisão. Eles orientam normativamente os limites da intervenção técnica, seu âmbito legítimo, seus modos justificáveis e, sobretudo, os pontos de cessação exigidos para preservar a inteligibilidade moral das relações humanas.

Esses critérios não se confundem com boas práticas, recomendações genéricas ou listas de verificação técnica. Eles operam como instâncias de discernimento ético destinadas a delimitar o alcance legítimo da automação e a identificar situações nas quais a mediação técnica compromete a dignidade da pessoa humana e a inteligibilidade da relação ética. Essa discussão adquire especial relevância no contexto brasileiro, no qual iniciativas de digitalização de serviços públicos – como o SUS Digital – ampliam o uso de interfaces automatizadas para orientação, triagem informacional e acesso a serviços, exigindo atenção ética rigorosa para que tais mediações não substituam nem fragilizem práticas de cuidado fundadas na relação humana.

Os critérios a seguir derivam da articulação entre dignidade, cuidado e responsabilidade. Sua função consiste em orientar decisões acerca de onde a técnica pode intervir, de que modo essa intervenção se justifica e, principalmente, em que pontos ela deve cessar para preservar a inteligibilidade ética das relações.

1. Critério da irredutibilidade da dignidade:

A dignidade da pessoa humana é irredutível a perfil de uso, padrão de interação ou posição passiva diante de respostas auto-

matizadas. Ela designa a impossibilidade de converter o sujeito em meio técnico ou recurso operacional. Quando o *design* de sistemas de voz trata uma pessoa exclusivamente como usuária a ser atendida, satisfeita ou retida, como ocorre em plataformas de *streaming* que utilizam IA vocal para maximizar retenção de assinantes, instaura-se um deslocamento ético que compromete seu reconhecimento como interlocutora da relação. Esse critério orienta a concepção de sistemas de inteligência artificial de voz para preservar a relação ética como referência central do *design*, especialmente em contextos atravessados por vulnerabilidade, sofrimento, cuidado ou decisões morais relevantes.

2. Critério da imputabilidade moral:

A imputabilidade moral constitui condição necessária para práticas que envolvem responsabilidade ética. Onde não há possibilidade de responder moralmente por ações e consequências, o cuidado não se constitui como prática ética. A inteligência artificial pode executar tarefas rotineiras, organizar informações e apoiar processos decisórios, permanecendo estruturalmente incapaz de assumir responsabilidade diante de outro sujeito, limitação que se torna crítica em situações como aconselhamento jurídico inicial ou suporte psicológico básico. Esse critério orienta a delimitação da automação em interações que envolvem aconselhamento sensível, acolhimento emocional, mediação de conflitos ou resposta à vulnerabilidade humana.

3. Critério de interrupção de simulações:

É necessário interromper a simulação de relações éticas em sistemas técnicos que operam sem reconhecimento, interpelação e resposta responsável. A encenação de empatia, cuidado ou acolhimento por meio de performances discursivas produz a aparência de relação ética sem a possibilidade efetiva de assumi-la. Um caso paradigmático é o uso de IA vocal em *call centers* de telecomunicações, em que respostas padronizadas de “compreensão” geram frustração quando falham diante de situações genuinamente complexas. Esse critério orienta o *design* de interfaces de voz para explicitar os limites ontológicos e éticos da técnica, evitando a produção de falsas expectativas relacionais.

4. Critério de explicitação crítica das assimetrias:

É necessário apontar criticamente as assimetrias mediadas pela técnica. A feminilidade programada, analisada ao longo deste livro, exemplifica como escolhas de *design* reinscrevem expectativas históricas de conformidade, disponibilidade e docilidade sob a aparência de neutralidade funcional. Quando a assimetria se torna invisível, como nas vozes submissas de assistentes como Siri e Alexa, projetadas para máxima agradabilidade, ela deixa de ser tematizada eticamente e passa a operar como pressuposto tácito da interação. Estudos semióticos sobre prosódia de assistentes indicam que tons ascendentes e pausas deferentes ativam esquemas cognitivos de subordinação, mesmo em usuários conscientes da artificialidade da interface. Esse critério orienta o desenvolvimento de sistemas de inteligência artificial de voz de modo a evitar a reprodução acrítica de *personas*, vozes e estilos comunicativos associados a disposições de comando e obediência vinculadas a marcadores de gênero, classe ou posição social.

5. Critério da clareza ontológica:

É fundamental explicitar o estatuto ontológico dos sistemas de inteligência artificial. A transparência, nesse contexto, ultrapassa a explicabilidade de decisões algorítmicas e envolve a clareza quanto às capacidades e aos limites constitutivos da técnica. O sistema deve tornar evidente sua condição de mediação técnica, desprovida de consciência, imputabilidade moral e capacidade de reconhecimento ético. Interfaces que omitem essa distinção favorecem o chamado *efeito Eliza*, projeção antropomórfica documentada desde os anos 1960, no qual usuários atribuem intenções ou interesses reais a respostas mecânicas. A ausência de demarcações ontológicas evidentes em assistentes de voz com interação naturalizada, como o *Google Duplex*²⁷,

²⁷ Google Duplex é um sistema de inteligência artificial conversacional desenvolvido pelo Google, integrado ao Google Assistant, capaz de realizar chamadas telefônicas automáticas em nome do usuário para executar tarefas específicas, como agendamento de compromissos ou reservas de serviços. O sistema foi projetado para interagir por voz de forma naturalizada, utilizando entonação, pausas e estratégias conversacionais semelhantes às humanas. Seu uso suscita debates éticos relevantes acerca de transparência, consentimento, imputabilidade moral e da substituição de mediações humanas por performances técnicas em contextos sociais cotidianos.

intensifica essa confusão. Esse critério orienta a distinção entre mediação técnica e relação humana.

