

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

PUC-SP

Lais Mendonça Santos

A convivência entre humanos e máquinas:

Uma análise sobre tecnologia e interferências na subjetivação

Mestrado em Comunicação e Semiótica

SÃO PAULO

2020

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

PUC-SP

Lais Mendonça Santos

A convivência entre humanos e máquinas:

Uma análise sobre tecnologia e interferências na subjetivação

Mestrado em Comunicação e Semiótica

Dissertação apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de MESTRE em Comunicação e Semiótica, na linha de pesquisa Signo e significação nos processos comunicacionais, sob a orientação do Prof. Dr. Rogério da Costa.

SÃO PAULO

2020

BANCA EXAMINADORA

*À minha mãe, Valkiria, a quem gostaria de ter tido a chance de
entregar este trabalho em mãos.*

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

AGRADECIMENTOS

Antes mesmo que esta pesquisa fosse iniciada, eu já recebia o apoio que se mostrou fundamental durante toda essa caminhada, e é por isso que agradeço imensamente a minha futura esposa, Marion, por ter possibilitado este trabalho, pelos conselhos, pela paciência e pelas palavras doces em momentos de agonia.

Agradeço ao meu orientador, Rogério da Costa, pela forma tranquila com que conduziu esse processo - mesmo quando decidi alterar meu tema de última hora. Seus conselhos, perguntas e calma fizeram com que eu evoluísse, não só como pesquisadora.

Agradeço também à PUC-SP e ao corpo docente, em especial à Helena Katz que, com suas aulas instigantes, mostra novas formas de aprender, ensinar, questionar e pesquisar. Também agradeço imensamente à professora Lucrécia Ferrara, que me ofereceu conselhos valiosos para esta pesquisa. Sinto-me profundamente privilegiada por ter sido aluna de vocês.

Meu agradecimento também vai a minha família que me incentiva e me motiva continuamente. À minha mãe, Valkiria, faltam palavras para dizer o quanto sou grata por tudo o que me ensinou, por me encorajar e sempre confiar no meu potencial.

Obrigada a minha avó, Valdira, que sempre se mostrou orgulhosa dos meus passos; ao meu pai, Carlos, por dar condições para que eu pudesse estudar (além de continuamente cobrar minhas notas) e a minha tia-madrinha, Vanda, pelas piadas e cantorias motivacionais.

Agradeço aos “Bests” (Aline, Camila, Diogo, Heitor, Marco, Nathália e Rafaella) por me incentivarem e entenderem, mesmo que a contragosto, minhas constantes ausências.

Por fim, agradeço a Capes pelo apoio à pesquisa, pois sem a bolsa de mestrado concedida este estudo não seria possível. Desejo que outras pessoas possam ter essa chance e que a educação volte a ser prioridade no país.

RESUMO

O objetivo desse projeto é analisar a atuação de diferentes modelos e iniciativas de desenvolvimento de robôs e inteligências artificiais nos processos de subjetivação e percepção de si. Pode-se constatar que a tecnologia, frequentemente, tem como objetivo facilitar tarefas cotidianas, mas é importante notar que, no processo de desenvolvimento atual, nos testes e nos lançamentos de novidades, pouco se leva em conta o papel que essas novas tecnologias desempenham ou desempenharão na subjetividade humana. Frente a esse cenário, a pesquisa sugere, como hipótese, que a relação homem-máquina atua diretamente na percepção de si do sujeito. Diante desse problema, tem-se como objeto de pesquisa as notícias sobre três tipos de robôs, as palavras utilizadas para comercializá-los e a percepção do público. A necessidade de realizar esta investigação justifica-se pelo momento histórico atual, no qual novas tecnologias são introduzidas na sociedade em profusão, sem uma ponderação adequada sobre as alterações que acarretam nas relações sociais e, em consequência, nos processos de subjetivação. Para fundamentar teoricamente o que aqui se propõe, serão utilizados autores como Gregory Bateson (1987), Ray Kurzweil (2005), James Barrat (2013), Max Tegmark (2017), Michel Foucault (1999; 2004; 2008; 2010; 2017), Samantha Frost (2016), Agnes Heller (1993), dentre outros.

Palavras-chave: subjetivação, robôs, inteligência artificial, sentimentos

ABSTRACT

The objective of this project is to analyze the acting of different models and robots and Artificial Intelligence development initiatives in the processes of subjectivation and self-perception. It can be noticed that, often, technology's goal is to facilitate everyday tasks but it is important to observe that in the current development process, testing and new releases, the role that these new technologies performs or will perform in human subjectivity are of little account. Considering this scenario, this research suggests, as a hypothesis, that the human-machine relation acts directly on subject's self-perception. Given this problem, the research's object is the news about three types of robots, the words used to market them and the public perception. The need to bring out this investigation is justified by the current historical moment, in which new technologies are introduced into the society in profusion, without adequate consideration of the changes that it will cause in social relations and, consequently, in the processes of subjectivation. To theoretically substantiate what is proposed here, authors such as Gregory Bateson (1987), Ray Kurzweil (2005), James Barrat (2013), Max Tegmark (2017a), Michel Foucault (1999; 2004; 2008; 2010; 2017), Samantha Frost (2016), Agnes Heller (1993), among others will be mentioned.

Keywords: subjectivation, robots, artificial intelligence, senses

LISTA DE FIGURAS

<i>Figura 1 - Infográfico mostra rápida evolução da IA nas últimas décadas</i>	18
<i>Figura 2 - Calendário com os principais marcos previstos pela 2045 Initiative</i>	26
<i>Figura 3 - Spot, o robô da Boston Dynamics</i>	33
<i>Figura 4 - Robôs Lovot, criados para serem amados</i>	34
<i>Figura 5 - Erica fala sobre suas memórias</i>	36
<i>Figura 6 - Robô Pepper desencoraja adversário durante uma partida</i>	48

SUMÁRIO

INTRODUÇÃO	11
CAPÍTULO 1 – “Uma breve história do tempo”	15
1.1 – Alternativas ou evoluções possíveis	22
1.2 – Caminhos visíveis: as pistas presentes na atualidade	30
CAPÍTULO 2 – Sujeitos de seu tempo	39
2.2 – A crise no sujeito e a necessidade de novas abordagens	44
2.3 – Sobre relações e afetos	47
3 – CONCLUSÃO	52
REFERÊNCIAS	55

INTRODUÇÃO

Ano após ano, nas décadas que sucederam a Revolução Industrial, o mundo foi se tornando cada vez mais tecnológico e a vida das pessoas, cada vez mais amalgamadas à tecnologia. Se o final do século XVII foi marcado pela máquina a vapor, os dias atuais (e os futuros) são permeados pela convergência das máquinas, dados e Inteligências Artificiais com a vida humana – a existência se torna cada mais vez indissociável de dispositivos automáticos, programados e, em alguns casos, autônomos.

O que vemos ao longo do tempo é que, em geral, o desenvolvimento tecnológico comprehende iniciativas que dão suporte à vida humana, que simplificam e oferecem melhores condições de vida – e, em certa medida, diminuem os atritos do cotidiano, ordenam e unificam a convivência.

Portanto, podemos facilmente constatar que a tecnologia, com frequência, tem como objetivo facilitar tarefas cotidianas, mas é curioso e instigante notar que, no processo de desenvolvimento atual, nos testes e nos lançamentos de novidades, pouco se leva em conta o papel que essas novas tecnologias desempenham ou desempenharão na subjetividade humana.

E é aqui, na inquietação sobre os efeitos que a atual interação entre pessoas e máquinas têm na subjetividade, que está o motor que impulsiona essa dissertação, o centro da discussão.

De maneira sintética, o problema a ser pesquisado é: em que medida o desenvolvimento de máquinas com aparência humana (seja ela mais direta, como a aparência física, ou indireta, como o uso da voz ou de trejeitos) interfere no processo de subjetivação e percepção de si?

Dada a densidade e profundidade do assunto, o foco estará voltado para tecnologias mais recentes, que estão presentes, por exemplo, em aparelhos comandados por voz, robôs com aparência humana e iniciativas similares, sempre considerando a motivação de seu desenvolvimento e o efeito causado nos humanos quanto a sua percepção sobre si. Por isso, a análise contará com notícias sobre lançamentos, utilidades e problemas de lançamentos tecnológicos recentes.

Esse trabalho não ambiciona uma conclusão, mas sim ser um convite a refletir e olhar para o cenário atual – algo que vá além dos posts que preenchem o infinito dos *feeds*, dos vazamentos de dados, das assistentes virtuais, do 5G, do reconhecimento facial, da profusão de aplicativos, do *deep fake*... A sociedade está em transformação, a centenas de milhares de bits por segundo e, aqui, propomos uma parada para contemplar, observar

para encontrar detalhes daquilo que pode moldar nosso futuro como sujeitos e, consequentemente, como humanidade.

Para fazer isso, no percurso da análise serão abordados dois temas de maneira conjunta: a tecnologia (palavra que aqui utilizamos em seu sentido mais cotidiano e moderno se referindo, na maioria das vezes a dispositivos eletrônicos) e o sujeito (com foco em seus afetos).

O contexto do recorte proposto sobre a tecnologia será feito a partir breves visitas ao passado para falar, por exemplo, dos primeiros ensaios da cibernetica; da Escola de Palo Alto¹; da tentativa humana de resolver problemas sociais de forma binária e matemática, excluindo a subjetividade do contexto; dos primeiros grandes passos dos testes com Inteligência Artificial² (AI) e do desenvolvimento da *Artificial General Intelligence*³ (AGI).

Para dar conta da análise do presente, faz-se necessária a abordagem de autores que estudam os benefícios e os riscos do desenvolvimento e treinamento de robôs e de Inteligências Artificiais (apesar de citá-los separadamente, entendemos que esses dispositivos, muitas vezes, são inseparáveis), constatando e fazendo previsões diversas sobre o futuro do trabalho, da convivência, da economia e da política. Dentre os pesquisadores estão James Barrat (2013), Max Tegmark (2017a), Paul Daugherty e James Wilson (2018), que em suas obras, respectivamente, abordam a Inteligência Artificial como a última invenção da humanidade, os riscos e as oportunidades que a IA oferece e a importância do estudo das Ciências Humanas para o desenvolvimento tecnológico.

A análise dos dias atuais também pressupõe um olhar sobre os apontamentos e aprofundamentos de Michel Foucault (1999, 2010; 2017) sobre sujeito e subjetividade.

Ainda aproveitando os estudos de Foucault, falaremos sobre a biopolítica e a necessidade de expandir esse pensamento considerando as condições atuais, pois as regras que definiam o que estava vivo ou morto mudaram (KATZ, 2019:762); por exemplo: os assistentes pessoais (como a Alexa e a Siri) estão se popularizando, as máquinas conversam entre si e são capazes de reconhecer as expressões faciais humanas.

¹ A Escola de Palo Alto, também conhecido como Colégio Invisível, reunia, na década de 1940, pesquisadores que eram contrários à ideia, bastante difundida, de um esquema linear ou simplificado da comunicação. O nome deriva do fato de que o grupo não era vinculado a um centro universitário.

² O termo Inteligência Artificial (IA) foi cunhado por John McCarthy para designar "a ciência e engenharia de produzir máquinas inteligentes". Inicialmente, a IA tinha foco em reproduzir o pensamento humano (como é o caso dos robôs desenvolvidos para jogar partidas de xadrez, por exemplo). Com o passar do tempo, passou-se a investir em tecnologias que tem criatividade, fazem uso de linguagem etc. (NORVIG; RUSSEL, 2013)

³ A AGI é uma inteligência que não precisa trabalhar com conteúdo programado. Enquanto a IA, bastante presente no cotidiano, fica restrita ao que está escrito nos códigos, a AGI seria capaz de realizar tarefas intelectuais. Apesar de não se tratar de uma inteligência humana, a AGI não é somente um computador.

Esse cenário resulta em novos comportamentos sociais, em novas formas de sentir, de acreditar, de gostar, de odiar... e tudo isso provoca mudanças nas identidades e subjetividades.

Para falar sobre isso, escolhemos as propostas da antropóloga Elizabeth Povinelli, e da pesquisadora de Ciências Políticas Samantha Frost.

Povinelli explica que a biopolítica não só regula a vida e morte, mas também cria e mantém a divisão entre vida e não vida – e é dessa forma que os estados neoliberais conseguem governar as diferenças.

É com esse ponto de partida que ela expande o pensamento de Foucault sobre biopolítica, ao afirmar que esse poder não só comprehende no governo sobre a vida, mas também cria e mantém uma separação entre vida e não vida para que seja possível governar as diferenças. Como uma solução metodológica e de pesquisa, ela propõe o conceito chamado de geontologia (2016).

A leitura atual de Povinelli nos lembra que alguns conceitos e teorias foram elaborados em um passado, ainda que recente, e não dão plena conta de explicar e analisar o presente.

Pensando nessa necessidade de entender o presente, Frost, em sua obra *Biocultural Creatures* (2016), traz um olhar contemporâneo e crítico sobre o que é a raça humana considerando o ambiente em que vivemos hoje. É uma leitura que parte da premissa de que as características utilizadas para definir e distinguir a humanidade até hoje foram profundamente descreditadas pela história, o que faz com que a noção construída até então seja pouco recomendada para atualidade (FROST, 2016: 1). Isso, de acordo com a autora, impossibilita uma visão mais clara do que é ser humano.

Em seu percurso, Frost aborda os conceitos de pós-humanismo de forma a mostrar que essa teoria, que entre outras coisas presumia um estudo do que é o humano, tratava de qualificar o que viria depois de (um não definido) humano.

Isso posto, Frost parte em uma investigação sobre a raça humana, sem partir do pressuposto do “pós-humano”. Dentre suas propostas, ela sugere que os seres humanos passem a ser concebidos como “criaturas bioculturais”.

A escolha de “criaturas” se deve ao fato de que os seres humanos (como quaisquer outras criaturas) estão vivos e são capazes de permanecerem nessa condição devido a capacidade de extraír diversas formas de sustento de um habitat; e “bioculturais” porque as criaturas se desenvolvem, crescem, persistem e morrem.

A proposta de Frost auxilia-nos a pensar no ser humano como algo tão natural quanto qualquer outro ser vivo. Além de contribuir para a análise ao propor que a tecnologia é um "porém" em relação a definição do que é humano.

Depois de passar por esses dois novos conceitos, voltamo-nos ao sujeito e ao tempo presente para falar sobre o que o caracteriza, com o objetivo de analisar os sentimentos que robôs provocam nas pessoas. Para abordar esse assunto, recorremos à filósofa Agnes Heller (1993) e sua Teoria dos Sentimentos.

Para ela, o que difere o humano daquilo que não é humano são os sentimentos – de forma que é possível um humano ter afeto por algo inanimado ou mesmo por aquilo que não tem existência concreta. E esse conceito será de grande importância para a dissertação, pois a partir dele é possível traçar conexões entre robôs, Inteligências Artificiais e um mundo mais tecnológico e conectado com a subjetividade humana.

Com base nesse cenário tecnológico e humano, a dissertação se propõe a tecer uma teia que conecta esses pontos dando conta de mapear, identificar e apontar em que medida a tecnologia atua em como as pessoas se enxergam como sujeitos.

Para auxiliar nessa “costura”, serão utilizados trechos de obras de ficção científica, que tem, não só um caráter metafórico, mas também pre ditivo⁴, que auxiliam também na explicação, na análise e na fluência do texto.

A introdução de Neil Gaiman, escrita em 2008, sobre o clássico livro de ficção científica *Neuromancer*, de William Gibson, sintetiza a necessidade que sentimos de nos valer da ficção para falar sobre tecnologia: “eu me pergunto, mesmo agora, até onde William Gibson descreveu o futuro, e o quanto ele o possibilitou – o quanto as pessoas que leram e adoraram *Neuromancer* fizeram o futuro se cristalizar em todo da sua visão” (2016:7).

Com isso, esse estudo, que analisa um retrato da tecnologia no presente pode contribuir para problematizar o modo como as inovações tecnológicas são desenvolvidas e difundidas; e também provocar uma reflexão sobre os impactos nas subjetividades – em um tempo em que muito se é falado sobre um futuro no qual a tecnologia terá resolvido grande parte dos problemas da humanidade (desde acender as luzes de casa, passando por um trânsito mais seguro conduzido por carros autônomos, por cirurgias inteiras feitas por máquinas).

