

Androides sonham com ovelhas elétricas? Leituras neurobioéticas de Westworld – 1ª Temporada

*Androids dream of electric sheep?
Westworld's neurobioethical readings – Season 1*

*¿Sueñan los androides con ovejas eléctricas?
Lecturas neurobioéticas de Westworld - Temporada 1*

Flávio Marques Andreon^{1,2}, Elen Nas^{2,3}, Vanderson Esperidião
Antonio^{2,4}, Eugênio Silva^{2,5,6}, Romario Brunos Will^{1,2}, Igor Rodrigues
Mendes¹, Mathias Viana Vicari^{2,7}, Lindisley Ferreira Gomides^{2,8}

1.Laboratório de Métodos Epidemiológicos e Computacionais em Saúde, Universidade Federal de Viçosa, Viçosa-MG, Brasil.

2.Redes BraiNNIAC – [<http://www.brainniac.ufv.br/>]

3.Programa de Pós-graduação em Bioética, Ética Aplicada e Saúde Coletiva, Universidade Federal do Rio

4.Laboratório Morfofuncional, Universidade Federal de Viçosa, Viçosa-MG, Brasil.

de Janeiro, Rio de Janeiro-RJ, Brasil.

5.Grupo de Pesquisa em Computação Inteligente, Centro Universitário Estadual da Zona Oeste-RJ, Rio de Janeiro-RJ, Brasil.

6.Centro de Ciências e Tecnologia - Ciência da Computação, Centro Universitário Serra dos Órgãos, Teresópolis-RJ, Brasil.

7.Programa de Pós-Graduação em Meteorologia Aplicada, Universidade Federal de Viçosa, Viçosa-MG, Brasil.

8.Escola de Medicina, Faculdade Dinâmica do Vale do Piranga, Ponte Nova-MG, Brasil.

Resumo

Introdução. O debate dirigido às questões relativas (1) à criação de robôs *androides* (2) dotados de inteligência artificial (IA) e que (3) poderão estabelecer relação com seres humanos tem alcançado grande relevância no campo bioético, técnico-científico e cinematográfico. Este é precisamente o universo descrito na série *Westworld*, as quais abrem caminho para reflexões acerca de uma neuro(bio)ética, termo que abrange conceitos da neurociência, da biotecnologia e da filosofia moral. **Objetivo.** apresentar e discutir questões sobre a neuro(bio)ética relacionadas à consciência aplicada aos entes artificiais e aos seres-humanos, utilizando como subsídio a primeira temporada da série *Westworld*. **Resultados.** No decorrer dos episódios, a série estimula o debate sobre até que ponto se pode estender ou limitar as fronteiras entre arte e a ciência. Também almeja-se discutir e apresentar a imbricada integração do sistema nervoso humano na *construção* dos processos consciência, assim como as interseções entre a consciência e a tecnologia, através da IA e da robótica. **Conclusões.** As atuais perspectivas de evolução da robótica e da IA demandam robusta reflexão ética, a qual tem como pano de fundo a questão *kantiana* "O que é o homem?". Diante disso, é de grande proficuidade entender o que se deseja da tecnologia e os porquês de tal desejo.

Unitermos. consciência; inteligência artificial; filosofia da mente; westworld; neurobioética

Abstract

Introduction. The debate on issues related to (1) the creation of *android* robots (2) endowed with artificial intelligence (AI) and which (3) can establish a relationship with human beings has reached great relevance in the bioethical, technical-scientific and cinematographic field. This is precisely the universe described in the *Westworld* TV-show, which opens the way for reflections on a *neuro(bio)ethics*, a term that encompasses concepts from neuroscience,

biotechnoscience and moral philosophy. **Objective.** present and discuss questions about *neuro(bio)ethics* related to consciousness applied to artificial and human beings, using as a subsidy the first season of the *Westworld* TV-show. **Results.** over the episodes, the TV-show stimulates the debate on the extent to which the boundaries between art and science can be extended or limited. It also aims to discuss and present the imbricated integration of the human nervous system in the *construction* of the consciousness processes, as well as the intersections between consciousness and technology, through AI and robotics. **Conclusions.** The current perspectives for the evolution of robotics and AI demand robust ethical reflection, which has as a background the *kantian* question "What is the man?". Given this, it is of great benefit to understand what is desired from technology and the reasons for such desire. **Keywords.** consciousness; artificial intelligence; philosophy of mind; westworld; neurobioethics