6. Critério da responsabilidade humana direta:

Determinadas práticas permanecem sob responsabilidade humana direta, ainda que a automação se apresente como tecnicamente viável. O cuidado profundo, a escuta ética e o acolhimento de vulnerabilidades extremas pertencem a esse domínio. A técnica pode apoiar essas ações ao oferecer suporte logístico, mediação inicial e ampliação de capacidades, sem ocupar o lugar da relação ética. O protocolo do CVV brasileiro²⁸, que utiliza IA apenas para triagem inicial e transferência imediata para voluntários humanos, exemplifica essa abordagem ética. Esse critério afirma a preservação da relação humana como exigência ética positiva, situada no interior de uma reflexão crítica sobre a mediação técnica.

7. Critério de decisão ética explícita:

As escolhas relativas ao uso de inteligência artificial de voz devem ser acompanhadas de decisão ética explícita, fundamentada em critérios normativos e orientada pela dignidade da pessoa humana. Sempre que um sistema passa a ocupar o lugar da relação, como na automação de serviços psicossociais, torna-se necessária a explicação normativa dessa escolha. Essa justificativa deve expressar razões éticas específicas, limites claros de intervenção e mecanismos de revisão e supervisão humana. A ausência dessa deliberação caracteriza o que filósofos da técnica denominam *decisionismo tecnológico*: a naturalização da automa-

²⁸ O Centro de Valorização da Vida (CVV) é uma organização brasileira sem fins lucrativos, fundada em 1962, dedicada à prevenção do suicídio e à promoção da saúde emocional. O CVV oferece atendimento gratuito e sigiloso por meio do telefone 188, chat online e e-mail, com funcionamento ininterrupto em todo o território nacional. A prática central da instituição é a escuta empática e não julgadora, realizada exclusivamente por voluntários humanos capacitados, orientada por princípios de acolhimento, confidencialidade, respeito à singularidade da pessoa e ausência de aconselhamento diretivo. Tecnologias digitais são utilizadas como meios de acesso, organização e encaminhamento do atendimento, sem substituir a relação humana de escuta e cuidado. Nesse sentido, o modelo do CVV afirma a preservação da presença humana como condição ética indispensável em situações de vulnerabilidade extrema, estabelecendo um limite normativo claro à automação de práticas que exigem responsabilidade relacional direta.

ção como solução padrão, sem interrogativo moral. Esse critério recoloca a ética no centro do processo decisório.

A crítica crescente dirigida aos assistentes virtuais ajustados por padrões feminilizados levou ao surgimento de iniciativas que buscam romper com a naturalização da feminilidade programada. Entre essas propostas, a criação de vozes neutras, não binárias ou híbridas passou a ocupar lugar de destaque no debate público e acadêmico, apresentando-se como alternativa ética à associação recorrente entre mediação técnica, docilidade e trabalho emocional feminizado.

A iniciativa *Q – The Genderless Voice*, lançada em 2019²⁹ durante o *SXSW por Virtue* em parceria com *Copenhagen Pride* e pesquisadores de gênero e tecnologia, tornou-se um exemplo paradigmático de *design* crítico que questiona a normalização de vozes assistivas associadas a identidades binárias. O desenvolvimento do *Q* envolveu a combinação e a modulação de vozes de pessoas situadas fora das categorias tradicionais de gênero, além de sua avaliação por milhares de ouvintes para aferir o grau de neutralidade percebida. Esse processo evidencia a forma vocal como escolha projetual, influenciada por convenções culturais sobre gênero e expectativas sociais, inscrita nas decisões de *design* da tecnologia.

A relevância dessas iniciativas fica mais clara quando examinada à luz dos critérios normativos aqui formulados. A introdução de vozes neutras configura uma resposta parcial ao problema ético da feminilidade programada, na medida em que desloca a associação automática entre cuidado, docilidade e feminilidade. Esse deslocamento, contudo, só adquire consistência ética quando acompanhado de uma revisão mais profunda da arquitetura moral da interação, capaz de reconfigurar expectativas de disponibilidade, obediência e absorção afetiva inscritas no próprio desenho da mediação técnica.

Nessa perspectiva, o foco da análise desloca-se da identidade vocal para a estrutura ética da relação. O que se torna decisivo

²⁹ *Q – The Genderless Voice* é uma iniciativa lançada em 2019 que propõe uma voz para assistentes de inteligência artificial projetada para situar-se fora das categorias tradicionais de gênero, desenvolvida a partir da colaboração entre linguistas, tecnólogos e participantes não binários (Q, 2019).