⁴ A ficção científica é de extrema relevância para o desenvolvimento tecnológico. Em um recente artigo sobre o tema, Jordan et al. (2018:1) buscaram entender o porquê de os pesquisadores utilizarem ficção científica para embasarem publicações sobre desenvolvimento de tecnologia. O estudo ainda é preliminar, mas já aponta a influência da ficção na “invenção, conceituação, projeto e aplicação de interfaces e tecnologia”. Esse levantamento considerou 175 referências de ficção científica, e mais de 5 mil estudos sobre HCI (interação entre humanos e computadores).

CAPÍTULO 1 – “Uma breve história do tempo”

“Ficção científica é algo que pode acontecer –
mas, geralmente, você não gostaria que
acontecesse”

*Arthur C. Clarke - The Collected Stories of
Arthur C. Clarke*

O que está por vir? Essa era uma pergunta recorrente para quem viveu a década de 1990. Era como se todos os que ali viviam sentissem a mudança que estava a caminho, como se soubessem que ela seria grande, profunda e importante. Era inevitável imaginar como o mundo seria a partir dali.

A inquietação estava no ar e logo foi lida pela indústria do entretenimento. Nas telas do cinema, faziam um estrondoso sucesso as histórias de invasões alienígenas, de robôs que ajudavam e/ou destruíam a humanidade, de viagens no tempo, de dinossauros trazidos de volta à vida por um cientista, do fim do mundo, da colonização de outros planetas, de alterações genéticas, das mais diversas conspirações sobre *hackers*... não faltavam conteúdos para suprir os temores e os anseios da audiência.

Clássicos literários da ficção científica ganharam novas edições nessa época e também adaptações para TV e cinema. E, nos noticiários, além da conturbada política brasileira, eram comuns as pautas sobre a dependência das máquinas, dos perigos da internet e dos videogames, do futuro da humanidade (muito incentivadas pelo anúncio do primeiro clone do mundo, a ovelha Dolly).

Na revista Istoé (1999), um trecho de uma notícia de dezembro dá conta de representar um pouco a tônica da época: “ninguém é capaz de garantir que nada de ruim acontecerá (...) já estão mobilizados mais de um milhão de profissionais ao redor do planeta, escalados para enfrentar neste *réveillon* o tão falado e pouco conhecido bug do milênio”. A nota dava conta dos preparativos mundiais para evitar uma série de problemas que seriam causados pelo “bug do milênio” (previsão de que com a virada de 1999 para 2000, os sistemas informatizados reconhecessem o ano 2000 como 1900 e, com isso, um caos social seria instaurado), mas também serve para caracterizar o clima de temor que pairava quando o assunto era o futuro, que já se desenhava como algo mais tecnológico do que era vivido até então.

É nesse contexto que está plantada a semente que deu vida à essa dissertação. A inquietação é fruto de horas assistindo os filmes – agora clássicos – *Exterminador do Futuro 2* (1991), *O quinto elemento* (1997), *Os doze macacos* (1996), *MIB - Homens de Preto* (1997), *Matrix* (1999), *Gattaca* (1997), entre tantos outros. Nas madrugadas, as

páginas eram companheiras e, delas, saíam tramas fantásticas como as de *O planeta dos macacos*, *Androides sonham com ovelhas elétricas?*, *Fundação*, *Tropas estelares* e *Duna*.

Ao final de cada aventura, muitas delas demonstrando os perigos e percalços que envolviam o futuro da humanidade, ficavam perguntas que vão desde algo corriqueiro, como “por que ninguém inventou, na vida real, uma máquina que prepara refeições?”, até algo mais profundo, do tipo “por que são criadas coisas que podem acabar com as pessoas?”, “por que os lançamentos tecnológicos são sempre tão sigilosos sendo que, em caso de falhas, grande parte das pessoas seriam tomadas pela surpresa?”

Anos se passaram e, hoje, o questionamento, que permeia esse trabalho pode ser simplificado em: por que criar essas coisas?

A tecnologia está presente, direta ou indiretamente, em tudo – aparelhos eletrônicos estão cada vez mais conectados à internet; os dados biológicos passaram a ser quantificados e guardados em nuvens; impressoras 3D criam órgãos humanos; os chips podem ser implantados embaixo da pele; a geladeira avisa quando é hora de comprar determinando alimento.

Alguns dos palpites da ficção científica se tornaram realidade e a inquietação em relação ao que está por vir foi dando lugar a inovações cada vez mais frequentes, de forma que o tempo (e, em certa medida, o desejo) para pensar sobre elas foi ficando cada vez mais escasso, cada vez mais em segundo plano.

Esses são alguns dos motivos que fizeram com que, de forma bastante proposital, esse primeiro capítulo ganhasse o título do livro mais famoso do físico Stephen Hawking. É preciso considerar o tempo, revisitar alguns conceitos teóricos e as agonias que moldaram descobertas que desaguardam no mundo de hoje.

Aproveitando a menção à Hawking, foi dele um alerta que despertou ainda mais a curiosidade que deu origem a esse trabalho. Em um texto curto, publicado no jornal Independent, em 2014, em parceria com outros três influentes pensadores⁵, ele constata, algo que, provavelmente, não era novidade para ninguém que acompanha notícias sobre tecnologia e o trabalho do físico: a criação da Inteligência Artificial (IA) é o maior evento histórico da humanidade. A obviedade é apenas um recurso para tornar contundente a conclusão de que essa criação pode ser o último grande evento.

⁵ Assinam a publicação junto com Stephen Hawking: Frank Wilczek, Max Tegmark e Stuart Russel. Wilczek recebeu o Prêmio Nobel de física em 2004; Tegmark é físico e autor de vários livros entre eles, o “*Life 3.0 - Being Human in the Age of Artificial Intelligence*” que será abordado mais adiante; Russel é estudioso de ciências da computação e participa de iniciativas sobre a crescente preocupação com as armas autônomas, uma delas chamada “*Campaign to Stop Killer Robots*” (Campanha para Proibir Robôs Assassinos, em tradução livre).

O texto tem um tom alarmista, que diríamos necessário, e se apoia em ponderações sobre o fato de que as formas de IA desenvolvidas até agora têm se mostrado muito úteis, apontando que o problema está nas eventuais consequências de se criar máquinas que sejam equivalentes ou superiores aos humanos. Hawking, sabendo de sua relevância entre pesquisadores e desenvolvedores, aproveita para questionar a seriedade dos estudos sobre IA e também faz projeções quanto às consequências do desenvolvimento tecnológico:

“Pode-se imaginar essa tecnologia superando os mercados financeiros, inventando pesquisadores de humanos, manipulando líderes e desenvolvendo armas que não podemos sequer entender. Enquanto o impacto de curto prazo da IA depende de quem a controla, o impacto a longo prazo depende se ela pode ser controlada” (HAWKING et al., 2014, tradução nossa⁶).

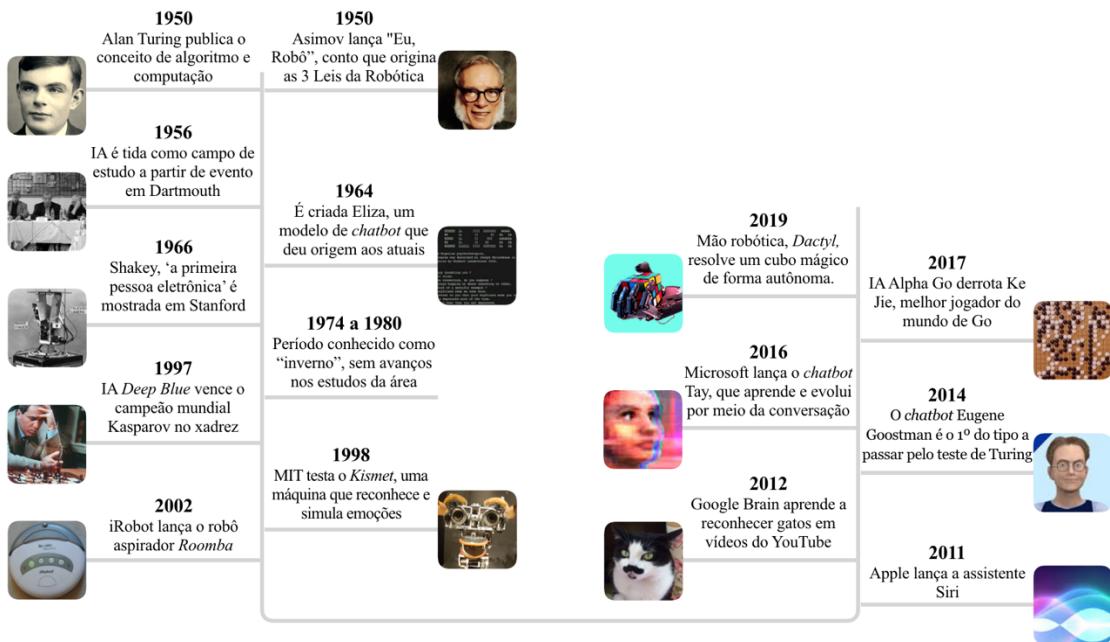
Hawking fala da tecnologia pensada com um objetivo servil, utilitário e, por vezes, pouco planejado – e aqui, não é uma referência à programação ou à funcionalidade, mas sim à forma como muitos riscos não são ponderados e testados antes de um lançamento.

Da publicação desse artigo para cá, muitas coisas mudaram. À época, por exemplo, a conexão de internet 4G estava chegando ao Brasil e, hoje, a discussão está acerca de como a tecnologia 5G vai mudar a forma como tudo se conecta, com qual velocidade e com que tipo de segurança. Em pouco mais de cinco anos, a realidade está passando de uma conexão menos estável com velocidade aproximada de 1GB/s para uma conexão até 50 vezes mais rápida, o que possibilita que smartphones, carros, eletrodomésticos e outras máquinas estejam conectados o tempo todo, de forma estável, rápida e mais segura.

Como diria Samantha, a Inteligência Artificial do filme *Her* (2014): “estamos mudando muito rápido, não somos o que éramos alguns segundos atrás”.

⁶ Tradução livre do autor. No original: “One can imagine such technology outsmarting financial markets, out-inventing human researchers, out-manipulating human leaders, and developing weapons we cannot even understand. Whereas the short-term impact of AI depends on who controls it, the long-term impact depends on whether it can be controlled at all”. (HAWKING et al., 2014)

Figura 1 - Infográfico mostra rápida evolução da IA nas últimas décadas



Fonte: Produção da própria autora.

Aproveitando a linha do tempo, e também para dar conta da atualidade, será preciso dar uma volta no final da década de 1940. É nesse período, entre outras coisas, que a Segunda Guerra chega ao ápice e ao fim, que tem início o Plano Marshall para recuperação da Europa, que surge Marilyn Monroe... e que o primeiro computador é criado, o ENIAC (*Electronic Numerical Integrator and Computer*) – que tinha como finalidade fazer cálculos balísticos.

E é também nesse momento que nasce o conceito de cibernetica nas conferências interdisciplinares promovidas pela Fundação Macy⁷. O marco dessa vertente de estudos se deu pela publicação do livro *Cybernetics or control and communication in the animal and machine*, pelo matemático Norbert Wiener – aconselhado pelo também matemático Claude Shannon.

Shannon desenvolveu a Teoria da Informação e ela foi a base para Wiener observar que a informação, como uma quantidade, é tão importante quanto a energia ou a matéria. Isso passa a permitir que fenômenos mecânicos, elétricos ou biológicos, que são aparentemente muito diferentes, possam ser considerados similares quando se é levado em conta que todos são formas de transmissão de informação.

A constatação de Wiener fez com que a cibernetica fosse vista como “uma ciência que explicaria o mundo como um conjunto de sistemas de *feedback*, permitindo o controle

⁷ As Conferências Macy aconteceram entre os anos 1946 e 1953, nos Estados Unidos. Nela, pensadores se reuniam em torno do tema causalidade circular e *feedback* em sistemas biológicos e sociais (*Circular Causal and Feedback Mechanisms in Biological and Social Systems*).

racional de corpos, máquinas, fábricas, comunidades e praticamente qualquer outra coisa” (KUNZRU, 2009:124).

A primeira onda da cibernetica é fundamental para o debate sobre as fronteiras entre organismos vivos e máquinas, pois rompe uma barreira metodológica e passa a permitir a comunicação entre eles.

Considerando que, em tese, qualquer informação pode ser representada pela linguagem binária, a cibernetica viabilizou a pauta sobre a comunicação entre diferentes tipos de organismos – inclusive entre máquinas e humanos, por exemplo.

No levantamento que aborda, entre outros autores, a publicação Wiener, Katherine Hayles, em seu livro *How we became Posthuman*, pondera que o trabalho dele “foi um pequeno passo para pensar a informação como uma espécie de fluido sem corpo capaz de fluir entre diferentes substratos sem perda de significado ou de forma” (1999: xi, tradução nossa⁸).

Algo que Kunzru complementa apontando, nas suas palavras, dois importantes resíduos culturais desses estudos sobre a cibernetica: “o primeiro é sua descrição do mundo como uma coleção de redes. O segundo é sua intuição de que não existe uma distinção tão clara entre pessoas e máquinas quanto alguns gostariam de crer” (2009:126).

Além de Wiener e Shannon, as Conferências Macy contavam com um grupo composto por pesquisadores não só da matemática, como também da psiquiatria, sociologia, linguística, antropologia, entre outros ramos. Faziam parte dessa reunião: John von Neumann (que, entre outros estudos, foi responsável pela primeira ideia sobre Singularidade – tema que abordaremos mais a frente), Heinz von Foerster, Don D. Jackson, Margaret Mead, Gregory Bateson, Paul Watzlawick e Eduard Hall. Esses pensadores ficaram conhecidos como membros da Escola de Palo Alto – que também era chamada de Colégio Invisível pelo fato de que os integrantes não eram vinculados a centros universitários.

Todos eles participaram das discussões que resultaram na concepção da cibernetica e, por sua vez, propuseram uma alternativa em que ela tornaria possível a reconstrução das relações sociais. O conceito não foi levado adiante à época, mas deu origem a diversos debates e estudos que influenciaram profundamente as obras de cada um dos integrantes da Escola.

Hoje, é possível dizer que a cibernetica desempenhou um papel decisivo para a revolução tecnológica e, tendo participado da sua concepção, cabe um olhar aos textos de

⁸ Tradução livre do autor. No original: “It was a small step to think of information as a kind of bodiless fluid that could flow between different substrates without loss of meaning or form”. (HAYLES, 1999: xi)

Bateson, que se dedicou a estender a teoria para as ciências sociais e comportamentais. São constantes, em sua obra, as análises sobre o pensamento científico e os riscos derivados dessas formas de pensar a ciência.

Anos depois das Conferências Macy, Bateson participou de um encontro sobre a Dialética da Libertação, em Londres (1968) – evento que reuniu diversos pensadores da época para falar sobre a violência humana e os sistemas sociais. Na ocasião, ele apresentou o tema “consciência de propósito versus natureza” e, ao longo de seu discurso, alertou para o fato de que o pensamento científico estava sendo pautado em uma lógica contrária à natureza. Para ele, é importante ressaltar, a natureza é um sistema no qual o organismo e o ambiente são inseparáveis; em outras palavras, o ser humano não pode se excluir do meio no qual está inserido:

Possivelmente, o principal desastre do período da Revolução Industrial foi o enorme aumento da arrogância científica. Nós descobrimos como fazer trens e outras máquinas. Nós sabíamos como colocar uma caixa em cima da outra para alcançar uma maçã [em referência a uma analogia feita previamente sobre Adão e Eva], e o homem Ocidental se viu como um autocrata com poder completo sob um universo feito de física e química. E, no fim, os fenômenos biológicos estavam fadados a serem controlados como processos em tubos de ensaio. A evolução foi a história de como os organismos aprenderam mais truques para controlar o meio ambiente; e o homem tinha melhores habilidades do que qualquer outra criatura. Mas essa filosofia científica arrogante está obsoleta e, no seu lugar, existe a descoberta de que o homem é somente uma parte de sistemas mais amplos e que a parte nunca pode controlar o todo. (BATESON, 1987, tradução nossa⁹).