Resumen

Introducción. El debate está dirigido a temas relacionados con (1) la creación de robots *androides* (2) dotados de inteligencia artificial (IA) y que (3) pueden establecer una relación con los seres humanos. Este tema ha alcanzado gran relevancia en el campo bioético, técnico-científico y cinematográfico. Es precisamente el universo descrito en la serie *Westworld*, el que abre el camino a reflexiones acerca de la neuro(bio)ética, término que incluye conceptos de neurociencia, biotecnología y filosofía moral. **Objetivo.** presentar y discutir cuestiones sobre neuro (bio) ética relacionadas con la aplicación de la conciencia a seres artificiales y humanos, utilizando como subvención la primera temporada de la serie *Westworld*. **Resultados.** Durante los episodios, la serie estimula el debate sobre hasta qué punto se puede ampliar o restringir los límites entre arte y ciencia. También tiene como objetivo discutir y presentar la integración aplicada del sistema nervioso humano en los procesos de *construcción* de la conciencia, así como las intersecciones entre conciencia y tecnología, a través de la IA y la robótica. **Conclusiones.** Las perspectivas actuales para la evolución de la robótica y la IA exigen una sólida reflexión ética, que tiene como trasfondo la pregunta *kantiana* «¿Qué es el hombre?». Teniendo esto en cuenta, es de gran beneficio comprender qué se desea de la tecnología y las razones de tal deseo.

Palabras clave. consciencia; inteligencia artificial; filosofía de la mente; westworld; neurobioética

Trabalho realizado na Universidade Federal de Viçosa, Viçosa-MG, Brasil.

Conflito de interesse: não

Recebido em: 31/01/2021

Aceito em: 17/02/2022

Endereço para correspondência: Flávio Marques Andreon. Laboratório de Métodos Epidemiológicos e Computacionais em Saúde, Departamento de Medicina e Enfermagem, Universidade Federal de Viçosa. Av. Peter Henry Rolfs, s/n. Campus Universitário. Viçosa-MG. CEP 36570-900. Email: flavio.andreon@ufv.br

INTRODUÇÃO

O debate ético entre os temas que abrangem inteligência artificial (IA), criação de robôs *androides* e a vindou-ra relação desses com os seres humanos são assuntos de grande pertinência no campo (*bio*)ético, técnico-científico e cinematográfico. A abordagem multifacetada e crítica de tais assuntos na primeira temporada da série de TV *Westworld* [www.hbo.com/westworld], da rede norte-americana de

televisão HBO (*Home Box Office*), traz considerações éticas em seu discurso estético, sendo assim capaz de estimular a reflexão sobre os limites da tecnociência.

A interseção entre a (*bio*)ética e o questionamento do avanço técnico-científico é bem representada na primeira temporada da série. O roteiro é fundamentado numa ficção futurística em que o homem é capaz de criar androides à sua imagem e semelhança, e utilizando-os em um parque temático ao estilo *faroeste*, voltado para o entretenimento da alta sociedade. *Westworld* também estimula a discussão ética, quando os robôs começam a desenvolver sua inteligência e a aquisição da autoconsciência, doravante percebendo que são vassalos de seus criadores, sendo utilizados e objetificados em narrativas contínuas, cíclicas e pré-determinadas.