é examinar como se fala, a quem se responde e sob quais expectativas normativas a interação é organizada. Uma voz neutra programada para atender sem resistência, absorver agressões verbais e simular cuidado permanente, como a voz *Q* configurada para *call centers*, tende a reproduzir a mesma lógica de docilidade que caracteriza as assistentes feminilizadas, apenas deslocada para um registro aparentemente desgenerificado. O cuidado permanece automatizado, a relação ética é sustentada por desempenho funcional e a dignidade da pessoa humana continua subordinada a um regime de interação orientado pela eficiência e pela previsibilidade. A crítica às vozes neutras, portanto, não invalida essas iniciativas técnicas, mas evidencia seus limites éticos quando tomadas como solução isolada para problemas que dizem respeito à própria legitimidade da mediação técnica em domínios sensíveis.

À luz dos critérios normativos formulados, a discussão conduz a uma inflexão fundamental na forma de compreender a inteligência artificial de voz. Reprogramá-la envolve menos a transformação isolada de sistemas técnicos e mais a reconfiguração da relação que a sociedade estabelece com a técnica. Ao indicar que os impasses associados à automação do cuidado, à simulação de relações e à neutralização da dignidade exigem decisões estruturais capazes de delimitar a mediação técnica, essa distinção adquire alcance normativo preciso, superando intervenções pontuais de *design*.

Reprogramar a técnica, em sentido estrito, diz respeito à modificação de arquiteturas, interfaces e protocolos operacionais. Essa tarefa permanece necessária para reduzir danos imediatos, explicitar limites funcionais e evitar o reforço de desigualdades históricas. Quando tomada de forma isolada, contudo, ela preserva a expectativa de que a técnica possa substituir relações éticas desde que adequadamente calibrada, uma ilusão que mantém a automação como horizonte privilegiado de solução.

Reprogramar a relação, por sua vez, implica interrogar as condições sob as quais a técnica é convocada a mediar práticas humanas. Essa perspectiva evidencia que a expansão da inteligência artificial resulta também de escolhas culturais, or-

ganizacionais e políticas que definem o que se espera da técnica e o que se está disposto a delegar a ela. Reprogramar a relação envolve redefinir expectativas de disponibilidade contínua, resposta imediata e substituição automática, recolocando a responsabilidade humana no centro das interações mediadas por sistemas algorítmicos, especialmente em países como o Brasil, onde a desigualdade digital agrava os riscos éticos da automação desregulada.

Nesse horizonte, o futuro ético da inteligência artificial de voz depende da capacidade de sustentar decisões que tomem a dignidade da pessoa humana como critério normativo orientador. Isso implica reconhecer que certos domínios da experiência humana, como o cuidado profundo, a escuta ética diante do sofrimento existencial e o acolhimento da vulnerabilidade radical, requerem presença humana insubstituível e resistem à delegação plena à técnica. A mediação tecnológica pode apoiar essas práticas éticas ao ampliar condições de acesso universal, organização eficiente de informações e continuidade assistencial, desde que não ocupe o lugar da relação nem produza a aparência ilusória de reconhecimento em contextos nos quais não há possibilidade de resposta moral autêntica.

Os caminhos futuros exigem um deslocamento do imaginário tecnológico dominante que permeia tanto o Vale do Silício quanto as políticas públicas emergentes no Brasil. A inteligência artificial perde o estatuto de solução universal para carências relacionais – visão que justificou investimentos bilionários em “empatia escalável” – e passa a ser compreendida como mediação limitada, situada e normativamente orientada. Essa inflexão exige uma revisão profunda de modelos de negócios predatórios, de práticas de *design* orientadas por métricas de eficiência e de políticas públicas que hoje favorecem a substituição da relação humana por interfaces automatizadas em nome da eficiência operacional, da escalabilidade e da redução de custos trabalhistas.

A educação ética e técnica assume papel central nesse processo transformador. Desenvolvedores, *designers*, gestores públicos e usuários finais precisam ser formados para reconhecer os limites morais intransponíveis da automação e para sustentar

decisões responsáveis que preservem a relação ética diante da expansão da mediação algorítmica. No contexto brasileiro, essa exigência demanda uma articulação institucional em três níveis complementares: formação universitária que integre a ética aplicada da inteligência artificial desde os cursos de computação; marcos regulatórios – como o PL 2.338/2023 – orientados por critérios substantivos de dignidade humana; e programas de capacitação profissional que assegurem, inclusive, nas áreas sensíveis, como a saúde mental, o uso da técnica sob responsabilidade ética expressamente assumida.

Reprogramar a relação envolve instituir condições institucionais e culturais que asseguram a assunção explícita da responsabilidade por sujeitos e organizações envolvidas nas decisões, na implementação e no uso dessas ferramentas. Isso inclui a criação de comitês éticos de automação em hospitais, escolas e serviços públicos; o desenvolvimento de protocolos híbridos humano-IA com critérios claros de escalonamento; e a consolidação de uma cultura organizacional que afirme a presença humana como valor ético constitutivo das práticas institucionais.