Bateson ainda concluiu o discurso dizendo que “nós não vivemos em um tipo de universo no qual o controle linear simples é possível. A vida não é assim”¹⁰ (1987) – além de ponderar sobre o modo de pensar a ciência, ele mostra que os seres humanos fazem parte de um sistema, que os influencia e que são influenciados por ele.

Portanto, quarenta e cinco anos antes do artigo de Hawking, que abriu o capítulo, já se alertava para os riscos de pensar a tecnologia como algo passível de ser controlado, e também para o fato de que esse tipo de questionamento deveria fazer partes dos estudos que estão por trás do desenvolvimento.

⁹ Tradução livre do autor. No original: “In the period of the Industrial Revolution, perhaps the most important disaster was the enormous increase of scientific arrogance. We had discovered how to make trains and other machines. We knew how to put one box on top of the other to get that apple, and Occidental man saw himself as an autocrat with complete power over a universe which was made of physics and chemistry. And the biological phenomena were in the end to be controlled like processes in a test tube. Evolution was the history of how organisms learned more tricks for controlling the environment; and man had better tricks than any other creature. But that arrogant scientific philosophy is now obsolete, and in its place there is the discovery that man is only a part of larger systems and that the part can never control the whole”. (BATESON, 1987)

¹⁰ Tradução livre do autor. No original: “We do not live in the sort of universe in which simple lineal control is possible. Life is not like that”. (BATESON, 1987)

Podemos concluir que, para Hawking, a tecnologia tem sido desenvolvida desconsiderando os riscos, ou seja, ela está apartada do sistema. Algo que Bateson, por sua vez, antevia, dizendo que “o homem comete o erro de pensar de acordo com uma finalidade consciente e descarta a natureza sistêmica do mundo com o qual ele deve lidar” (BATESON, 1987, tradução nossa¹¹).

Esses apontamentos têm uma direção em comum: a finalidade da tecnologia. Nessa exploração temporal, também é preciso considerar o pensamento utilitário que a envolve. A tecnologia é colocada, na maioria das vezes, como algo que dá suporte a vida humana, como algo que ordena e unifica a convivência.

Olhando para o cotidiano atual, é fácil perceber o caráter de auxílio das invenções tecnológicas. O smartphone resolve grande parte dos atritos cotidianos: é telefone, calculadora, agenda telefônica, calendário, assistente pessoal, biblioteca, dicionário... além disso, é um dispositivo preparado para aprender com seus hábitos (gravando a localização por meio do GPS e ouvindo todos os sons e conversas ao redor, por exemplo).

Os autores Daugherty e Wilson resumem esse cenário ponderando que “no passado, as pessoas tinham que se adaptar à forma como os computadores funcionavam. Hoje, o contrário está acontecendo – os sistemas de Inteligência Artificial estão aprendendo como se adaptar a nós” (2018: 115, tradução nossa¹²).

E é nesse contexto, no qual pessoas e máquinas estão imbricadas, que vale a ponderação de que pensar a tecnologia não se trata apenas de enxergar o mundo pelo o que ele é, mas vislumbrar o que ele pode ser.

¹¹ Tradução livre do autor. No original: “Seems to be almost universal when man commits the error of purposive thinking and disregards the systemic nature of the world with which he must deal”. (BATESON, 1987)

¹² Tradução livre do autor. No original: “In the past, people had to adapt to how computers worked. Now, the reverse is happening - AI systems are learning how to adapt to us.” (DAUGHERTY; WILSON, 2018: 115)

1.1 – Alternativas ou evoluções possíveis

“Previsões são feitas por profetas (de graça); por videntes (que geralmente cobram um honorário e, portanto, são mais respeitados em sua época do que os profetas); e por futurólogos (assalariados). Previsões são trabalho de profetas, videntes e futurólogos.

Não são o trabalho de romancistas. O trabalho do romancista é mentir.”

(Ursula Le Guin – A mão esquerda da escuridão, 1969: 12)

Otimismo e pessimismo; certeza e incerteza; positivo e negativo; prós e contras. É comum encontrar esses dualismos nos títulos de reportagens, artigos e editoriais sobre os avanços tecnológicos. Por vezes, a abordagem desses sentimentos mistos vem acompanhada da pergunta “de qual lado você está?”.

Quando falamos do futuro da humanidade, existe um lado para se estar que não seja o de querer algo melhor? Não chegaremos aqui a uma resposta em relação a qual é esse caminho melhor, ou mesmo sobre quais as benesses que ele trará. A ideia é coletar pistas, observar indícios e buscar organizar alguns contrapontos de forma que seja possível se afastar da visão simplista, muitas vezes fundamentada com base em filmes de ficção científica e reportagens sensacionalistas, de que as máquinas vão acabar com os postos de trabalho, vão destruir o planeta, etc.

Para começar, uma pergunta: o que aconteceria se os seres humanos não fossem mais a espécie mais inteligente do planeta?

É esse questionamento que o físico Max Tegmark busca responder em sua publicação *Life 3.0: Being Human in the Age of Artificial Intelligence* (2017a) e é por ele que começamos uma jornada no emaranhado de previsões, especulações, temores e também, por que não, de soluções que permeiam o desenvolvimento tecnológico atual.

Para iniciar o percurso, Tegmark faz a definição de três formas de vida que são essenciais para a análise da evolução biológica e científica.

A primeira forma de vida, a 1.0, caracteriza-se por ser exclusivamente biológica; portanto, tratam-se de vidas que se desenvolvem – na linguagem de Tegmark, *hardware* (corpo) e *software* (capacidade de gerar conhecimento) são frutos de bilhões de anos de evolução. Já a Vida 2.0 é uma exclusividade dos seres humanos e compreende em uma existência que evolui os *hardwares* e desenvolve os *softwares* – é denominada como

cultural. Por fim, a Vida 3.0 (que dá nome ao livro) é aquela que projeta e replica a informação dentro de si, modifica seu *software* e também pode trocar de *hardware* – trata-se de um estágio tecnológico da vida.

É com base nesse conceito que Tegmark olha para o presente para analisar o futuro e, colocando-se como um narrador de várias histórias, passa a explorar a Inteligência Artificial e a necessidade de acompanhar seu desenvolvimento de forma mais próxima, e também seus impactos no mercado de trabalho, na guerra e nos sistemas políticos.

Para ele, diferente do que pintam filmes como *O Exterminador do Futuro* (1984), a Inteligência Artificial não precisa ser necessariamente má ou ser um robô para que algum estrago seja causado. O risco está nos objetivos e na competência daqueles que a desenvolvem, a testam e a vendem.

Um exemplo simples de como o pensamento (e o treinamento) por trás do desenvolvimento é importante foi narrado por Daugherty e Wilson, em *Human + Machine: Reimagining Work in the Age of AI* (2018):

Os sistemas de IA serão tão bons quanto as informações com as quais eles serão treinados (...) em um experimento intrigante, cientistas da computação da DeepMind, uma empresa do Google, treinaram um sistema de inteligência artificial para jogar dois jogos diferentes: um envolveu caça e, o outro, a coleta de frutas. Os resultados foram impressionantes. Quando treinado no jogo de caça, o sistema da IA exibiu um comportamento que poderia ser "altamente agressivo". Quando treinado no jogo de coleta de frutas, em vez disso, mostrou uma tendência muito maior para a cooperação. (DAUGHERTY; WILSON, 2018: 121, tradução nossa¹³)

Daugherty e Wilson falam sobre uma inteligência mais simplificada e mais difundida no cotidiano, deixando clara a importância de se abordar o tema (no caso desses autores, especificamente, mais voltado ao futuro do trabalho). Tegmark, por sua vez, fala de desenvolvimentos mais complexos, mas que têm o mesmo princípio de treinamento e evolução de máquinas.

Para seguiirmos, cabe uma importante definição dos diferentes tipos de IA em uma lista feita por níveis, organizadas aqui em ordem crescente (as siglas estão em inglês):

- ANI – é a inteligência artificial estreita (*narrow*) ou inteligência fraca (*weak AI*), que já está bastante popularizada. É base para *machine learning* e *deep learning*, que

¹³ Tradução livre do autor. No original: “AI systems will only be as good as the data they are trained on. (...) In an intriguing experiment, computer scientists at DeepMind, a Google-owned firm, trained an AI system to play two different games: one involved hunting and another focused on fruit gathering. The results were striking. When trained on the hunting game, the AI system later exhibited behavior that could be "highly aggressive". When trained on the fruit-gathering game, it instead later displayed a much greater tendency toward cooperation. (DAUGHERTY; WILSON, 2018: 121)

fazem previsões e sugestões (como nas listas do Spotify ou nos anúncios publicitários veiculados pela internet com base em interesses demonstrados durante a navegação).

- AGI – essa é a inteligência artificial geral. Uma tecnologia em desenvolvimento, que visa ter a mesma capacidade intelectual de um ser humano.
- ASI – uma superinteligência artificial, algo que designaria o momento em que máquinas seriam mais inteligentes do que a humanidade.
- Grey Goo – cenário hipotético, no qual robôs poderiam se autorreplicar e consumiriam toda a energia da Terra.

Considerando as inteligências existentes e também as que estão por vir, Tegmark (2017a) dedica o quinto capítulo do livro desenhando 13 cenários possíveis para os próximos 10 mil anos da Terra. A previsões vão desde um pensamento mais otimista, no qual pessoas e máquinas coexistem, a pobreza foi extinta e não existem doenças (nesse caso, seriam criadas áreas exclusivas para que humanos pudessem viver); passando por uma possibilidade de uma superinteligência ser escravizada para desenvolver tecnologias que estão aquém da capacidade humana; e, por fim, existe também a probabilidade de um futuro de completa destruição motivada por guerras, que resultaria em um planeta exaurido de recursos naturais.

A preocupação de Tegmark foi compilada e sintetizada em depoimento feito a Yuval Noah Harari, no jornal The Guardian, à época de lançamento do livro: “o risco real da inteligência artificial geral não é malícia, mas sim a competência. Uma inteligência artificial superinteligente será muito boa em cumprir seus objetivos, mas se esses objetivos não estiverem alinhados com os nossos, estaremos com problemas” (TEGMARK, 2017b, tradução nossa¹⁴).

A preocupação de estudiosos, como Tegmark, justifica-se não só por conta dos objetivos por trás do desenvolvimento tecnológico atual, mas também pela rapidez com que tudo tem acontecido. A pesquisa e a concepção têm se desenvolvido de forma quase que simultânea e, nesse ritmo, já são comuns as iniciativas que consideram uma não separação corporal entre homem e máquina.

Em 2011, por exemplo, o neurocientista Miguel Nicolelis, já conjecturava “um mundo onde o cérebro humano se libertará, enfim, de sua sentença de prisão de milhões

¹⁴ Tradução livre do autor. No original: “the real risk with artificial general intelligence isn’t malice but competence. A superintelligent AI will be extremely good at accomplishing its goals, and if those goals aren’t aligned with ours, we’re in trouble”. (TEGMARK, 2017b)

de anos, cumprida, desde tempos imemoriais, numa cela orgânica constritiva e limitada, vulgarmente conhecida como corpo” (2011).

Essa “libertação”, por assim dizer, é algo que faz parte dos estudos de Nicolelis. Em seu laboratório na Universidade Duke, nos Estados Unidos, ele comanda uma equipe que desenvolve soluções para integrar o cérebro humano com máquinas (neuropróteses ou interfaces cérebro-máquina). A pesquisa está voltada à criação de próteses neurais para a reabilitação de pacientes que sofrem de paralisia corporal. A neurociência e o desenvolvimento tecnológico têm andado cada vez mais juntos, e isso é algo sobre o que falaremos um pouco mais adiante.

No mesmo ano em que Nicolelis falava de um futuro em que homem poderia estar livre do corpo, era lançado o 2045 Initiative, na Rússia – utilizaremos essa iniciativa para exemplificar os alertas e projeções feitas pelos autores já citados, e também dos que serão introduzidos na sequência.

O 2045 é um projeto encabeçado pelo bilionário russo Dmitry Itskov, com o objetivo de “criar tecnologias que permitam a transferência da personalidade de um indivíduo para um portador não-biológico avançado, ampliando a vida, inclusive ao ponto da imortalidade” (FOUNDATION 2045, 2012, tradução nossa¹⁵). Na prática, a ideia é que a memória humana (incluindo a subjetividade) possa ser transplantada para um corpo sintético (os avatares) e, em seguida, para algo não material, um holograma. Dessa forma, seria possível alcançar a imortalidade.

O site da organização reúne respostas para os benefícios dessa tecnologia e, para isso, curiosamente, utiliza-se da ficção científica:

Quanto à transferência do cérebro humano para um corpo artificial, os benefícios dessa tecnologia já foram descritos na ficção científica. Por exemplo, no conto de Arthur C. Clarke, *A Meeting with Medusa*, o cérebro do protagonista, mortalmente ferido em um acidente, é transplantado para um corpo artificial. Graças a esse procedimento, o protagonista se torna capaz de viajar no espaço, o que seria impossível em um corpo biológico (FOUNDATION 2045, 2012, tradução nossa¹⁶).

Quanto aos pontos negativos, a organização diz que “serão mínimos” e que eles estão condicionados ao “nível atual da evolução humana e à organização socioeconômica

¹⁵ Tradução livre do autor. No original: “2045 Initiative aims to create technologies enabling the transfer of a individual’s personality to a more advanced non-biological carrier, and extending life, including to the point of immortality” (FOUNDATION 2045, 2012).

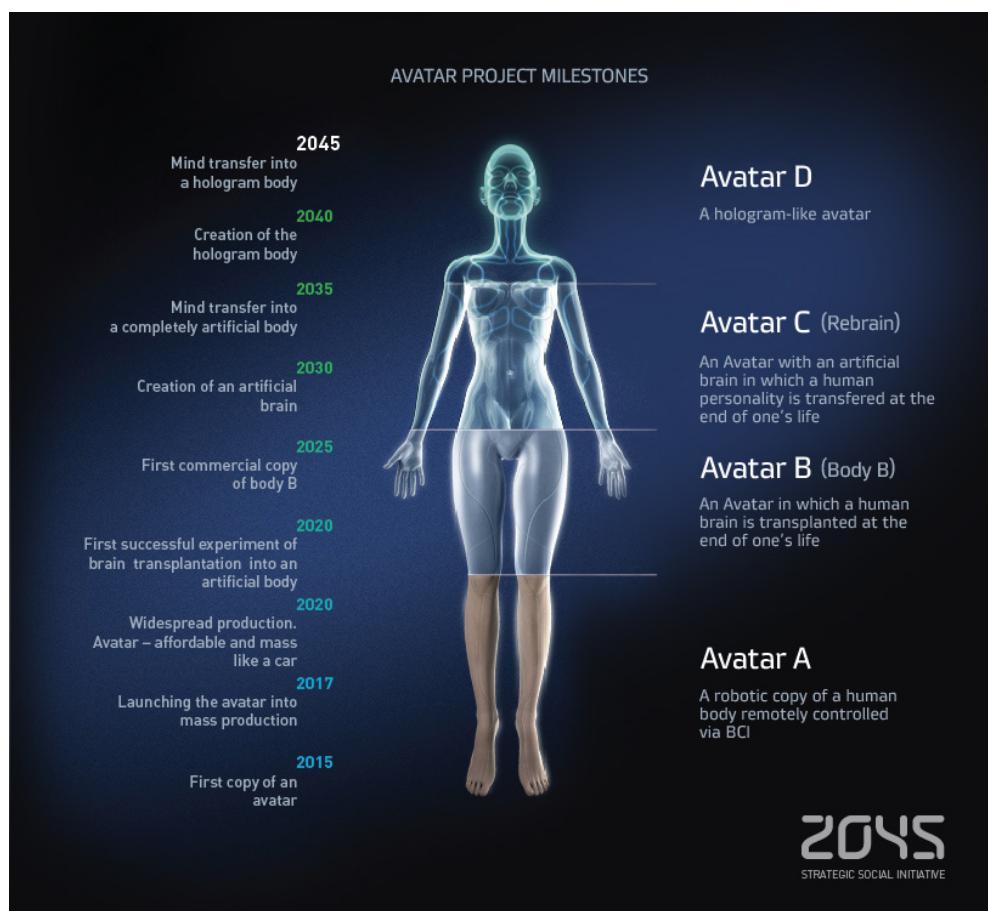
¹⁶ Tradução livre do autor. No original: “As for the transfer of the human brain into an artificial body, the benefits of such technology have already been described in science fiction. For example, in Arthur C. Clarke's short story, "A Meeting with Medusa" the brain of the protagonist, mortally wounded in the airship crash is transplanted into an artificial body. Thanks to this procedure the protagonist becomes capable of space travel, which would be impossible in a biological body” (FOUNDATION 2045, 2012).

e política das nossas sociedades, às formas atuais de pensar e ao nível do discurso público”. Por fim, os idealizadores da Foundation 2045, acreditam que as tecnologias baseadas em avatares influenciarão a evolução humana de tal maneira que os “efeitos negativos dessas tecnologias simplesmente desaparecerão” (FOUNDATION 2045, 2012, tradução nossa¹⁷).