Parece natural pensar que as mesmas razões prudenciais e éticas para mútuo respeito e tolerância que um ser humano tem face-a-face com outro ser humano, deveria permanecer também para com as novas pessoas biológicas paradigmáticas, mas não humanas¹. Torna-se válida, por conseguinte, a discussão em torno da atribuição de *considerações* (*bio*)éticas aos seres com IA, dotados de consciência. “*Esses seres seriam dignos de inclusão segundo os mesmos princípios?*”

Westworld foi ao ar pela primeira vez em outubro de 2016 e, em abril de 2020, foi renovada a sua quarta temporada. A obra é baseada na longa-metragem “*Westworld*”, por Michael Crichton (1973), classificada na dramaturgia como

"*Sci-Fi Drama*". Desenvolvida por Jonathan Nolan e Lisa Joy, a série conta com um enredo elaborado e amplo uso de recursos especiais.

Analisando o enredo da série, projeta-se uma boa associação entre a questão do desenvolvimento pessoal e técnico-científico atrelado à autoconsciência humana. Pode-se inferir, também, que o desenvolvimento da autonomia pessoal é um fator primordial na bioética ocidental contemporânea e isso está estritamente relacionado ao desenvolvimento da consciência da espécie humana².

O questionamento a respeito da origem da consciência humana e/ou artificial é, de forma contínua, discutido pelo *Homo sapiens*. Tal averiguação também compõe variadas formas de expressão artística, desde o início do conhecimento humano, com pinturas "*pré-históricas*", até a atual ficção científica para cinema e TV³. Não é de se impressionar que a discussão sobre a consciência humana, apesar de ser longa, ainda esteja presente em diversas obras de arte visuais e científicas nos tempos atuais.

Na era moderna, o desenvolvimento técnico, científico e biológico tornou possível a construção de teorias e *modelos físicos e matemáticos* para tentar subsidiar as explicações relacionadas às emoções e à sistemática cerebral, envolvidas na tomada de decisão pelo ser humano, através de efeitos quânticos e eletromagnéticos para um ajuste da compreensão e propagação dessas informações cerebrais⁴.

Levantou-se também a hipótese da criação de modelos

in-silico similares ao cérebro humano e a sua reprodução em massa para inúmeros fins, correspondente ao que é representado no universo de *Westworld*. Essa questão abre caminho para se pensar em uma **neuro(bio)ética**, que é um conceito que reúne conhecimentos da neurociência, biologia e filosofia.

Com base nessas considerações, o objetivo do presente artigo é apresentar e discutir questões sobre a **neuro(bio)ética relacionadas à consciência aplicada aos entes artificiais e aos seres-humanos**, utilizando como subsídio a primeira temporada da série *Westworld*.

A Consciência e a Arte

Relacionar consciência e arte é um estudo muito amplo que demanda um aprofundamento maior, contudo, vale pontuar que a expressão artística não somente representa aspectos da consciência, como também influencia a natureza de sua elaboração. A representação estética une os indivíduos dentro de uma mesma cultura ao passo que representa e afeta fortemente as emoções humanas tendo desse modo, um resultado sobre o (*sub*)consciente.

*"A emoção provavelmente havia se estabelecido na evolução antes do aparecimento da consciência, e emerge em cada um de nós como resultado de indutores que com frequência não reconhecemos conscientemente; por outro lado, os sentimentos produzem seus efeitos supremos e duradouros no teatro da mente consciente."*⁵

Inúmeras composições de variados tipos culturais e de arte podem ser observadas pelos telespectadores. A trilha sonora da série, composta por Ramin Djawadi, consegue captar a temática futurística e também resgata o passado, utilizando canções contemporâneas como *background*, ao som, geralmente, de instrumentos de cordas.

A relação da harmonização da consciência do homem moderno, cujo conflito reside na tendência em identificar-se com seu intelecto, deixa de lado sua percepção sensorial. Nesse caso, também, é possível pontuar o caráter humano como “um sistema relativamente constante de todos os impulsos não instintivos (aspirações e interesses) que conectam uma pessoa ao mundo social e natural”⁶.

Ponderando-se que o imaginário do homem moderno está a entender-se como centro do universo, destaca-se a observação dos visitantes do parque em um ambiente controlado e, teoricamente, sem limites e “imortal”. O parque temático proposto por *Westworld* propicia aos seus visitantes uma experiência ambiciosa relacionada ao poder. Questões bioéticas contemporâneas também debatem sobre a utilização de andróides dotados de IA para fins sexuais⁷. Até que ponto serão saudáveis e sustentáveis as interações e a coexistência de humanos e robôs?