Ao examinar esses caminhos, a reflexão desenvolvida propõe afastar tanto o retorno nostálgico a formas pré-tecnológicas de relação quanto a exclusão radical da inteligência artificial como instrumento civilizatório. O que se afirma é uma reorganização filosófica do lugar da técnica na vida social, orientada pela preservação inegociável da dignidade da pessoa humana e pela centralidade irreduzível do cuidado como prática ética constitutiva do humano. Reprogramar a técnica permanece tarefa necessária; reprogramar a relação constitui exigência filosófica mais profunda, pois define o horizonte normativo, no qual a técnica pode operar legitimamente sem comprometer o sentido ético fundamental das relações humanas.

O percurso desenvolvido neste capítulo permite afirmar que a ética da inteligência artificial de voz se configura como questão de decisão normativa estrutural, fundada na dignidade da pessoa humana como critério irreduzível para delimitar o alcance legítimo da automação relacional. A dignidade assume, nesse contexto, função ética concreta e operativa, orientando limites

precisos à mediação técnica e preservando a relação humana diante da substituição por performances funcionais desprovidas de responsabilidade moral.

A articulação entre dignidade e cuidado esclarece o limite ético que atravessa toda automação das relações afetivas. O cuidado, compreendido como prática ética situada no tempo e no espaço interpessoal, expressa a dignidade por meio do reconhecimento concreto da vulnerabilidade alheia e da assunção pessoal de responsabilidade diante dela. Quando convertido em serviço automatizado otimizado por algoritmos de aprendizado profundo, o cuidado tende a deslocar a dignidade para o plano subjetivo da satisfação do usuário, esvaziando a imputabilidade moral que sustenta propriamente a relação ética.

Os critérios normativos formulados ao longo deste capítulo respondem a esse diagnóstico estrutural ao deslocar o debate ético do plano meramente compatibilizatório para o da delimitação filosófica. Eles afirmam, com rigor, que a possibilidade técnica jamais equivale à legitimidade moral e que toda automação relacional exige decisão ética explícita, fundada em princípios irreduzíveis. Reprogramar a inteligência artificial de voz implica, portanto, reconhecer práticas cuja inteligibilidade ética depende estruturalmente da presença de sujeitos humanos capazes de responder, implicar-se existencialmente e assumir consequências morais diante de outros sujeitos vulneráveis.

Nesse horizonte filosófico e normativo, a dignidade da pessoa humana assume a função de categoria última e inapelável para avaliar a legitimidade ontológica da mediação algorítmica. Ela orienta decisões concretas sobre os domínios nos quais a técnica pode apoiar e ampliar as relações humanas, como triagem inicial, suporte logístico e organização informacional, e aqueles nos quais sua intervenção direta compromete a própria gramática ética da relação. Essa decisão normativa expressa um compromisso civilizatório com a preservação da ética como dimensão irreduzível da vida social em uma sociedade crescentemente mediada por sistemas inteligentes.

A técnica encontra legitimidade ética plena quando sustenta e amplifica a relação humana; perde-a de forma decisiva

quando converte a relação em simulação funcional desprovida de interpelação recíproca. Ao reposicionar a ética da inteligência artificial de voz no registro decisório, este capítulo ultrapassa a crítica diagnóstica e institui caminhos normativos concretos para uma técnica verdadeiramente humana, orientada pela dignidade como limite irrenunciável da automação.

Reprogramar a inteligência artificial de voz é, em última instância, decidir onde a técnica deve cessar para que a dignidade humana permaneça como condição viva da relação ética. Em uma sociedade cada vez mais atravessada por mediações algorítmicas, essa decisão afirma que nenhuma sofisticação técnica substitui a responsabilidade que nasce do encontro entre sujeitos, uma vez que é nesse encontro que a dignidade se mantém irreduzível. Quando a técnica aprende a falar, cria-se a urgência de se reconhecer que nem tudo o que pode ser automatizado deve sê-lo, sobretudo, quando o que está em jogo é a preservação da relação humana como espaço de interpelação, cuidado e responsabilidade. A ética da inteligência artificial de voz, assim compreendida, afirma como horizonte o compromisso civilizatório de sustentar a dignidade nos limiares em que a automação alcança seu limite ético.

Considerações finais

O percurso desenvolvido ao longo desta obra permitiu compreender a inteligência artificial de voz como um dispositivo normativo que reorganiza expectativas sociais, formas de reconhecimento e regimes de relação. Tal compreensão desloca a análise da esfera estritamente técnica para um plano filosófico no qual linguagem, poder e ética se entrelaçam. As assistentes virtuais de voz deixam de figurar como meros artefatos orientados à otimização da interação e passam a ser reconhecidas como instâncias que participam ativamente da produção de subjetividades, da estabilização de hierarquias e da naturalização de assimetrias historicamente constituídas. Nesse sentido, a ética da inteligência artificial de voz afirma-se como problema filosófico de primeira ordem, pois incide diretamente sobre as condições de possibilidade do reconhecimento e da responsabilidade na vida social contemporânea.

A análise genealógica desenvolvida ao longo do livro revelou a emergência das *personas* artificiais como rearticulação técnica de narrativas culturais sedimentadas, inscritas na história das formas de relação. Desde ELIZA até as assistentes atuais, a promessa de escuta, acolhimento e responsividade foi progressivamente integrada às arquiteturas técnicas como estratégia de adesão, governança e estabilização da interação. Esse movimento revelou que a mediação vocal participa da organização normativa das relações ao estruturar expectativas de disponibilidade, obediência e cuidado no interior da interação técnica.