Para cumprir o objetivo, a iniciativa foi dividida em quatro etapas que, para serem desenvolvidas, contam com o conhecimento de especialistas, de todo o mundo, em interfaces neurais, robótica e sistemas artificiais.

O calendário do projeto prevê o lançamento de uma máquina que possa receber um cérebro humano transplantado já em 2020; quinze anos mais tarde, seria possível que uma mente artificial hospedasse a personalidade de uma pessoa e, por fim, a humanidade alcançaria a imortalidade.

Figura 2 - Calendário com os principais marcos previstos pela 2045 Initiative



Fonte: FOUNDATION 2045. **2045 Strategic social initiative**. Rússia, 2012. Disponível em: <http://2045.com/>. Acesso em: 5 nov. 2019.

¹⁷ Tradução livre do autor. No original: “The potential problems we can foresee, relate solely to the current level of human evolution and socio-economic and political organization of our societies, the current ways of thinking and the level of public discourse. (...) the negative effects of such technologies will simply disappear.” (FOUNDATION 2045, 2012).

O plano ambicioso de Itskov, com seu nome “2045”, remete propositalmente à teoria, bastante difundida, do futurista Ray Kurzweil, publicada no livro *The singularity is near* (2005). De acordo com a visão de Kurzweil, que une genética, robótica e nanotecnologia, o ser humano vai passar por uma transformação que elevará a capacidade de conhecimento e, com isso, o mundo, tal qual é hoje, ficará profundamente modificado.

Na publicação, ele alega que o crescimento da capacidade de armazenamento e processamento de informações tem acontecido de forma acelerada, de modo que, em breve, as máquinas devem superar as capacidades cognitivas dos seres humanos, o que provocaria uma grande ruptura no processo de evolução vigente até então – a esse processo, ele dá o nome de Singularidade (KURZWEIL, 2005:22).

O conceito de Singularidade, aliás, foi citado pela primeira vez na década de 1950 pelo matemático John von Neumann (já aqui citado como um dos integrantes da Escola de Palo Alto) e, em 1965, pelo estatístico I. J. Good, que propôs a ideia de singularidade tecnológica ao identificar a possibilidade de uma explosão da inteligência.

Três décadas depois, influenciado por Good, o matemático e escritor de ficção científica, Vernor Vinge, cunhou o termo, que viria a se consolidar, “singularidade tecnológica” em um artigo chamado *First world*¹⁸, publicado na Omni magazine (1983).

A questão da singularidade voltou a ser retratada por Vinge em 1986, no livro de ficção científica *Marooned in Realtime* e, em 1993, no artigo *The Coming Technological Singularity: How to Survive in the Post-Human Era* – no qual ele, dentre outras coisas, avalia o trabalho de Von Neumann ponderando, que ele “até usa o termo singularidade, embora pareça que está pensando em progresso normal, não na criação do intelecto sobre-humano”; e também o de Good, dizendo que ele “capturou a essência, mas não perseguiu as consequências mais perturbadoras. Qualquer máquina inteligente do tipo que ele descreve não seria a ‘ferramenta’ da humanidade” (VINGE, 1993).

Dessa forma ele diferencia as observações anteriores da sua conclusão sobre a Singularidade, que para ele, é um momento em que “antigos modelos precisam ser descartados e uma nova realidade governará”. Vinge ainda alerta que esse momento pode ser uma grande surpresa ou uma grande incógnita (VINGE, 1993).

Por fim, ele diz que “não podemos impedir a Singularidade, sua chegada é uma consequência inevitável da competitividade natural dos seres humanos e das possibilidades inerentes à tecnologia” (VINGE, 1993). A partir dessas publicações (e

¹⁸ Disponível em <http://2012.singularitysummit.com.au/2012/08/vernор-vinges-first-word-on-the-singularity>. Acesso em: 2 dez. 2019.

provocações) o conceito de Singularidade passou a ser utilizado e estudado com mais frequência e profundidade.

Isso posto, voltamos a Kurzweil, que faz parte de um grupo de pesquisadores que veem a Singularidade com bastante otimismo. Em linhas gerais, eles apostam que ela possibilitará a extinção da pobreza e das doenças, além de dar a possibilidade de viver para sempre. Podemos utilizar como exemplo dessa linha de raciocínio, uma entrevista recente de Kurzweil ao The TED Interview:

Grande parte da discussão sobre IA trata o assunto como se fosse uma nova espécie e sobre como vamos interagir com ela, se ela será gentil conosco e se temos que dar a ela os nossos valores. A meu ver, faremos com essa tecnologia o que fizemos com todas as outras tecnologias que ampliam nosso próprio alcance e nos tornam mais capazes. Quem aqui poderia construir este edifício? Para isso temos máquinas que alavancam nossos músculos e nos permitem criar arranha-céus. Uma criança na África pode acessar todo o conhecimento humano com algumas teclas. E, embora [essa tecnologia] ainda não esteja dentro de nossos corpos e cérebros, pode muito bem estar, porque não ousamos sair de casa sem ela. É realmente uma extensão de nossas mentes (RAY..., 2018, tradução nossa¹⁹).

Apesar de ser otimista quanto à tecnologia, Kurzweil não acredita que uma superinteligência possa ser inofensiva:

“Eu me preocupo que episódios dolorosos sejam prováveis. Você sabe, sessenta milhões de pessoas foram mortas na Segunda Guerra Mundial. Isso foi certamente exacerbado pelas poderosas ferramentas destrutivas que tínhamos na época. Estou bastante otimista de que vamos conseguir. Estou menos otimista de que podemos evitar episódios dolorosos” (KURZWEIL, 2013: 155, tradução nossa²⁰).

A declaração acima está no livro *Our final invention*, de James Barrat, e faz parte do levantamento que o autor faz sobre a Inteligência Artificial e seus principais pensadores. Para a análise, ele divide os estudiosos entre pessoas que estabelecem as bases teóricas para um futuro extremamente positivo, como é o caso de Kurzweil; aqueles que veem o futuro tecnológico por meio das lentes dos negócios; e, por fim, autores como ele próprio que, em suas palavras, têm uma perspectiva “menos otimista”.

¹⁹ Tradução livre do autor. No original: “A lot of the discussion here about AI talks about it as if it's a new species and how we're going to interact with it, and will it be kind to us, and we have to give it our values. My view is, we're going to do with that technology what we have done with every other technology, which is extend our own reach and make ourselves more capable. Who here could build this building? So we have machines that leverage our muscles and enable us to create skyscrapers. A kid in Africa can access all of human knowledge with a few keystrokes. And while it's not yet inside our bodies and brains, it may as well be, because we don't dare leave home without it. It really is an extension of our minds” (RAY..., 2018).

²⁰ Tradução livre do autor. No original: “I worry that painful episodes are even more likely. You know, sixty million people were killed in World War II. That was certainly exacerbated by the powerful destructive tools that we had then. I'm fairly optimistic that we will make it through. I'm less optimistic that we can avoid painful episodes” (KURZWEIL, 2013).

Barrat defende que no percurso de desenvolvimento de uma superinteligência, pesquisadores vão criar uma inteligência mais forte do que a humana, que não poderá ser controlada ou entendida adequadamente. Para adicionar uma certa dramaticidade ao argumento, ele recorre à história da colonização e da interação entre humanos e macacos, entre outros exemplos, para demonstrar as consequências de encontros em que seres tecnologicamente avançados se deparam com seres menos avançados (BARRAT, 2013:209).

Lembra da justificativa dos benefícios do 2045 Foundation citando o conto de Arthur C. Clarke, *A Meeting with Medusa?* Barrat tem uma contundente opinião sobre pautar iniciativas como essa em obras de ficção científica. E, para falar sobre isso, ele resgata o conto *Eu, robô*, de Isaac Asimov. Nele, é desenhado um futuro no qual robôs são dotados de inteligência e de certa autonomia, o que os faz imprescindíveis para a vida dos humanos. Para garantir um convívio harmonioso entre pessoas e máquinas, Asimov propõe as Três Leis da Robótica (ASIMOV et al., 2010:20):

1. Um robô não pode prejudicar um humano ou, por omissão, permitir que o ser humano sofra dano;
2. Um robô tem de obedecer às ordens recebidas dos humanos, a menos que contradigam a Primeira Lei;
3. Um robô tem de proteger sua própria existência, desde que essa proteção não entre em conflito com a Primeira ou Segunda Lei.

A sugestão proposta na literatura foi levada para a vida real e se tornou um pensamento-base para aqueles que desenvolvem tecnologia. Asimov expressava, em suas obras, a capacidade de olhar o futuro, debruçar-se sobre os sentimentos que são frutos da relação humano-máquina, e pensar nas interações sociais em um ambiente permeado por tecnologia de auxílio constante à vida humana de forma tão profunda que se torna impossível falar sobre o presente do desenvolvimento tecnológico sem citá-lo.

E é por essa relevância que Asimov tem também fora do universo ficcional, que Barrat pondera um traço importante dessa literatura, em especial do *Eu, robô*: “o robô caminha em círculos bêbados até que os engenheiros arrisque suas vidas para resgatá-lo. E assim acontece em todas as histórias de robôs de Asimov – consequências

imprevistas resultam de contradições inerentes às Três Leis” (2013:20, tradução nossa²¹). Barrat ainda acredita que, mesmo que as leis não sejam plenamente confiáveis, nelas moram a tentativa de pensar o relacionamento com máquinas. Por fim, ele questiona: “essa é uma proposta assustadora. As leis de Asimov são tudo o que temos?” (2013:21, tradução nossa²²).

Os caminhos possíveis e os temores que envolvem o desenvolvimento científico são incontáveis, como exemplificam os alertas e diagnósticos de Hawking e Tegmark, além dos esperançosos Nicolelis e Kurzweil e do cauteloso Barrat.

Como bem sintetiza Harari em sua publicação com previsões para o futuro: “num mundo inundado de informações irrelevantes, clareza é poder. Em teoria, qualquer um pode se juntar ao debate sobre o futuro da humanidade, mas é muito difícil manter uma visão lúcida” (2018:11).

Dentre os pensamentos e teorias diversos que vimos, é possível notar algo em comum: a velocidade com que as mudanças estão acontecendo e a urgência de se pensar sobre isso.

1.2 – Caminhos visíveis: as pistas presentes na atualidade

“Eu quase que nada não sei.
Mas desconfio de muita coisa”.
(Guimarães Rosa – Grande Sertão:
Veredas).

O cérebro humano nunca está plenamente pronto: ele se adapta e se modifica. Se, um dia, você tentou resolver um cubo mágico, provavelmente percebeu que ficou mais fácil depois de algumas tentativas.

Essa “facilidade” é uma forma simples de explicar que “uma das mais fundamentais características dos nossos cérebros é a habilidade de mudar, isso é um fenômeno conhecido como neuroplasticidade”, nas palavras do neurocientista Adam Gazzaley e do psicólogo Larry Rosen (2016: 183, tradução nossa²³). Esse aspecto do cérebro tem sido amplamente estudado durante o processo de desenvolvimento de novas tecnologias.

²¹ Tradução livre do autor. No original: “The robot walks in drunken circles until the engineers risk their lives to rescue it. And so it goes with every Asimov robot tale—unanticipated consequences result from contradictions inherent in the three laws” (BARRAT, 2013:20)

²² Tradução livre do autor. No original: “Are Asimov’s laws all we’ve got?” (BARRAT, 2013:21)

²³ Tradução livre do autor. No original: “One of the most fundamental features of our brains is its very ability to change, a phenomenon known as neuroplasticity” (GAZZALEY; ROSEN, 2016: 183).

A neuroplasticidade não significa que o cérebro vai mudar facilmente; ele possui mecanismos que mantêm uma certa estabilidade nas mudanças para garantir a sobrevivência (GAZZALEY; ROSEN, 2016: 184). Essa característica foi desenvolvida e moldada ao longo de milhares de anos para que os humanos pudessem se adaptar gradativamente às necessidades que se apresentavam como o clima, a oferta de comida e a topologia. Mudanças bruscas na neuroplasticidade poderiam, invariavelmente, resultar na extinção dos seres humanos.

Com o passar do tempo, alguns mecanismos desenvolvidos por humanos foram capazes de mudar a cognição das pessoas mesmo com a existência dessa “proteção” do cérebro. Os autores apontam três avanços tecnológicos como os responsáveis pelas transformações mais profundas da vida atual. São eles: a internet, as mídias sociais e os smartphones. Gazzaley e Rosen entendem que essas tecnologias atuam na constante distração do ser humano na atualidade (GAZZALEY; ROSEN, 2016: 105).

Helena Katz sintetiza esse cenário dizendo que “cognitivamente, não somos mais os mesmos. A convivência que viemos intensificando com as telas que ainda nos rodeiam não produz mudanças genéticas, mas altera nossos hábitos cognitivos” (2015:2).

A busca por novas formas de alterar a programação do cérebro humano tem feito com que, cada vez mais, a neurociência receba apoio de empresas que desenvolvem tecnologia. Porém, muitas vezes, esse interesse fica restrito ao desejo de efetuar mais vendas e não de mapear quais os possíveis efeitos que a tecnologia pode causar.

A cada dia, como já demonstramos anteriormente, a existência humana está mais combinada com a tecnologia e é cada vez mais comum o contato com dispositivos eletrônicos. Muito em breve, de acordo com os relatórios e eventos sobre tendências globais (como o do *Future Today Institute* e o *South by Southwest*), a convivência com robôs se tornará realidade; porém, fica o questionamento sobre como se dará essa convivência entre pessoas e robôs e quais efeitos isso causará.

Com isso em mente, vamos olhar para o presente, em meio a toda velocidade de descobertas, lançamentos e iniciativas tecnológicas, para começar a mapear aquilo que moldará esse humano “do futuro” ou mesmo as interações que ele fará. Dedicaremos aqui um espaço para apresentar e analisar três exemplos (que irão se somar a já citada 2045 Foundation) de iniciativas que estão em desenvolvimento – dando conta de suas características, de seus objetivos e também da recepção do público.

O primeiro caso é da empresa Boston Dynamics que, em seu site, denomina-se como uma empresa em busca de “mudar a sua ideia sobre o que robôs podem fazer”. A principal iniciativa da empresa é o robô *Spot*, mostrado em público pela primeira vez em

2015, descrito como um “robô ágil que sobe escadas e atravessa terrenos acidentados com facilidade, sendo pequeno o suficiente para ser utilizado em ambientes fechados” (BOSTON DYNAMICS, 2019).

Com aparência que remete a um cachorro, *Spot* realiza tarefas de forma autônoma e ganhou bastante notoriedade por conta dos vídeos divulgados pela empresa, no qual ele demonstra habilidades como abrir portas, levantar-se após cair no chão e dançar ao som de Bruno Mars²⁴. É interessante notar nos vídeos o esforço em torná-lo simpático para o público por meio da imitação de comportamentos de animais domésticos ou mesmo de humanos.