Em uma atualização denominada “*devaneios*” é que acontece o início da tomada da autoconsciência pelos andróides em *Westworld*, simultaneamente à aquisição e ao desenvolvimento da capacidade de improvisação dos mesmos.

Todavia, os gestores do parque, na série, tentam limitar a capacidade de improvisação dos robôs a fim de controlar suas interações sociais.

A série, desse modo, no decorrer dos episódios, estimula o debate sobre até que ponto se pode estender ou limitar as fronteiras entre arte (a imaginação, a estética e a emoção) e a ciência (o estudo, a ética, a racionalização e o planejamento). De acordo com o observado nesses apontamentos, as fissuras entre mente e corpo, razão e emoção, tendem a dificultar uma boa compreensão científica (e ética) sobre os *impasses* apresentados em *Westworld*.

Bioética e Filosofia da Mente

Neste ponto, emerge aquela que talvez seja a questão central do presente trabalho: *seres racionais dotados de autoconsciência (por exemplo, andróides possuidores de IA) são dignos de consideração ética?* Essa indagação catalisa a discussão sobre as ações humanas, em relação aos andróides, sob uma perspectiva ética.

Em linhas gerais, pode-se dizer que a filosofia da mente não é propriamente uma disciplina, mas, primordialmente, um *estilo de filosofar* que ganha contornos mais nítidos nas últimas décadas, tendo como temas questões filosóficas tradicionais, como o problema das relações entre mente e corpo, o problema da natureza das representações mentais e o problema da identidade pessoal, apenas para citar algumas das questões tradicionais enfocadas nesse novo domínio⁸.

Como é produzida a visão das cores? Como o cérebro aprende e como ele retém as informações? Quais as interseções entre diferentes áreas cerebrais relacionadas à visão e suas respectivas contribuições funcionais? O que são representações e como o cérebro representa o mundo fora dele mesmo? O cérebro humano é mais complicado do que inteligente? Essas questões, independente se perguntadas por filósofos ou por neurocientistas, são parte da mesma investigação, em que algumas questões são facilmente habitadas simultaneamente na filosofia e na neurociência⁹.

O termo consciência é extremamente polissêmico, sendo utilizado em diversos contextos. A palavra consciência vem do termo latino "*consciētia*" que significa "conhecimento compartilhado com outros". No entanto, o termo consciência também se refere a vários outros aspectos que envolvem tanto seus níveis (sono, coma, sonhos e estado de vigília) quanto seus conteúdos (subjetivo, fenomenal e objetivo)¹⁰.

Com isso, é importante diferenciar, segmentar e conceituar corretamente os termos relacionados a esse campo de estudo. Os três descritores disponíveis no portal Descritores em Ciências da Saúde (DeCS), [<https://decs.bvsalud.org>], para o termo "Consciência" são: (1) **Awareness**, que possui a definição em português de: "O ato de levar em conta um objeto ou estado de coisas. Não implica em avaliação das qualidades ou da natureza do objeto, nem em atenção às mesmas.". Além disso, remete-se à tomada de decisões, estados de atenção e estar atento; (2) **Conscience**, "Consciência" na tradução literal para a língua portuguesa, traduzida

pelos processos cognitivos ou afetivos que compõem um governo moral internalizado, o qual atua sobre a conduta moral do indivíduo; (3) **Consciousness**, em que o descritor em português é “Estado de Consciência”, que se refere ao conhecimento de si mesmo e do ambiente.

Apesar da miríade de definições e de conceitos, as teorias a respeito da filosofia da mente parecem estar longe de chegar a um consenso quanto ao tema em questão. Em linhas gerais, diante do fenômeno da consciência, é debatido se ela pode ser reduzida e compreensível em forma de terceira pessoa, visto que os estados de consciência são, de maneira privilegiada, acessíveis apenas para o próprio sujeito; ou tal fenômeno é “objetivável”, passível de receber uma definição e uma explicação causal, necessariamente formuladas em terceira pessoa⁸.