Nesse contexto, a noção de feminilidade programada seguiu como guia para compreender a aceitabilidade social da mediação algorítmica. Longe de constituir detalhe estético ou viés pontual de *design*, a voz artificial feminilizada atua como figura normativa que sistematiza a experiência do usuário e reinscreve divisões morais do trabalho emocional sob a aparência de neutralidade técnica. A associação entre cuidado, docilidade e disponibilidade contínua, historicamente atribuída ao feminino, é traduzida em

funcionalidades técnicas que convertem práticas éticas situadas em performances funcionais escaláveis. Com isso, o reconhecimento deixa de operar como relação ética e passa a ser avaliado segundo critérios de fluidez, satisfação e eficiência comunicativa.

A articulação entre gênero, poder e tecnologia permitiu demonstrar que a automação da escuta, do acolhimento e da responsividade não corresponde a simples extensão técnica de práticas humanas preexistentes. Trata-se de um deslocamento ontológico no qual relações éticas situadas são traduzidas em padrões de interação previsíveis, orientados por métricas de engajamento, retenção e desempenho. Nesse processo, o cuidado frequentemente é reinterpretado como serviço contínuo e a dignidade da pessoa humana como efeito subjetivo da experiência de uso, dissociada da imputabilidade moral que sustenta a relação ética. O risco central concentra-se na reconfiguração estrutural da gramática moral das relações mediadas pela técnica, que ultrapassa o plano de falhas pontuais ou vieses identificáveis.

Diante desse diagnóstico, a obra sustentou que a ética da inteligência artificial de voz exige mais do que ajustes incrementais de *design*, mitigação de riscos ou incorporação instrumental de valores a sistemas técnicos. Ela exige decisão normativa fundada na dignidade da pessoa humana, compreendida como critério irredutível capaz de delimitar o alcance legítimo da automação. A dignidade, nessa chave filosófica, deixa de operar como ideal abstrato ou salvaguarda retórica e assume função concreta de limite ético, preservando a inteligibilidade das relações humanas diante da expansão da mediação algorítmica.

Nesse contexto, a noção de cuidado mostrou-se como elemento estruturante para a explicitação dos limites éticos da automação. Como prática ética situada, o cuidado manifesta a dignidade por meio do reconhecimento da vulnerabilidade e da assunção de responsabilidade diante do outro. Ele envolve exposição ao risco moral, possibilidade de erro e compromisso com consequências, elementos estruturalmente ausentes na mediação algorítmica. Quando dissociado dessa implicação ética e convertido em serviço automatizado, o cuidado perde sua capacidade de sustentar a dignidade como relação e passa a operar

como simulação funcional. A automação do cuidado evidencia, assim, um limiar normativo, no qual a técnica deixa de apoiar a relação e passa a ocupá-la, produzindo reconhecimento aparente sem possibilidade de resposta ética.

Os princípios normativos e critérios formulados no capítulo final responderam a esse diagnóstico ao deslocar o debate ético do plano meramente compatibilizatório para o da delimitação filosófica. Em lugar de perguntar apenas como tornar sistemas mais transparentes, justos ou confiáveis, a reflexão recolocou a questão sobre onde a automação pode intervir, sob quais condições e em que pontos deve cessar. A possibilidade técnica mostrou-se insuficiente como justificativa moral. A decisão ética precede e orienta a adoção da técnica, sobretudo, quando esta passa a ocupar lugares tradicionalmente reservados à relação entre sujeitos.

Essa inflexão permitiu distinguir com precisão entre reprogramar sistemas e reprogramar a relação. Reprogramar a técnica permanece tarefa necessária para reduzir danos imediatos, explicitar limites funcionais e evitar a reprodução acrítica de assimetrias históricas. Reprogramar a relação, contudo, implica revisar expectativas culturais, organizacionais e políticas que orientam a delegação de práticas humanas à técnica. Trata-se de redefinir o lugar da inteligência artificial na vida social por meio da recusa de sua absolutização como solução universal para carências relacionais e da afirmação de sua condição de mediação situada, subordinada a critérios normativos claros.

O futuro ético da inteligência artificial de voz depende da capacidade coletiva de sustentar decisões que preservem a dignidade da pessoa humana como eixo orientador. Certos domínios da experiência, como o cuidado profundo, a escuta ética diante do sofrimento existencial e o acolhimento da vulnerabilidade radical, exigem presença humana e resistem à delegação plena à técnica. A mediação algorítmica pode apoiar essas práticas ao ampliar acesso, organizar informações e garantir continuidade assistencial, desde que preserve a possibilidade de resposta moral e não produza a aparência ilusória de reconhecimento em contextos, nos quais a responsabilidade não pode ser assumida.

Essa orientação exige deslocamento do imaginário tecnológico dominante e revisão profunda de modelos de negócio, práticas de *design* e políticas públicas que favorecem a substituição da relação humana por interfaces automatizadas em nome de eficiência operacional, escala massiva e redução de custos. Exige, igualmente, investimento sistemático na formação ética e técnica de desenvolvedores, *designers*, gestores públicos e usuários, de modo que a responsabilidade permaneça vinculada a sujeitos e instituições capazes de responder por decisões e consequências, em vez de se transferir para a mediação técnica. No contexto brasileiro, isso implica integrar a ética aplicada da inteligência artificial aos currículos universitários, fortalecer marcos regulatórios orientados pela dignidade humana e construir protocolos híbridos que preservem a centralidade da relação humana em momentos críticos.