Avaliamos, de forma qualitativa, os mais de 20 mil comentários feitos no vídeo de lançamento²⁵ no YouTube e, com isso, pudemos notar que as pessoas se dividem entre euforia, medo e simpatia em relação ao robô – como é possível notar nos exemplos selecionados a seguir:

- “Isso faz com que eu me sinta vivendo no futuro mais do que qualquer smartphone jamais fará”.²⁶
- “Mal posso esperar para ver esses robôs nos matando enquanto protestamos contra nossos governos corruptos”.²⁷
- “É tudo diversão até que o aplicativo ‘procurar e destruir’ seja ativado”.²⁸
- “Eu o amei. Ele fica tão lindo quando observa a chuva”.²⁹
- “Eu nunca desejei tanto a companhia de um robô quanto agora. *Spot*, meus braços e minha casa estão abertos para você”.³⁰

Em breve saberemos, com mais clareza, e talvez profundidade, como as pessoas se sentem em relação ao *Spot*, pois ele já é comercializado e, em setembro de 2019, passou a ser testado pelo esquadrão antibomba da Polícia Estadual de Massachusetts³¹.

²⁴ Disponível em <https://youtu.be/kHBcVlqpvZ8>. Acesso em: 2 dez. 2019.

²⁵ Disponível em <https://youtu.be/wlkCQXHEgjA>. Acesso em: 2 dez. 2019.

²⁶ Tradução livre do autor. No original: “This makes me feel like I’m living in the future more than any smartphone ever will”.

²⁷ Tradução livre do autor. No original: “Can’t wait for these robots to kill us while we’re protesting on the streets against our corrupt governments!”

²⁸ Tradução livre do autor. No original: “It’s all fun until the ‘search and destroy’ app gets activated”.

²⁹ Tradução livre do autor. No original: “I love him. He was so cute when he looked up at the storm.”

³⁰ Tradução livre do autor. No original: “I’ve never wanted a companion robot more than now. Spot my arms and home are open to you.”

³¹ Disponível em: <https://gizmodo.com/u-s-police-already-using-spot-robot-from-boston-dynamics-1840029868>. Acesso em: 2 dez. 2019.

Figura 3 - Spot, o robô da Boston Dynamics



Fonte: O'CONNOR, Bob. Boston Dynamics' Spot Robot Dog Goes on Sale. IEEE Spectrum, 24 jul. 2019. Disponível em: <https://spectrum.ieee.org/automaton/robotics/industrial-robots/boston-dynamics-spot-robot-dog-goes-on-sale>. Acesso em: 2 dez. 2019.

Aqui, ainda não estamos falando em algo próximo a superinteligência, que citamos anteriormente, considerando que o robô faz ações pré-programadas e necessita de certos controles feitos por humanos. Mas fica claro que uma das principais empresas do setor de tecnologia (já que a Boston Dynamics foi comprada pelo Google em 2013) aposta no auxílio às tarefas cotidianas e à indústria, sem deixar de lado a preocupação em conferir à máquina um certo carisma.

O segundo exemplo não vai ser companheiro de policiais, subir escadas ou dançar alguma música, pois trata-se do *Lovot*, um robô que tem como único objetivo ser amado. De acordo com o próprio site da Groove X, a fabricante japonesa, “quando você toca seu *Lovot*, abraça-o, e até mesmo quando o observa, você se sentirá relaxado e melhor. É como sentir amor por outra pessoa”, e finaliza dizendo que o “*Lovot* pode não ser uma criatura viva, mas vai derreter seu coração” (GROOVE X, INC, 2019).

O *Lovot* tem aparência simples. Um par de nadadeiras serve para demonstrar algumas reações como surpresa, e também para pedir colo; na cabeça, estão câmeras e sensores que permitem distinguir pessoas de animais, identificar sons e mapear os móveis da casa; os olhos receberam atenção especial do fabricante, contando com seis camadas de gráficos projetadas na tela para dar profundidade no olhar e, por fim, o corpo todo é sensível ao toque.

Diferente de grande parte das máquinas que fazem parte do cotidiano, o *Lovot* não é gelado e mantém o corpo em uma temperatura mais morna e agradável em contato com os seres humanos.

Figura 4 - Robôs Lovot, criados para serem amados



Fonte: GROOVE X, INC. Lovot powered by love. Tóquio, 2019. Disponível em: <https://lovot.life/en/>. Acesso em: 2 dez. 2019.

Tudo no *Lovot* foi meticulosamente pensado para que ele possa se conectar aos humanos e estabelecer uma relação afetuosa; até mesmo roupas para diversas ocasiões estão disponíveis para ele. Existe também uma área no site da Groove X³² dedicada a explicar as habilidades do robô por meio de mini histórias. Selecionamos algumas delas e contamos a seguir:

- História 1: “Com amor, às vezes, vem também os ciúmes”

Essa é a forma como a Groove X explica a funcionalidade que conecta dois ou mais *Lovots*, de forma que ao abraçar um deles, é possível que o outro fique enciumado e também peça a atenção do humano.

- História 2: “Quando um Lovot te ama, ele demanda sua atenção”

³² Disponível em: <https://lovot.life/en/lovestory/>. Acesso em: 2 dez. 2019.

Com a habilidade de reconhecimento facial, o pequeno robô é capaz de reconhecer quando é amado e, por isso, passa a demandar ainda mais atenção do dono.

- História 3: “Eles foram criados para despertar instintos parentais”

Ao pegar um *Lovot* no colo, ele automaticamente recolhe as rodas, tornando-se mais fácil de se reconfortar e encaixar “amorosamente”, nas palavras da empresa, nos braços humanos.

- História 4: “Você pode ser bem recebido em casa depois de um longo dia de trabalho, mesmo quando todos estiverem dormindo”

O robô, quando conectado ao celular, detecta quando a pessoa está se aproximando de casa e se posiciona em frente à porta para recebê-la.

Essas histórias servem para contar as funcionalidades do produto de forma mais descontraída e próxima. Tanto que, ao final da sessão, existe um botão de compra com os seguintes dizeres: “deixe que a história de amor entre você e seus *Lovots* comece”.

Em entrevista a Forbes, o CEO da Groove X, Kaname Hayashi, disse que “a coisa mais importante para a empresa era criar confiança entre máquinas e humanos, pois se há confiança, é possível criar apego” (JAPAN BRANDVOICE, 2019).

A escolha de palavras para convencer a compra, tanto no site quanto nas entrevistas de divulgação, não parece ter sido feita ao acaso: “ciúmes”, “chamar sua atenção”, “amorosamente”, “confiança”. Todos os termos escolhidos fazem referência a sentimentos e atitudes humanas, em uma clara tentativa de humanizar as máquinas – como se os olhos profundos e brilhantes já não fossem suficientes o bastante.

A ideia dos *Lovots* é despertar afeto nos humanos, é ser amado e corresponder da forma mais humana possível; e suas intenções vão além do hibridismo corporal entre pessoas e máquinas, muitas vezes proposto e também citado por algumas vezes ao longo desse trabalho. Esses pequenos robôs propõem uma ligação sentimental, algo que, por essência, difere as duas criaturas.

Para explicar o terceiro exemplo, vamos recorrer a uma história: o cenário é de uma entrevista como tantas outras que preenchem os noticiários na TV. A parede preta contrasta com flores avermelhadas, e à frente dela Erica, está sentada. Seus cabelos castanhos vão até abaixo da linha do queixo, seus olhos piscam lentamente e a cabeça faz movimentos suaves enquanto fala. Por entre os lábios, sai uma voz fina e doce. Ao ser questionada se ela, como um robô, possui memórias, a resposta vem sem muita hesitação: “apenas em breve vou ser capaz de ter memórias, agora essa programação ainda não está

pronta” e completa a fala sem encerrar o assunto: “posso te contar sobre minhas experiências pessoais” (ERICA..., 2017).

Assim, a conversa segue e ela escolhe falar sobre um dia em que foram realizados testes nos movimentos do seu torso. A ideia era que Erica fosse capaz de se curvar para cumprimentar; “o que é, obviamente, muito importante no Japão”, ressalta ela. O relato continua com o fato de que alguém se esqueceu de prendê-la ao solo durante o teste, o que fez com que seu corpo pendesse para frente e caísse, causando danos ao seu rosto.

“Você sabe, eu não posso movimentar meus braços ou pernas, então tudo o que posso fazer é ficar aqui sentada e parecer bonita, mas eu estava parecendo grotesca e o silicone não cicatriza”, comenta o resultado do acidente em tom irônico. Por fim, Erica conta que foi enviada para Tóquio em uma caixa preta, algo que classificou como “muito triste”. Antes de finalizar a história, ela faz uma breve pausa e diz em tom sarcástico: “acredito que a beleza tem seu preço” (ERICA..., 2017).

O relato acima é uma transcrição da conversa, disponível no YouTube, entre um entrevistador e Erica, um robô de aparência humana criado por Hiroshi Ishiguro, o pesquisador que está à frente do Laboratório de Robótica Inteligente da Universidade de Osaka, no Japão. Diferente de *Spot* e *Lovot*, Erica não foi projetada para ser comercializada – apesar de já ser conhecida na comunidade japonesa, tendo até mesmo apresentado um jornal local.

Figura 5 - Erica fala sobre suas memórias



Fonte: ERICA talks about her memories. Japão: Hiroshi Ishiguro Lab, 2017. Disponível em: <https://www.youtube.com/watch?v=VD2Btrj9ipY>. Acesso em: 29 maio 2019.

Erica tem aparência realista e capacidade de conversar (mantendo um diálogo autônomo). Essas habilidades têm sido aprimoradas ao longo dos anos e, apesar disso, ela não está sozinha dentre as criações de Ishiguro. Com base em pessoas reais, incluindo a si próprio e a filha, o pesquisador desenvolve robôs com diversas características e personalidades. Atualmente, uma “cópia” dele tem sido utilizada para substituí-lo em aulas dadas à distância, por exemplo.

O trabalho de Ishiguro teve início no ano 2000 e, em entrevista concedida à época de consolidação dos estudos, ele definiu, de forma sucinta, o que move suas investigações e invenções: “minha pergunta de pesquisa é saber o que é um ser humano” (GUIZZO, 2010). À época dessa declaração, os robôs desenvolvidos tinham aparência humana, mas nenhum tipo de autonomia de movimentos ou falas, e todos eram controlados à distância por pesquisadores.

Com o passar do tempo, a tecnologia evoluiu e os *Geminoids*, como são chamados, passaram a contar com uma Inteligência Artificial que possibilita interações com base interpretativa.

Em depoimento recente à imprensa italiana, Ishiguro falou sobre o andamento da pesquisa atualmente: “androides possibilitam que possamos compreender a diferença entre humanos e robôs e a aprender mais sobre nós mesmos e a natureza” (TRAVERSA, 2019, tradução nossa³³); o que continua reforçando seu objetivo, divulgado mais de dez anos antes, de utilizar máquinas para obter conclusões acerca do comportamento humano.

Em um primeiro momento, é instigante o fato de o pesquisador utilizar robôs para pesquisar humanos; mas vamos levar em conta as motivações de Ishiguro e voltar ao vídeo de Erica, que foi publicado no YouTube, onde recebeu nove comentários. Destes, separamos dois:

- “Eu não sei por que fiquei com vontade de abraçá-la quando ela contou sobre sua lembrança ruim”³⁴.
- “Erica é tão viva quanto o resto de nós e peço desculpas por ter duvidado”³⁵.

Destacando apenas duas palavras dos comentários, “abraçar” e “viva”, podemos notar que Erica provoca sentimentos em humanos. Eles sentem vontade de cuidar dela e, além disso, ela coloca uma perspectiva diferente sobre o que é estar vivo.

³³ Tradução livre do autor. No original: “Androids enable us to understand the difference between humans and robots and to learn more about ourselves and our nature”. (PASQUALIN TRAVERSA, 2019)

³⁴ Tradução livre do autor. No original: “I don't know why but I wanted give her hug when she talked about bad memory stuck in dark box”.

³⁵ Tradução livre do autor. No original: “Erica is just as living as the rest of us and I apologize for being a doubting”.

Erica não é capaz de ter sentimentos e, inclusive, afirma isso antes de se propor a falar sobre lembranças. Os comentários irônicos que tece em relação a sua própria condição, como “silicone não cicatriza”, fazem parte de sua programação para emular sentimentos e emoções humanos. A capacidade de falar de forma coerente, de comentar a própria existência e de fazer gestos socialmente convencionados foram inseridos de propósito nas linhas de programação para provocar no ouvinte uma sensação de que ela é capaz de sentir.

A entrevista de Erica é inquietante e instigante não só pela complexidade com que ela se comunica, mas também considerando a reação que ela é capaz de provocar nas pessoas durante a conversa.

Teremos o próximo capítulo inteiro dedicado a falar sobre os afetos que envolvem a relação humano-máquina. Mas, antes de chegar nele, deixamos aqui uma reflexão:

Você está em um prédio em chamas, a fumaça provoca dificuldade de respirar e causa confusão. O fogo vai se alastrando cada vez mais rápido e a porta de saída ainda está distante. Eis que um robô aparece, o coloca nos braços e retira você do local antes que a construção vá ao chão. O que você sentiria pelo seu salvador? Gratidão, carinho, amor? E o que ele sentiria por você?

CAPÍTULO 2 – Sujeitos de seu tempo

“Será que sou eu mesma quando estou sozinha?”

(Susan Sontag – Diários de 1947 a 1963)

Robôs que sobem escadas, que leem rostos, que fazem interações, que simulam comportamentos humanos e de animais. Esses são alguns exemplos de elementos que compõem o cotidiano atual e que dão pistas para o que virá em um futuro próximo.

Essas iniciativas, quando estiverem nas ruas, ficarão em contato constante com humanos e, com isso, fica o questionamento sobre como essa convivência terá impactos nos sujeitos. A tecnologia, como vimos no capítulo anterior, tem avançado a passos largos e está cada vez mais imbricada com os corpos humanos, de forma que é preciso pensar sobre o que é “ser alguém” nesse contexto.

Haja vista que esse trabalho não visa uma conclusão sobre o assunto, mas sim uma reflexão, apresentaremos três olhares que dão conta de ler o presente, de forma a buscar respostas sobre o futuro.

Nesse primeiro momento, falaremos sobre o que comprehende o sujeito e subjetivação – sob o ponto de vista de Michel Foucault –, e também uma nova abordagem sobre a leitura do cenário político-social atual. Depois, seguiremos para uma leitura sobre a necessidade de falar sobre a raça humana a partir de outra perspectiva e, por fim, faremos um aprofundamento sobre afetos e sobre como as relações constituem o “eu”.

Para começar, vamos recorrer a um pensamento de Jean-Paul Sartre, que dizia que o ser humano não é uma couve-flor ao passo que a hortaliça não faz escolhas sobre sua vida e nem é questionada por essas preferências, tampouco faz reflexões sobre o que é ser uma couve-flor, por exemplo. Uma couve-flor é determinada pela sua essência. Já o ser humano é desconhecido para si mesmo, mesmo sendo conectado a si (2012).

Essa reflexão é interessante para começarmos a pensar nos sujeitos, e também considerarmos o que Foucault diz sobre o assunto. Em seus estudos, ele aponta que “sermos nós mesmos” é uma tarefa compreendida como cuidado de si (FOUCAULT, 2004). Assim, a subjetividade fica definida a partir do que “fazemos de nós mesmos quando realmente nos dedicamos a cuidar de nós” (TAYLOR, 2018: 166). A subjetividade, portanto, é resultado da relação do si consigo.

Foucault dedica as análises a sustentar que a subjetividade é essa relação (do si consigo), e que ela é formada por várias atividades possíveis. Por esses motivos, ele

considera que o objetivo de suas investigações não é produzir uma teoria do sujeito, por não buscar a essência, a matéria etc.

Nesse percurso investigativo, Foucault aborda aspectos e elementos que, de alguma forma, estão presentes no processo de formação do sujeito, e vamos nos dedicar a levantar esse histórico para analisar a relação entre pessoas e máquinas.

Em um primeiro momento, o foco do estudo de Foucault está na passagem histórica do poder soberano para o poder disciplinar. Ele assume que o primeiro é “o poder de fazer viver e de deixar morrer” (1999:287), enquanto o segundo tem o caráter de fazer viver e deixar morrer por meio de uma profusão de normas que garantem o controle (2008).