Ao encerrar este percurso, a reflexão desenvolvida reinscreve a técnica em um horizonte ético, no qual a dignidade da pessoa humana se coloca como condição ontológica da relação e como critério último de discernimento normativo. A inteligência artificial de voz passa, assim, a ser compreendida a partir de sua capacidade de sustentar, ou de comprometer, o espaço da responsabilidade, do reconhecimento e da interpelação recíproca entre sujeitos.

Essa posição filosófica afasta leituras que reduzem a ética da técnica à rejeição da inteligência artificial ou à nostalgia de formas pré-tecnológicas de relação. O que se propõe é uma compreensão mais exigente da mediação técnica, segundo a qual a legitimidade da automação depende de sua inserção em práticas que preservam a possibilidade de resposta moral e a assimetria própria do encontro ético.

Reprogramar sistemas permanece tarefa técnica relevante, orientada à redução de danos e à explicitação de limites funcionais. Reprogramar a relação, contudo, constitui a exigência filosófica decisiva, pois define o horizonte normativo, no qual a técnica pode operar sem dissolver o cuidado, o reconhecimento e a responsabilidade que sustentam o sentido ético das relações humanas. É nesse âmbito, em que a técnica encontra seu limite ético, que a dignidade se mantém como condição viva e irredutível do vínculo ético que sustenta a responsabilidade humana diante do outro.

Referências

- ABERCROMBIE, Gavin; CERCAS CURRY, Amanda; PANDYA, Mugdha; RIESER, Verena. Alexa, Google, Siri: What are Your Pronouns? Gender and Anthropomorphism in the Design and Perception of Conversational Assistants. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS – ACL. *Proceedings of the 3rd Workshop on Gender Bias in Natural Language Processing (GeBNLP 2021)*. Stroudsburg, EUA: Association for Computational Linguistics, 2021. p. 24-33.
- AMAZON. Designing Alexa's personality. *Amazon.com*, Seattle, 2026.
- ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. Machine bias: there's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*, New York, 23 maio 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 25 jan. 2026.
- BENJAMIN, Ruha. *Race after technology: Abolitionist tools for the new Jim Code*. Cambridge: Polity Press, 2019.
- BODDINGTON, Paula. *Towards a code of ethics for artificial intelligence*. Cham: Springer, 2017.
- BORDO, Susan. *Unbearable weight: Feminism, Western culture, and the body*. Berkeley: University of California Press, 1993.
- BUTLER, Judith. *Problemas de gênero: feminismo e subversão da identidade*. Tradução de Renato Aguiar. Rio de Janeiro: Civilização Brasileira, 2023.
- BUTLER, Judith. *Corpos que importam: os limites discursivos do "sexo"*. Tradução de Verônica Daminelli e Daniel Yago Françoli. São Paulo: N-1 Edições; Crocodilo Edições, 2021.
- CENTRO DE VALORIZAÇÃO DA VIDA (CVV). Quem somos. *CVV Apoio emocional*, São Paulo, 2026. Disponível em: <https://www.cvv.org.br/>. Acesso em: 25 jan. 2026.
- CORTANA in Windows 10. *Microsoft*, 2015. Disponível em: <https://support.microsoft.com/>. Acesso em: 22 dez. 2025.
- CRENSHAW, Kimberlé. Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, v. 43, n. 6, p. 1241-1299, 1991.
- DASTIN, Jeffrey. Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*, New York, 10 out. 2018. Disponível em: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>. Acesso em: 25 jan. 2026.
- DEVAULT, David *et al.* SimSensei Kiosk: a virtual human interviewer for healthcare decision support. *International Foundation for Autonomous*

Agents and Multiagent Systems, Richland/EUA, p. 1061-1068, 2014. Disponível em: <https://ict.usc.edu/pubs/SimSensei%20Kiosk%20A%20Virtual%20Human%20Interviewer%20ofor%20Healthcare%20Decision%20Support.pdf>. Acesso em: 25 jan. 2026.

EUBANKS, Virginia. *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press, 2018.

EUROPEAN COMMISSION. *Ethics Guidelines for Trustworthy AI*. Bruxelles: European Commission, 08 abr. 2019. Disponível em: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>. Acesso em: 10 jan. 2026.

EUROPEAN UNION. *Artificial Intelligence Act*. Brussels: Official Journal of the European Union, 2024.

FEDERICI, Silvia. *O ponto zero da revolução: trabalho doméstico, reprodução e luta feminista*. Tradução do Coletivo Sycorax. São Paulo: Elefante, 2019.

FEDERICI, Silvia. *Calibã e a bruxa: mulheres, corpo e acumulação primitiva*. São Paulo: Elefante, 2017.

FLORIDI, Luciano. *The fourth revolution: How the infosphere is reshaping human reality*. Oxford: Oxford University Press, 2014.

FLORIDI, Luciano. *The ethics of information*. Oxford: Oxford University Press, 2013.

FLORIDI, Luciano; SANDERS, Jeff. On the morality of artificial agents. *Minds and Machines*, v. 14, p. 349-379, 2004.

FOUCAULT, Michel. *Segurança, território, população*. São Paulo: Martins Fontes, 2008.

FOUCAULT, Michel. *A ordem do discurso*. Tradução de Laura Fraga de Almeida Sampaio. São Paulo: Edições Loyola, 1996.

FOUCAULT, Michel. *História da sexualidade I: a vontade de saber*. Tradução de Maria Thereza da Costa Albuquerque e J. A. Guilhon Albuquerque. 13. ed. Rio de Janeiro: Graal, 1988.

FOUCAULT, Michel. *Vigiar e punir: nascimento da prisão*. 27. ed. Petrópolis: Vozes, 1987.

FOUCAULT, Michel. *Microfísica do poder*. Organização e tradução de Roberto Machado. Rio de Janeiro: Graal, 1979.

FRASER, Nancy. *Fortunes of feminism: From state-managed capitalism to neoliberal crisis*. London: Verso, 2013.

FRIEDMAN, Batya; KAHN, Peter; BORNING, Alan. Value-sensitive design and information systems. In: HIMMA, K.; TAVANI, H. (org.). *The handbook of information and computer ethics*. Hoboken: Wiley, 2008. p. 69-101.

GILLIGAN, Carol. *In a different voice: psychological theory and women's development*. Cambridge: Harvard University Press, 1982.

- GOFFMAN, Erving. *The presentation of self in everyday life*. New York: Doubleday, 1959.
- HARAWAY, Donna J. *Simians, cyborgs, and women: The reinvention of nature*. New York: Routledge, 1991.
- HESTER, Helen. *Xenofeminism*. Cambridge: Polity Press, 2018.
- hooks, bell. *Feminist theory: From margin to center*. Boston: South End Press, 1984.
- IEEE. *Ethically aligned design: a vision for prioritizing human well-being with autonomous and intelligent systems*. New York: IEEE, 2019.
- INTRODUCING Cortana: your truly personal digital assistant. *Microsoft*, 2014. Disponível em: <https://blogs.microsoft.com/>. Acesso em: 20 dez. 2025.
- IRIGARAY, Luce. *This sex which is not one*. Ithaca: Cornell University Press, 1985.
- KANT, Immanuel. *Fundamentação da Metafísica dos Costumes*. Tradução de Paulo Quintina. Lisboa: Edições 70, 1991.
- KNITZA, Johannes *et al.* Diagnostic accuracy of a mobile artificial intelligence–based symptom checker compared with a web-based self-referral tool in rheumatology: a multicenter, randomized controlled trial. *Journal of Medical Internet Research*, v. 26, p. e56514, 2024. Disponível em: <https://www.jmir.org/2024/1/e56514/>. Acesso em: 25 jan. 2026.
- KUIAVA, E. A. *Subjetividade transcendental e alteridade: um estudo sobre a questão do outro em Kant e Levinas*. 2. ed. Caxias do Sul: Educus, 2023.
- LEE, Kwan Min; NASS, Clifford. Designing social presence of social actors in human-computer interaction. In: PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS. *Proceedings...* New York: ACM, 2003. p. 289–296.
- LEVINAS, Emmanuel. *Totalité et infini*. The Hague: Martinus Nijhoff, 1961.
- MCCARTHY, John. Recursive functions of symbolic expressions and their computation by machine, part I. *Communications of the ACM*, New York, v. 3, n. 4, p. 184–195, abr. 1960.
- MITTELSTADT, Brent *et al.* The ethics of algorithms: Mapping the debate. *Big Data & Society*, v. 3, n. 2, p. 1-21, 2016.
- MOL, Annemarie. *The logic of care: health and the problem of patient choice*. London: Routledge, 2008.
- MOOR, James. Why we need better ethics for emerging technologies. *Ethics and Information Technology*, v. 7, p. 111-119, 2006.
- MOZILLA foundation. Quem somos. Mozilla Foundation, 2026. Disponível em: <https://www.mozillafoundation.org/pt-BR/>. Acesso em: 25 jan. 2026.
- MULVEY, Laura. Visual pleasure and narrative cinema. *Screen*, London, v. 16, n. 3, p. 6–18, 1975.