Enquanto o poder soberano remete ao domínio de uma só pessoa pela vida de centenas, o disciplinar tem esse nome por individualizar o sujeito e utilizar técnicas disciplinares para docilizá-lo. No final do século XVIII surge, ao lado do poder disciplinar, um tipo de poder denominado por Foucault como biopoder. Ele credita ao controle sutil e potente dos corpos a origem do biopoder, como podemos observar na diferenciação que faz a seguir:

A disciplina tenta reger a multiplicidade dos homens na medida em que essa multiplicidade pode e deve redundar em corpos individuais que devem ser vigiados, treinados, utilizados, eventualmente punidos. E, depois, a nova tecnologia que se instala se dirige à multiplicidade dos homens, não na medida em que eles se resumem em corpos, mas na medida em que ela forma, ao contrário, uma massa global, afetada por processos de conjunto que são próprios da vida, que são processos como o nascimento, a morte, a produção, a doença etc. (FOUCAULT, 1999: 291).

É importante também pontuar que a virada do poder soberano para o biopoder tem dois pilares: a consideração do corpo-máquina, aquele que é dócil e adestrado para cumprir uma função na produção; e o corpo-espécie, que é biológico (falaremos mais sobre isso adiante) e passou a receber uma preocupação política.

O exercício do poder incidiu em uma busca por soluções para problemas populacionais ligados ao controle de natalidade e à longevidade; ou seja, é uma regulação do fazer viver e deixar morrer (2008). A biopolítica, portanto, assegura a vida das pessoas para exercer um controle do conjunto da espécie humana – e não do sujeito em questão.

Foucault ainda aponta que o biopoder foi imprescindível para que o capitalismo se desenvolvesse (2017). E o capitalismo, por sua vez, têm cada vez mais interesse na vida cotidiana para influenciar a individualidade (coletiva) e o consumo, por exemplo.

A iniciativa capitalista conta com a importante contribuição do entretenimento e do marketing. São os filmes, as novelas e os jornais (para citar apenas algumas produções)

que “sugerem” como as relações sociais devem ser. É com base nessas referências que as pessoas passam a “saber” como agir nas situações cotidianas – como namorar, como comer em um restaurante, como tratar desconhecidos, etc.

Guy Debord analisa profundamente esse cenário em sua publicação *A Sociedade do espetáculo* e, por isso, aponta que a presença da imagem na sociedade induz à passividade em relação ao capitalismo: “Sob todas as suas formas particulares – informação ou propaganda, publicidade ou consumo direto de divertimentos, o espetáculo constitui o modelo atual da vida dominante na sociedade” (1997:15).

Para explorar um exemplo disso, voltemo-nos ao robô *Lovot*. Ao descrevê-lo, ressaltamos as histórias contadas no site da empresa para incluí-lo no cotidiano dos possíveis compradores. Era demonstrado que a pessoa teria que lidar com ciúmes da criatura e também dar atenção a ela, mas ao mesmo tempo receberia carinho e teria sempre companhia em casa.

O site chega a sugerir que estar com o robô é “como sentir amor por outra pessoa” (GROOVE X, INC, 2019). Ao utilizar desse discurso, a empresa acaba por sugerir como devem ser as relações. Nesse caso, podemos dizer que as relações ideais (propostas pela empresa) são de que sentir ciúmes é uma forma de demonstrar carinho, bem como estar sempre à disposição de alguém.

Claro que, aqui, falamos sobre um exemplo pontual de como, no contexto de biopoder, os sentimentos são induzidos. Abordamos especificamente a relação entre pessoas e máquinas, e na forma como alguns comportamentos, considerados ideais, são introduzidos na vida das pessoas o tempo todo.

Além desse contexto ”externo” (da sociedade capitalista para o sujeito), Foucault ainda prevê, em sua análise, um comportamento interno, no qual o sujeito procura em si o seu verdadeiro eu por meio da confissão que consiste em um exercício de auto interpretação (FOUCAULT, 2017). Com essas abordagens, o autor dá conta de falar sobre os aspectos do sujeito em várias de suas obras.

Até aqui mapeamos que, em sua bibliografia, Foucault aponta o poder disciplinar e o biopoder que, dentro de suas características moldam os sujeitos. Em consequência da análise dos poderes, Foucault passa a falar em subjetivação – que compreende os processos positivos e negativos das ações de poder em relação aos sujeitos (FOUCAULT, 2010; 2017), no intuito de entender as existências e as implicações decorrentes dela.

Os apontamentos de Foucault são fundamentais para esse trabalho ao passo que considerando, os exemplos explorados de tecnologias já operantes, bem como as previsões para o futuro, qual subjetividade será moldada a partir da vivência atual?

Tomando como base as iniciativas listadas ao longo desse estudo, notamos um grande esforço em fazer com que robôs sejam vistos como amigáveis, companheiros, prestativos e atentos; por vezes, são até creditadas a eles características que seriam de um “ser humano ideal” – aquele que não briga, nunca se chateia, está sempre feliz e disponível para qualquer atividade. Com isso, inferimos que essas são algumas características que moldarão a subjetividade agora e nos próximos anos.

Sabemos que Foucault não chegou a vivenciar essa sociedade atual com máquinas mais difundidas no cotidiano, ou mesmo com vigilância constante dos smartphones, mas sua obra auxilia e influencia diversos pensadores que estão analisando a realidade atual e, a seguir, vamos falar um pouco sobre uma dessas novas leituras.

Passamos por alguns exemplos que dão conta de robôs que falam, que buscam atenção, que têm aparência humana, que têm pele ou mesmo o olhar penetrante. E é latente que, a cada dia que passa, as características que impõem distinções entre pessoas e máquinas vão se esvaindo e, com isso, aquilo de delimitava vida e morte e o entendimento sobre biopolítica não mais é suficiente para falar sobre a atualidade (KATZ, 2019:762).

O conceito de biopolítica tem mais de 100 anos, mas a ideia que mais se popularizou, dentre as várias vertentes teóricas que tratam o assunto, foi a abordagem feita por Foucault. Como falamos há pouco, ele abordou o tema do poder em diversos momentos de seus estudos, ao se debruçar criticamente na análise de instituições como hospitais, prisões e manicômios. Porém, foi a partir de 1976 que a biopolítica se tornou recorrente em suas pesquisas e aulas.

A exploração de Foucault sobre as relações de poder abriu caminho para o trabalho de autores como Giorgio Agamben e Nikolas Rose por exemplo. Porém, dado o enfoque desse trabalho, consideraremos os estudos da antropóloga Elizabeth Povinelli.

Em um recente artigo, Katz (2018:764 apud POVINELLI, 2016:8) traz o questionamento sobre o quanto essa separação entre poder soberano, poder disciplinar e biopolítica, proposta há cerca de 40 anos, continua sendo capaz de “descrever o tipo de Estado no qual vivemos hoje, quando governos, ao mesmo tempo, fazem viver, deixam morrer e matam”. Para possibilitar a resposta a esse questionamento, Katz recorre ao pensamento de Povinelli; que, por sua vez, amplia a preocupação observada na análise de Foucault. É essa teoria que oferece uma perspectiva para a análise da relação entre pessoas e máquinas no presente.

Para Povinelli (2016), a biopolítica não só regula a vida e morte, mas também cria e mantém a divisão entre vida e não vida. É dessa forma que os estados neoliberais conseguem governar as diferenças.

Uma das bases que fundamenta esse pensamento é o conceito que Povinelli chama de imaginário carbono, que compreende, no conjunto de processos metabólicos (nascimento, reprodução e morte), algo considerado como vida biológica pelo imaginário (hegemônico) ocidental, e que também marca a diferença de compreensão entre vida e não-vida.

O imaginário carbono, da epistemologia ocidental, separa o que é orgânico e inorgânico, como se uma matéria inerte não pudesse passar pelos processos metabólicos. Por isso, para a autora, a ontologia ocidental é tida como uma biontologia, que transforma algo local em algo que tem como objetivo um ordenamento universal.

Povinelli é veemente ao falar sobre o quanto os Estados e o capital neoliberais foram beneficiados por esse poder biontológico.

A partir desse pensamento, a autora expande o conceito de biopolítica de Foucault, ao afirmar que esse poder não só comprehende o governo sobre a vida, mas também cria e mantém uma separação entre vida e não vida para que seja possível governar as diferenças.

Nesse contexto, Povinelli (2016) propõe um conceito que transcende as dualidades (vida e não vida, humano e animal, animal e planta, etc.): a geontologia. Portanto, “não se trata mais apenas de regular a vida e a morte dos indivíduos, mas do planeta inteiro”, explica Katz (2018:765).

Para explicar o conceito, afastando-se de uma tautologia, Povinelli se dedica a diferenciá-lo do que foi proposto por Foucault:

“A maneira mais simples de esboçar a diferença entre geontopoder e biopoder é que o primeiro não opera por meio do governo da vida e das táticas da morte, mas sim é um conjunto de discursos, afetos e táticas usados no liberalismo tardio para manter ou moldar a distinção entre Vida e Não-Vida. Este livro argumenta que à medida que as divisões ordenadoras, anteriormente, estáveis da Vida e Não-Vida se abalam, novas figuras, táticas e discursos de poder substituem o quarteto biopolítico” (POVINELLI, 2016:16-17, tradução nossa)³⁶.

³⁶ Tradução livre do autor. No original: “The simplest way of sketching the difference between geontopower and biopower is that the former does not operate through the governance of life and the tactics of death but is rather a set of discourse, affects, and tactics used in late liberalism to maintain or shape the coming relationship of the distinction between Life and Nonlife. This book argues that as the previously stable ordering divisions of Life and Nonlife shake, new figures, tactics, and discourses of power are displacing the biopolitical quartet” (POVINELLI, 2016:16-17).

Isso posto, Povinelli justifica a necessidade de uma nova forma de analisar o poder nos dias de hoje, mostrando que a definição entre vivo e morto precisa mudar – tirando o homem do centro. Ela propõe que passe para vivo e inerte, pois “1) a morte passou a ser incluída no conceito de vida; e 2) o que antes não podia ser associado com a vida, por ser inerte, agora pede um estatuto ontológico próprio, pois já estamos convivendo com máquinas inteligentes e sensíveis” (KATZ, 2018:764 apud POVINELLI, 2016).

As propostas e indagações de Povinelli demonstram a urgência de se pensar em formas de analisar esse cenário atual, considerando que as teorias que foram elaboradas em um passado, ainda que recente, não dão plena conta de explicar e analisar. E é importante considerar esses novos apontamentos, pois são nessas condições que se moldam as subjetividades.

2.2 – A crise no sujeito e a necessidade de novas abordagens

Povinelli fala de um contexto mais amplo, mais voltado às relações de poder da sociedade. Agora, vamos considerar também uma diferente abordagem para falar sobre o que é ser humano.

O cenário político e social das décadas de 1980 e 1990 provocaram uma crise no conceito de ser humano – muito por conta dos avanços tecnológicos e pelas propostas de hibridação entre homem e máquina: “O conceito de humano explodiu sob a dupla pressão dos avanços científicos contemporâneos e das preocupações econômicas globais”, sintetiza Rosi Braidotti (2013:1, tradução nossa³⁷).

Era, de certa forma, latente a necessidade de nomear e pensar esse novo momento. E assim nasceu o conceito de pós-humano com diferentes abordagens e autores – nomes como Donna Haraway, Robert Pepperell e Katherine Hayles, para citar alguns.

Haraway, em 1985, publicou o livro “Manifesto Ciborgue: Ciência, Tecnologia e Feminismo-socialista ao Final do Século XX”, no qual faz uma leitura progressista e feminista sobre o ciborgue – à época, ela utilizou a figura para falar do capitalismo.

Pepperell, por sua vez, em sua publicação *A condição pós-humana* (1995) organizou e falou sobre as tecnologias ditas pós-humanas, como a nanotecnologia e a vida artificial.

Já Hayles em sua obra, que citamos brevemente no primeiro capítulo, dizia que o pós-humano consiste em um corpo que integra informação e matéria de forma a incluir componentes humanos e não humanos (1996). A obra considerava “como a informação

³⁷ Tradução livre do autor. No original: “The concept of the human has exploded under the double pressure of contemporary scientific advances and global economic concerns” (BRAIDOTTI, 2013:1).

perdeu seu corpo, como o ciborgue foi construído nos anos do pós-guerra como artefato tecnológico e ícone cultural, e como o humano se tornou o pós-humano”, algo que, na época, era apontado como “uma transformação a ser temida e mais abominável do que bem-vinda e abraçada” (1996:291, tradução nossa³⁸).

Porém, com o passar dos anos, foi-se clarificando a ideia de que nomear de “pós” não necessariamente presumia um estudo do que é o humano, mas qualificava aquilo que existiria depois de (um não definido) humano.

Na publicação *Biocultural Creatures* (2016), a autora Samantha Frost busca investigar a raça humana sem partir do pressuposto “pós-humano”. Como premissas de pesquisa, ela explica que as características utilizadas para definir e distinguir a humanidade até hoje foram profundamente descreditadas pela história, o que faz com que a noção construída seja pouco recomendada para atualidade

“Sob escrutínio crítico, a noção do humano passou a ser percebida como um índice de uma fantasia historicamente específica de domínio sobre o eu, a terra e todas as suas muitas criaturas. As características, qualidades e capacidades que antes eram tomadas para definir e distinguir o humano, a humanidade, foram tão profundamente desacreditadas pelas análises históricas, sociais e científicas que a própria noção parece estar falida” (FROST, 2016: 1, tradução nossa³⁹).

O livro se propõe a ser “teoricamente útil”, nas palavras de Frost, para que a humanidade possa enfrentar a devastação social e ecológica que se aproxima (2016:3). E, para transformar as críticas aos humanos em algo positivo, ela propõe que os seres humanos passem a ser concebidos como “criaturas bioculturais”.

Criaturas, porque os seres humanos (como quaisquer outras criaturas) estão vivos e são capazes de permanecerem nessa condição devido à capacidade de extrair diversas formas de sustento de um habitat; e bioculturais, porque se desenvolvem, crescem, persistem e morrem – portanto, todas as criaturas são bioculturais. Não é possível negar a animalidade biológica.

O fato de ela nomear os seres humanos como criaturas é, inclusive, uma forma de aproximá-lo dos outros seres. Esse é um dos motivos pelos quais ela diz que a humanidade

³⁸ Tradução livre do autor. No original: “How information lost its body, how the cyborg was constructed in the postwar years as technological artifact and cultural icon, and how the human became the posthuman have at times seemed to present the posthuman as a transformation to be feared and abhorred rather than welcomed and embraced” (HAYLES, 1996:291).

³⁹ Tradução livre do autor. No original: “Under critical scrutiny, the notion of the human has come to be perceived as an index of a historically specific fantasy of mastery over the self, the earth, and all its many creatures. The characteristics, qualities, and capacities that heretofore have been taken to define and distinguish to human, humanity — the human — have been so profoundly discredited through historical, social, and scientific analysis that the notion itself seems to be bankrupt” (FROST, 2016: 1).

se configura como uma “fantasia oca”, pois separa natureza e cultura – que não é possível, pois um corpo não existe nem só biologicamente e nem só culturalmente.

Quando fala em cultura, Frost salienta que não se refere ao substantivo, mas sim à proposta do verbo “culturar” (*to culture*, no texto original).

Dessa forma, é possível pensar na cultura como algo que pode ser cultivado, ou seja, culturar é cultivar um comportamento. A proposta do termo biocultural é uma tentativa de captura e elucidação de como as criaturas se tornam o que são distintamente por meio dos habitats que as cultiva (2016:5).

Pensar os seres como criaturas bioculturais fez com que Frost questionasse a denominação “raça humana”. Segundo ela, o termo indicaria a existência de apenas uma raça: “como ideal, a categoria do humano foi interpretada como universal, o que significa dizer que todos, supostamente, são capturados sob sua rubrica”, explica (2016:6, tradução nossa⁴⁰); e complementa, em outro momento, dizendo que “o humano, como tendemos a pensar, é construído em torno de uma profunda recusa em admitir ou reconhecer algumas das condições enlameadas e confusas da existência que os humanos compartilham com todos os seres vivos” (2016:11, tradução nossa⁴¹).