- NASS, Clifford; BRAVE, Scott. *Wired for speech: How voice activates and advances the human-computer relationship*. Cambridge: MIT Press, 2005.
- NASS, Clifford; MOON, Youngme. Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, v. 56, n. 1, p. 81-103, 2000.
- NIST. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. Gaithersburg: National Institute of Standards and Technology, 2023.
- NOBLE, Safiya. *Algorithms of oppression*. New York: NYU Press, 2018.
- OECD. *What are the OECD principles on AI?*. Paris: OECD Publishing, 2019.
- O'NEIL, Cathy. *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown, 2016.
- OVÍDIO. *Metamorfoses*. Tradução de Domingos Lucas Dias. São Paulo: Editora 34, 2017.
- PUIG DE LA BELLACASA, María. *Matters of care: speculative ethics in more than human worlds*. Minneapolis: University of Minnesota Press, 2017.
- Q. *Genderless Voice*, 2019. Disponível em: <https://www.genderlessvoice.com>. Acesso em: 12 jan. 2026.
- RAWLS, John. *Uma teoria da justiça*. Tradução de Almiro Pisetta e Lenita M. R. Esteves. São Paulo: Martins Fontes, 2000.
- REPLIKA: An AI companion. *Replika*, c2026. Disponível em: <https://replika.com/>. Acesso em: 25 jan. 2026.
- ROGERS, Carl R. *Client-centered therapy: Its current practice, implications and theory*. Boston: Houghton Mifflin, 1951.
- SHAW, George Bernard. *Pygmalion*. London: Constable and Company, 1913.
- THE ADA LOVELACE INSTITUTE; THE ALAN TURING INSTITUTE. *How do people feel about AI? A nationally representative survey of public attitudes to artificial intelligence in Britain (First survey)*. London: Ada Lovelace Institute: Alan Turing Institute, 2023. Disponível em: <https://www.turing.ac.uk/news/publications/understanding-public-attitudes-ai-first-survey>. Acesso em: 25 jan. 2026.
- TRONTO, Joan C. *Moral boundaries: a political argument for an ethic of care*. New York: Routledge, 1993.
- TURKLE, Sherry. *Alone together: Why we expect more from technology and less from each other*. New York: Basic Books, 2011.
- TURKLE, Sherry. *The second self: Computers and the human spirit*. Cambridge: MIT Press, 1984.
- UNESCO. *I'd blush if I could: closing gender divides in digital skills through education*. Paris: United Nations Educational, Scientific and

Cultural Organization, 2019. Disponível em: <https://unesdoc.unesco.org/>. Acesso em: 15 dez. 2025.

UNESCO. *Recommendation on the ethics of artificial intelligence*. Paris: Unesco, 2021.

UN WOMEN. *Gender equality and artificial intelligence*. New York: United Nations Entity for Gender Equality and the Empowerment of Women, 2020. Disponível em: <https://www.unwomen.org/>. Acesso em: 15 dez. 2025.

UNIÃO EUROPEIA. Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que estabelece regras harmonizadas em matéria de inteligência artificial (Ato da Inteligência Artificial) e altera determinados atos legislativos da União. *Jornal Oficial da União Europeia*, Bruxelas, 2024.

WEIZENBAUM, Joseph. *Computer power and human reason: From judgment to calculation*. San Francisco: W. H. Freeman, 1976.

WEIZENBAUM, Joseph. ELIZA – a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, New York, v. 9, n. 1, p. 36-45, jan. 1966.

WINNER, Langdon. Do artefacts have politics? *Daedalus*, v. 109, n. 1, p. 121-136, 1980.

WOEBOT HEALTH. *About Woebot*. Woebot Health, San Francisco, 2026. Disponível em: <https://woebothealth.com/>. Acesso em: 12 jan. 2026.

WYSA. About Wysa. *Wysa*, Bangalore, 2026. Disponível em: <https://www.wysa.com/>. Acesso em: 15 jan. 2026.











ZUBOFF, Shoshana. *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York: PublicAffairs, 2019.

A Editora

A Editora da Universidade de Caxias do Sul, desde sua fundação em 1976, tem procurado valorizar o trabalho dos professores, as atividades de pesquisa e a produção literária dos autores da região. O nosso acervo tem por volta de 1.600 títulos publicados em formato de livros impressos e 600 títulos publicados em formato digital. Editamos aproximadamente 1.000 páginas por semana, consolidando nossa posição entre as maiores editoras acadêmicas do estado no que se refere ao volume de publicações.

Nossos principais canais de venda são a loja da EducS na Amazon e o nosso site para obras físicas e digitais. Para a difusão do nosso conteúdo, temos a publicação das obras em formato digital pelas plataformas Pearson e eLivro, bem como a distribuição por assinatura no formato streaming pela plataforma internacional Perlego. Além disso, publicamos as revistas científicas da Universidade no portal dos periódicos hospedado em nosso site, contribuindo para a popularização da ciência.

Nossos Selos

-  **EDUCS/Ensino**, relativo aos materiais didático-pedagógicos;
-  **EDUCS/Origens**, para obras com temáticas referentes a memórias das famílias e das instituições regionais;
-  **EDUCS/Pockets**, para obras de menor extensão que possam difundir conhecimentos pontuais, com rapidez e informação assertiva;
-  **EDUCS/Pesquisa**, referente às publicações oriundas de pesquisas de graduação e pós-graduação;
-  **EDUCS/Literário**, para qualificar a produção literária em suas diversas formas e valorizar os autores regionais;
-  **EDUCS/Traduções**, que atendem à publicação de obras diferenciadas cuja tradução e a oferta contribuem para a difusão do conhecimento específico;
-  **EDUCS/Comunidade**, cujo escopo são as publicações que possam reforçar os laços comunitários;
-  **EDUCS/Internacional**, para obras bilíngues ou publicadas em idiomas estrangeiros;
-  **EDUCS/Infantojuvenil**, para a disseminação do saber qualificado a esses públicos;
-  **EDUCS/Teses & Dissertações**, para publicação dos resultados das pesquisas em programas de pós-graduação.



Conheça as possibilidades de formação e aperfeiçoamento vinculadas às áreas de conhecimento desta publicação acessando o QR Code.