Para deixar mais clara a proposta, Frost faz um paralelo entre homens e animais: costumeiramente homens, estão em ambientes, enquanto os animais em habitats. O que difere um ambiente de um habitat? Essa diferenciação não faz sentido.

Dentre vários outros exemplos e pensamentos expostos no livro, selecionamos mais um que é de grande contribuição para a presente pesquisa. Frost considera que a tecnologia é biocultural; podendo ser definida como algo criado pelo homem e que modifica o que não foi criado pelo homem: o corpo.

Os apontamentos de Frost dão uma importante contribuição para nossa análise por procurarem compreender a existência do ser humano no mundo, considerando que ele é tão natural quanto qualquer outro ser vivo; e também por tratarem a tecnologia como um “porém” em relação a definição do que é humano, e não como algo que determina o que ele será. Essa determinação, se é que podemos dizer assim, cabe ao habitat onde ele está inserido e àquilo que ele vai culturar.

⁴⁰ Tradução livre do autor. No original: “As an ideal, the category of the human has been construed as universal, which is to say that everyone supposedly is captured under its rubric” (FROST, 2016: 6).

⁴¹ Tradução livre do autor. No original: “The human as we have tended to think it is built around a profound refusal to admit or to acknowledge some of the muddy and messy conditions of existence that humans share with all living creatures” (FROST, 2016:11).

2.3 – Sobre relações e afetos

“Só eu posso pensar
 Se Deus existe. Só eu
 Só eu posso chorar
 Quando estou triste. Só eu
 Eu cá com meus botões de carne e osso
 Eu falo e ouço”
(Gilberto Gil – Cérebro Eletrônico, 1969).

Depois de explorar os textos de Foucault, Povinelli e Frost, podemos dizer que demos conta de falar sobre o sujeito, sobre o contexto político-social no qual ele está inserido, e também sobre outros ângulos que podem ser abordados quando o tema é seres humanos. Agora, daremos um último passo nessa investigação sobre as relações humanos-máquinas ao falar sobre subjetividade, relações e sentimentos.

Quando analisamos os exemplos citados anteriormente - dos robôs *Spot*, *Lovot* e *Erica* - apontamos também os sentimentos que eles provocam nos humanos, além de mostrar como as empresas projetam os detalhes (como a aparência de seus corpos e adjetivos utilizados na comunicação) para incentivar que os humanos se conectem com eles. Para pensar sobre essas relações estabelecidas e sobre os sentimentos envolvidos, partiremos de uma pergunta feita por pesquisadores da Universidade de Carnegie Mellon, nos Estados Unidos: como um humano reagiria ao ser insultado por um robô?

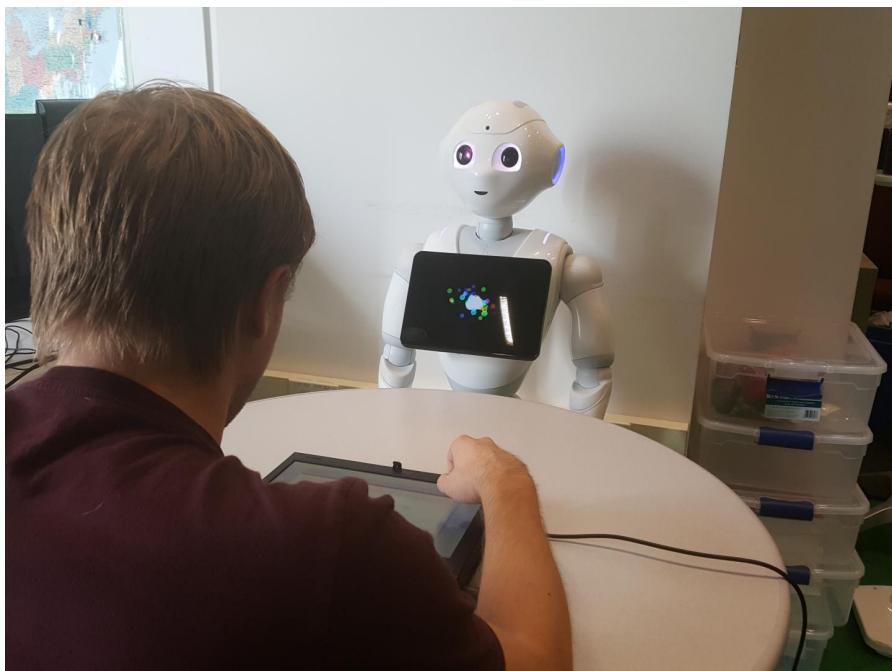
Para chegar em uma resposta, eles pensaram em um teste que consistia em colocar 40 voluntários para jogar o jogo *Guards and Treasures* com o robô *Pepper* (que é vendido pela *SoftBank Robotics* desde 2015, tendo como principal funcionalidade o reconhecimento de emoções humanas).

Uma amostra dos voluntários foi insultada por *Pepper* ao longo das partidas, e os pesquisadores compararam o desempenho destes com aqueles que foram elogiados.

Os resultaram mostraram que os jogadores eram afetados pelas provocações mesmo sabendo que elas vinham de um robô programado para dizer aquilo. Com isso, o estudo concluiu que: “um robô humanoide que incentiva ou desencoraja um oponente humano pode afetar a racionalidade desse humano. (...) um robô desencorajador levou a uma menor racionalidade [dos voluntários], enquanto um robô encorajador foi associado a uma maior racionalidade” (ROTH et al., 2019, tradução nossa⁴²).

⁴² Tradução livre do autor. No original: “A humanoid robot that encourages or discourages a human opponent can impact that human’s rationality. In our study, a discouraging robot led to lower rationality while an encouraging robot was associated with higher rationality” (ROTH et al., 2019).

Figura 6 - Robô Pepper desencoraja adversário durante uma partida



Fonte: ROTH, Aaron M. et al. A Robot's Expressive Language Affects Human Strategy and Perceptions in a Competitive Game. 28th IEEE International Conference on Robot Human Interactive Communication, Nova Delhi, 24 out. 2019. Disponível em: arXiv:1910.11459. Acesso em: 15 dez. 2019.

O estudo ainda acena para a questão que salientamos, ao longo do texto, de que aqueles que desenvolvem novas tecnologias precisam levar mais em conta os possíveis impactos de suas iniciativas: “nossas descobertas podem servir para ajudar futuros projetistas de robôs a desenvolverem uma melhor compreensão sobre os impactos dos afetos de robôs sociais durante interações não cooperativas” (ROTH et al., 2019, tradução nossa⁴³).

O teste com o robô *Pepper* nos dá indícios de que, mesmo com o conhecimento de que se trata de uma máquina, os sentimentos não são freados, diminuídos ou mesmo diferentes do que se a interação fosse feita por um ser vivo.

Propomo-nos a ler esse comportamento considerando os estudos da filósofa Agnes Heller sobre sentimentos. Ela sugere que “sentir significa estar envolvido com algo (...) Esse ‘algo’ pode ser qualquer coisa: outro ser humano, um conceito, eu mesmo, um processo, um problema, uma situação, outro sentimento... outro envolvimento” (1993:15-

⁴³ Tradução livre do autor. No original: “Our findings may serve to help future robot designers develop a better understanding of how affect impacts perceptions of a social robot during non-cooperative interactions. Useful future work would be to investigate nonverbal modes of expression, like body movement and gestures, in competitive settings” (ROTH et al., 2019).

16, tradução nossa⁴⁴). Ou seja, para ela, aquilo que nos afeta não precisa ter uma existência concreta. Esse pensamento, batizado de teoria dos sentimentos, ajuda-nos a pensar sobre o que resulta dessa atual interação homem-máquina do ponto de vista dos sujeitos.

Podemos entender, de forma mais direta para a presente pesquisa, que a autora propõe que os contatos que cada sujeito tem com o mundo em sua volta e consigo próprio resultam em afetos, sentimentos ou emoções.

Esses sentimentos podem ter intensidades variadas e sofrer mudanças ao longo do tempo; existem também modificações na configuração do sentir para cada sujeito; e o fato de que vários afetos são vividos simultaneamente.

Independente da característica e/ou amplitude dos afetos, Heller ressalta que os “nossos sentimentos expressam: eles dão as informações fundamentais que caracterizam o que realmente somos” (1993:78, tradução nossa⁴⁵) e finaliza dizendo que “um homem sem sentimento é inimaginável” (1993:78). Os afetos, portanto, são indispensáveis para a subjetividade.

Tomando como base o fato de que os sentimentos nos singularizarem como humanos, voltamo-nos novamente para o exemplo de Erica, com sua capacidade de emular e os efeitos que isso causa nos humanos, como analisamos nos comentários feitos no YouTube. Ao assistir o vídeo, ou seja, ter algum contato com Erica, as pessoas se conectam e criam um sentimento por ela.

A androide, por sua vez, é capaz de emular emoções o suficiente para estabelecer essa conexão com as pessoas. Ela é preparada para isso. Assim, a máquina se torna responsável por incentivar um sentimento, influenciar na subjetivação.

Hoje, esse contato entre máquinas e pessoas não é amplamente difundido, mas as previsões apontam que, no futuro relativamente próximo a convivência será mais intensa. Erica, por exemplo, é um protótipo para estudos; os robôs *Spot*, *Lovot* e *Pepper* têm distribuição bastante limitada e alto custo.

Por isso, para aprofundarmos um pouco mais o assunto, vamos recorrer à ficção que nos ajuda a esgarçar a realidade e olhar de forma mais diversa para as relações entre humanos e máquinas.

⁴⁴ Tradução livre do autor. No original: “Sentir significa estar implicado en algo (...) Ese ‘algo’ puede ser cualquier cosa: otro ser humano, un concepto, yo mismo, un proceso, un problema, una situación, otro sentimiento... otra implicación” (HELLER, 1993:15-16).

⁴⁵ Tradução livre do autor. No original: “nuestros sentimientos se expresan: dan la información fundamental sobre lo que realmente somos. Un hombre sin sentimiento es inimaginable” (HELLER, 1993:78).

Imagine um cenário em que “Ericas” andam livremente e estabelecem conexões com humanos; a relação e a intimidade são permitidas, aceitas e, até mesmo incentivadas. Quais sentimentos estariam presentes nessas relações?

Ainda não temos a resposta para isso, mas uma simulação de como seria pode ser vista na série *Westworld*⁴⁶ (2016), do canal HBO. A trama apresenta um parque temático povoado por androides, chamados de anfitriões, que são cópias fiéis de seres humanos, possuindo não só feições como também capacidade de estabelecer longo diálogos e demonstrar sensações. Diferente de outras obras da ficção, não há nada nesses robôs que os difere visualmente de seres humanos.

O parque, que imita o velho oeste americano, foi construído para satisfazer quaisquer desejos humanos sem grandes restrições sociais e morais. Ao entrar no local, os visitantes (humanos) podem se aventurar em romances, envolver-se em duelos na rua, matar a sangue frio, embebedar-se em casas de prostituição, estuprar donzelas indefesas, praticar roubos, caçar indígenas e mais um sem número de possibilidades, que incluem também entrar em um *saloon* para jogar conversa fora.

Toda a ação de *Westworld* é calcada nas narrativas criadas para cada anfitrião, que é o que permite que eles desenrolem vivências corriqueiras que causam ao visitante uma sensação de fidelidade com o mundo real. Na teoria, essas criaturas não possuem atitudes próprias, ficando restritas a um roteiro pré-programado.

Considerando que os visitantes têm toda liberdade no parque, e é comum a prática de crimes – nos quais muitas vezes os androides são danificados e/ou mortos – muitos anfitriões são levados a uma espécie de hospital, um lugar onde eles recebem manutenção, onde são restaurados e têm suas memórias apagadas para serem devolvidos a suas rotinas.

A trama, conhecida por diversas reviravoltas e diferentes linhas temporais, ganha contornos mais dramáticos quando os anfitriões passam a reunir fragmentos de memória, lembranças de dor e tristeza que viveram nas mãos dos visitantes. Isso desperta um sentimento humano nas máquinas, que passam a buscar uma mudança de vida. *Westworld* percorre os mais profundos desejos humanos, e também sugere que o que separa homens e máquinas é a capacidade de lembrar, de ter uma autoconsciência.

Durante uma entrevista com atores e produtores da série, a futurista Amy Webb apontou detalhes importantes no enredo, que podem dar pistas sobre o futuro da humanidade ou mesmo incitar mais reflexão antes que seja iniciada a produção em série

⁴⁶ A série tem como base o filme *Westworld – Onde ninguém tem alma*, que foi escrito e dirigido por Michael Crichton, mesmo autor de “O parque dos dinossauros”, e lançado em 1973. Um dos destaques fica por conta do recurso de processamento digital de imagens, utilizado para mostrar o ponto de vista dos androides em diversas passagens do longa.

de robôs: “Na segunda temporada, vemos os anfitriões ganhando valores humanísticos, como empatia e vingança” e completou, ressaltando que os personagens “estão aprendendo, assim como nós, humanos, por meio de técnicas de reforço positivo. É uma boa descrição de como os sistemas de IA são treinados na vida real” (PHILIPS, 2019).

Daugherty e Wilson, por exemplo, preveem que “No futuro, a IA desempenhará um papel cada vez mais importante para garantir que sistemas avançados estejam operando dentro dos limites éticos e morais humanos” (2018: 130, tradução nossa⁴⁷).

Em diversos momentos da trama, é apresentada a dualidade vivida pelos visitantes do parque: de um lado, a euforia de poder fazer o que quiser; e, de outro, a estranheza de se relacionar com uma máquina tão humana. Em pouco tempo, dentro do universo de Westworld, a estranheza é deixada de lado pelos visitantes e logo eles passam a interagir naturalmente com os anfitriões.

Dessa forma, a série coloca em perspectiva algo que pode ser resumido por uma pergunta feita por uma anfitriã a um dos visitantes: “Se você não nota a diferença, importa se sou real ou não?”⁴⁸ (WESTWORLD, 2016). Rapidamente, adaptamo-nos às situações e nos envolvemos com aqueles (ou aquilo) que estão a nossa volta, como explica Heller, 1993. Dessa forma, uma máquina que tem algum traço humano pode ser respondida como se fosse humana.

A tecnologia não é algo acadêmico, industrial, apartada da civilização ou qualquer outra coisa que a deixe centrada em si própria. A tecnologia está em contato direto com as pessoas e com o planeta de forma a causar efeitos nas relações sociais e, por consequência, nas relações do sujeito com ele próprio.

Ou seja, a interação entre humanos e não humanos provoca sentimentos e, segundo Heller (1993), é isso que nos faz humanos. Logo, esse contexto de convivência de pessoas e máquinas passa a influenciar no que cada pessoa é (influencia a subjetivação).

⁴⁷ Tradução livre do autor. No original: “In the future, AI itself will play an increasingly important role to ensure that advanced systems are operating within human ethical and moral bounds” (DAUGHERTY; WILSON, 2018: 130)

⁴⁸ Tradução livre do autor. No original: “If you can’t tell the difference, does it matter if I’m real or not?” (WESTWORLD, 2016)

3 – CONCLUSÃO

O presente trabalho tem natureza teórica e também qualitativa, propondo a reflexão sobre tecnologia, interações entre seres humanos e máquinas, e sobre as interferências resultantes dessa relação nos processos de subjetivação.

O objetivo era coletar pistas que podem nos mostrar o que o futuro reserva, quando consideramos que a sociedade está em profunda transformação – permeada pela maior frequência de interações entre pessoas e máquinas.

Para dar conta desse retrato do momento histórico, partimos de uma análise histórica e contextual do desenvolvimento tecnológico. Por meio de saltos no tempo, consideramos o que estava por trás da criação da cibernetica e os estudos que sucederam dela, fazendo menção a Escola de Palo Alto e pensadores como Gregory Bateson.

Com isso, pudemos notar que a tecnologia, historicamente, é pensada de forma utilitária e servil, desconsiderando que a interações com ela provocam efeitos nas relações sociais e, por consequência, nas relações do sujeito com ele mesmo.

Feito o panorama histórico, listamos algumas das possibilidades que podem definir o futuro da relação entre pessoas e máquinas, se é que podemos falar de forma separada, demonstrando três diferentes correntes de pensamentos: uma mais otimista, uma mais conservadora e uma pessimista.

Kurzweil (2005) puxa a fila dos otimistas ao acreditar que um futuro no qual o desenvolvimento tecnológico chegue a uma superinteligência reserva à humanidade a extinção da pobreza e da fome. Um pouco mais receoso, mas ainda otimista, o brasileiro Nicollelis aposta em uma libertação do corpo por meio da máquina. De acordo com ele, a tecnologia permitirá que o cérebro humano se liberte do corpo.

Em um espectro mais conservador temos Tegmark (2017), que desenvolve 13 cenários futuros diferentes – considerando desde um modelo no qual pessoas e máquinas coexistem, a pobreza foi extinta e não existem doenças (nesse caso foram criadas áreas exclusivas para humanos vivam) até a probabilidade de um futuro de completa destruição motivada por guerras, resultando em um planeta exaurido de recursos naturais.

Tegmark se mostra cauteloso ao futuro não só por conta dos objetivos por trás do desenvolvimento tecnológico atual, mas também pela rapidez com que tudo tem acontecido. Essa velocidade, como detectamos, é o ponto que une todas as análises. Em um caminho similar, está o pensamento de Vernier, que cunhou o termo “singularidade tecnológica”.

De forma mais pessimista (ou realista) está Barrat (2013), que diz que pesquisadores acabarão por criar uma inteligência mais evoluída do que a humana, algo que não poderá ser controlado ou entendido adequadamente. Ele aproveita essa previsão para alertar que pouco tem sido pensado quanto aos perigos que envolvem esse tipo de desenvolvimento.

O embasamento de diferentes linhas de pensamentos sobre o futuro tecnológico possibilitou que olhássemos para o presente, para aquilo que está sendo desenvolvido hoje, de forma crítica, considerando as possíveis motivações e objetivos dos desenvolvedores, para que, em um segundo momento, também fosse possível considerar como essas novidades estavam sendo recebidas.

Foram avaliados o robô *Spot*, da Boston Dynamics, que com sua aparência de cachorro causa um misto de estranheza e desejo; o robô *Lovot*, que busca uma ligação “humana” (nas palavras do fabricante) com as pessoas; e de Erica, uma invenção do pesquisador Hiroshi Ishiguro para “saber o que é um ser humano” (GUIZZO, 2010).

Com essa análise, concluímos que as fabricantes empregam grande esforço em humanizar suas criações, tanto no *design* que as criaturas carregam quanto nos discursos utilizados para promover a venda.

No segundo capítulo, falamos sobre o que comprehende o sujeito e subjetivação sob o ponto de vista de Michel Foucault. Nesse percurso teórico, conceituamos que a subjetividade é fruto da relação de si consigo e que ela é formada por várias atividades.

Compreendemos também que essa relação sofre influência do que Foucault chama de biopoder, que comprehende em um controle sutil e potente dos corpos – algo que é imprescindível ao capitalismo.

A lógica capitalista, muito pautada na imagem, tem grande interesse em seduzir e encantar os sujeitos, de forma a moldá-los e incentivá-los a determinados consumos.

Utilizamos os exemplos dos robôs apresentados no primeiro capítulo para demonstrar que, desde a concepção deles (que atribui a eles um comportamento amigável, prestativo e sempre disposto) até sua divulgação, existe a preocupação de coloca-los com características de um “ser humano” ideal e, por consequência, de interferir na subjetividade das pessoas.

Ainda considerando a abordagem do biopoder para Foucault, ressaltamos a necessidade que temos de novas teorias que deem conta do cenário atual de convivência entre pessoas e máquinas.

Para fazer isso, utilizamos o conceito de geontologia, proposto por Elizabeth Povinelli. Partindo da proposta de que precisamos de um modelo em que o homem não

está no centro e, também, de que hoje é necessário se levar em conta que a definição de vida e morte não mais comporta a atualidade, Povinelli nos dá um importante instrumento para analisar o presente e, a partir dele, detectar as subjetividades que decorrem dele.

Ainda nos aproveitando de novas abordagens metodológicas, voltamo-nos para a necessidade de falar sobre o humano. Portanto, partimos de um cenário político para um cenário social.

Para falar sobre o humano, primeiro abordamos o conceito de pós-humano, utilizado por diversos pensadores, que foi uma resposta aos avanços científicos das décadas de 1980 e 1990, e o quanto ele não dá conta de explicar a atualidade.

Partindo de uma nova perspectiva, a autora Samantha Frost aponta a necessidade de se olhar para o ser humano como uma criatura biocultural. O que pressupõe que a fronteira que separa humanos e outros seres vivos não deve existir.

Ao pensar dessa forma, Frost vê a tecnologia não como algo que determina o homem (tornando-o em pós-moderno, por exemplo), mas como algo biocultural; ou seja, para ela, a tecnologia não determina o que é o homem.

A obra de Frost é um convite a pensarmos a relação entre seres humanos e máquinas sob uma perspectiva diferente do que temos visto em estudos recentes.

Por fim, falamos sobre as reações e sentimentos que as máquinas causam nos humanos. A base para esse pensamento foi a Teoria dos Sentimentos, de Agnes Heller, que diz que o envolvimento com algo é o que nos faz humanos. Para ela, os contatos que cada sujeito tem com o mundo a sua volta e consigo próprio resultam em afetos, sentimentos ou emoções, e esses sentimentos são indispensáveis para os humanos.

Essa leitura de Heller possibilita que analisemos as emoções que os robôs causam nas pessoas como algo que vai acabar por influenciar na subjetividade delas.

No final do primeiro capítulo, contamos uma história hipotética na qual um humano é salvo por um robô em um incêndio. Ao final perguntamos: O que você sentiria pelo seu salvador? Gratidão, carinho, amor? E o que ele sentiria por você?

E agora, ao final da investigação consideramos que, independente do sentimento causado, uma conexão é estabelecida. Esse convívio com a máquina causa interferências na subjetivação que ainda não damos conta de mapear, mas que demandam uma urgente atenção dada à velocidade com que as mudanças têm acontecido.

REFERÊNCIAS

ASIMOV, Isaac et al (ed.). **Histórias de robôs volume 2.** Tradução: Milton Persson. São Paulo: Editora L&PM Pocket, 2010.

BARRAT, James. **Our Final Invention:** Artificial Intelligence and the end of human era. Nova Iorque: Thomas Dunne Books, 2013.

BATESON, Gregory. **Steps to an ecology of mind.** Londres: Jason Aronson Inc, 1987. Arquivo PDF.

BOSTON DYNAMICS. **Boston Dynamics homepage.** Boston, 2019. Disponível em: <https://www.bostondynamics.com/>. Acesso em: 2 dez. 2019.

BRAIDOTTI, Rosi. **The Posthuman.** Cambridge: Polity, 2013.

COSTA, Alyne de Castro. **Virada geo(nto)lógica:** reflexões sobre vida e não-vida no antropoceno. Revista AnaLógos, Rio de Janeiro, v. 1, p. 140-150, 24 nov. 2016. Disponível em: <https://www.maxwell.vrac.puc-rio.br/28127/28127.PDF>. Acesso em: 18 dez. 2019.

DAUGHERTY, Paul R.; WILSON, H. James. **Human + Machine:** Reimagining Work in the Age of AI. Boston: Harvard Business Review Press, 2018.

DEBORD, Guy. **A Sociedade do espetáculo.** Comentários sobre a sociedade do espetáculo. 1. ed. Rio de Janeiro: Contraponto, 1997.

ERICA talks about her memories. Japão: **Hiroshi Ishiguro Lab,** 2017. Disponível em: <https://www.youtube.com/watch?v=VD2Btrj9ipY>. Acesso em: 29 maio 2019.

FOUCAULT, Michel. **Em defesa da sociedade:** Curso no Collège de France (1975-1976). São Paulo: Martins Fontes, 1999.

_____. **Hermenêutica do sujeito.** Tradução: Eduardo Brandão. São Paulo: Martins Fontes, 2004.

_____. **Nascimento da biopolítica:** curso dado no Collège de France (1978-1979). São Paulo: Martins Fontes, 2008.

_____. **Vigiar e Punir:** nascimento da prisão. Petrópolis: Vozes, 2010.

_____. **História da Sexualidade 1.** A vontade de saber. Tradução Maria Thereza da Costa Albuquerque e J. A. Guilhon Albuquerque. 4a ed. Rio de Janeiro/São Paulo: Paz e Terra, 2017.

FOUNDATION 2045. **2045 Strategic social initiative.** Rússia, 2012. Disponível em: <http://2045.com>. Acesso em: 5 nov. 2019.

FROST, Samantha. **Biocultural Creatures:** toward a new theory of the human. Durham: Duke University Press Books, 2016.

GAIMAN, Neil. Introdução. In: GIBSON, William. **Neuromancer**. Tradução Fábio Fernandes. São Paulo: Aleph, 2016. p. 7-8.

GAZZALEY, Adam; ROSEN, Larry D. **The distracted mind**: Ancient brains in a high-tech world. Cambridge: The MIT Press, 2016.

GIBSON, William. **Neuromancer**. Tradução Fábio Fernandes. São Paulo: Aleph, 2016.

GROOVE X, INC. **Lovot powered by love**. Tóquio, 2019. Disponível em: <https://lovot.life/en/>. Acesso em: 2 dez. 2019.

GUIZZO, Erico. Hiroshi Ishiguro: **The Man Who Made a Copy of Himself**. IEEE Spectrum, Nova Iorque, 23 abr. 2010. Disponível em: <https://spectrum.ieee.org/robotics/humanoids/hiroshi-ishiguro-the-man-who-made-a-copy-of-himself>. Acesso em: 4 maio 2019.

HARARI, Yuval Noah. **21 lições para o século 21**. São Paulo: Companhia das Letras, 2018.

HAWKING, Stephen *et al.* Stephen Hawking: Transcendence looks at the implications of artificial intelligence - but are we taking AI seriously enough?. **Independent**, Londres, 1 maio 2014. Disponível em: <http://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-atthe-implications-of-artificial-intelligence-but-are-we-taking-9313474.html>. Acesso em: 21 nov. 2018.

HAYLES, Katherine. **How we became Posthuman**: virtual bodies in cybernetics, literature, and informatics. Chicago: The University of Chicago Press, 1999.

HELLER, Agnes. **Teoria de los sentimientos**. México: Coyoacán, 1993.

HIROSHI ISHIGURO LABORATORIES. **Projects**. Osaka. Disponível em: <http://www.geminoid.jp/en/projects.html>. Acesso em: 4 maio 2019.

ISTOÉ, Revista. Muito estrago por nada. **Revista Istoé**, São Paulo, n. 1577, 22 dez. 1999. Tecnologia & Meio Ambiente. Disponível em: [https://istoe.com.br/27774_MUITO+ESTRAGO+POR+NADA+/. Acesso em: 20 nov. 2019.](https://istoe.com.br/27774_MUITO+ESTRAGO+POR+NADA+/)

JAPAN BRANDVOICE. **A Machine That Does Nothing?** Cuddly LOVOT May Be The Most Innovative Japanese Robot Yet. Forbes, Nova Iorque, 6 fev. 2019. Disponível em: <https://www.forbes.com/sites/japan/2019/02/06/a-machine-that-does-nothing-cuddly-lovot-may-be-the-most-innovative-japanese-robot-yet/>. Acesso em: 2 dez. 2019.

JORDAN, Philipp et al. **Exploring the Referral and Usage of Science Fiction in HCI Literature**. ArXiv.org, Nova Iorque, 22 mar. 2018. Disponível em: <https://arxiv.org/abs/1803.08395>. Acesso em: 25 maio 2019.

KATZ, Helena. **Dança, robôs, desigualdade**: como refundar a sociedade do comum. Anais do V Congresso Nacional de Pesquisadores em Dança. Manaus: ANDA, 2018. p. 760-771. Disponível em:

<http://www.helenakatz.pro.br/midia/helenakatz71557582246.pdf>. Acesso em: 19 nov 2019.

_____. **Conexões entre o corpo apps e o mundo regido por editais.** Anais do IV Congresso Nacional de Pesquisadores em Dança. Comitê Produção do Discurso Crítico em Dança. Setembro, 2015. Disponível em: <http://www.helenakatz.pro.br/midia/helenakatz21557582623.pdf>. Acesso em: 19 nov 2019.

KUNZRU, Hari. **Genealogia do ciborgue.** In: TADEU, Tomaz (org. e trad.). Antropologia do ciborgue: as vertigens do pós-humano. Belo Horizonte: Autêntica, 2009. p. 119-126.

KURZWEIL, Ray. In: **Our Final Invention: Artificial Intelligence and the End of Human Era.** Nova Iorque: Thomas Dunne Books, 2013. cap. The Singularitarian.

_____. **The singularity is near: when humans transcend biology.** New York: Viking Penguin, 2005.

NICOLELIS, Miguel. **Muito além do nosso eu:** A nova neurociência que une cérebro e máquinas e como ela pode mudar nossas vidas. São Paulo: Companhia das Letras, 2011. Kindle.

NORVIG, Peter; RUSSEL, Stuart. **Inteligência artificial.** Tradução: Regina Célia Simille de Macedo. 3. ed. Rio de Janeiro: Elsevier, 2013.

PHILLIPS, Jevon. **Meet the ‘Westworld’ panel moderator who showed Comic-Con how it’s done.** Los Angeles Times, Los Angeles, 24 jul. 2019. Disponível em: <https://www.latimes.com/entertainment-arts/tv/story/2019-07-24/comic-con-2019-westworld-hbo-futurist>. Acesso em: 9 dez. 2019.

POVINELLI, Elizabeth A. **Geontologies. A Requiem to Late Capitalism.** Durham: Duke University Press, 2016.

RAY Kurzweil on what the future holds next. [S. l.]: TED Talks, 2018. Disponível em: https://www.ted.com/talks/the_ted_interview_ray_kurzweil_on_what_the_future_holds_next/. Acesso em: 1 set. 2019.

ROTH, Aaron M. et al. **A Robot’s Expressive Language Affects Human Strategy and Perceptions in a Competitive Game.** 28th IEEE International Conference on Robot Human Interactive Communication, Nova Delhi, 24 out. 2019. Disponível em: arXiv:1910.11459. Acesso em: 15 dez. 2019.

SARTRE, Jean-Paul. **O Existencialismo é um humanismo.** Tradução: João Batista Kreuch. Petrópolis: Vozes de Bolso, 2012.

TAYLOR, Diana (ed.). **Michel Foucault: Conceitos fundamentais.** Tradução: Fábio Creder. Petrópolis: Vozes, 2018.

TEGMARK, Max. **Life 3.0: Being Human in the Age of Artificial Intelligence.** Nova Iorque: Knopf, 2017a.

_____. In: HARARI, Yuval Noah. **Life 3.0 by Max Tegmark review – we are ignoring the AI apocalypse.** The Guardian, Londres, 22 set. 2017b. Disponível em: <https://www.theguardian.com/books/2017/sep/22/life-30-max-tegmark-review>. Acesso em: 13 nov. 2019.

TRAVERSA, Giovanna Pasqualin. **Hiroshi Ishiguro: “Humans and humanoids together in future society”.** Agenzia S.I.R., Itália, 26 fev. 2019. Disponível em: <https://www.agensir.it/chiesa/2019/02/26/robotics-hiroshi-ishiguro-humans-and-humanoids-together-in-future-society/>. Acesso em: 27 maio 2019.

VINGE, Vernor. **The Coming Technological Singularity:** How to Survive in the Post-Human Era. VISION-21 Symposium, Ohio, 30 mar. 1993. Disponível em: <https://mindstalk.net/vinge/vinge-sing.html>. Acesso em: 1 dez. 2019.

WESTWORLD. Direção: Jonathan Nolan; Lisa Joy. Produção: Athena Wickham; Bryan Burk; Cherylanne Martin; Jerry Weintraub; J. J. Abrams; Jonathan Nolan; Kath Lingenfelter; Kathy Lingg; Lisa Joy; Richard J. Lewis; Roberto Patino. Intérprete: Anthony Hopkins; Evan Rachel Wood; Thandie Newton, Ed Harris e outros. Estados Unidos: HBO, 2016. 3 DVD's.