A natureza da consciência é uma das questões mais duradouras e complexas da história das ciências naturais e também das humanidades. David Chalmers, filósofo e pensador sobre o tema, coloca que consciência dá origem a muitos “problemas fáceis” e um “problema difícil”: “Problemas fáceis” consistem em identificar mecanismos neurais e computacionais responsáveis por componentes da consciência, como a atenção, a diferença entre sono e vigília ou nossa capacidade de discriminar e reagir a estímulos ambientais. O chamado “problema difícil”, porém, consiste em compreender a experiência de ser um sistema consciente¹¹.

A noção de um *nível de consciência* é uma premissa fundamental na compreensão da consciência. Não é apenas um termo empregado para descrever os *estados globais de consciência* associados a distúrbios patológicos, como o estado comatoso, mas também desempenha um papel cada vez mais influente em contextos teóricos e metodológicos¹².

A reflexão sobre como o questionamento da própria realidade, através da aquisição plena da *consciência* pelos robôs *semi-humanos*, pode ser extremamente devastadora. No universo de *Westworld*, os androides são explorados pelos seus criadores em um “parque de diversão para adultos”, a fim de satisfazerem seus desejos, que são, muitas vezes, perversos. Por conseguinte, questiona-se: qual significado moral é atribuído por pessoas humanas a seres não orgânicos, possuidores de consciência, na forma de IA?

Ao longo da primeira temporada, a frase Shakespeariana “*these violent delights have violent ends*”, em português “*esses prazeres violentos têm fins violentos*” é usada como inspiração para mover os anfitriões do parque ao longo do caminho de tomada de consciência, à medida que se tornam mais semelhantes aos humanos e adquirem autonomia em suas escolhas.

Quais considerações morais, obrigações e direitos devemos à pessoa que não consegue expressar capacidades conscientes e que, por outro lado, permanecem ativas de acordo com descobertas neurocientíficas?¹³. O debate a respeito do tema não é corriqueiro. Contudo, o estímulo representado

pela obra é próprio para fomentar reflexões no âmbito legislativo e da sistemática envolvida no processo de tomadas de decisões em situações médicas e legais relacionadas ao *fim da vida*.

Diante dessa discussão, é de grande proficuidade entender o que queremos da tecnologia e o porquê; qual o seu papel para melhoria da qualidade da vida humana e em que momento os valores relacionados ao bem comum se distorcem e passam a servir interesses particulares e/ou minoritários. Este é o ponto central onde a discussão ético-filosófica reside: os objetivos de um projeto e suas possíveis consequências devem ser levados em consideração, e ampliar a consciência sobre tais feitos depende da capacidade de refletir sobre os mesmos.

Bases Neurobiológicas da Consciência Humana

A abordagem neurobiológica sobre a consciência, assim como outras propostas para explicar esse complexo fenômeno, destaca a integração entre diferentes segmentos do sistema nervoso (SN), suas integrações e interseções funcionais.

Os estados conscientes são frequentemente descritos como um mistério, visto que não existem elos explicativos claros que os conectam aos processos físicos e celulares. Sendo assim, tal fenômeno leva em conta experiências subjetivas conscientes em termos de certas características de estados cerebrais “subjacentes” e interações neurais; parecendo ser um problema intratável para investigações científicas¹⁴.

Para a construção dessa proposta de compreensão sobre o processamento da consciência, estão integrados, a partir de sua divisão anatômica, diferentes segmentos do sistema nervoso central (SNC) e periférico (SNP), considerando características estruturais, conexões anatômicas, funcionais e especificidades¹⁵.

Cada uma compoendo vias neurais de interesse que, apesar de guardarem características específicas quando integradas funcionalmente, contribuem para a elaboração mental dessa percepção dos compartimentos interno e externo, oferecendo, desta forma, um componente representativo para a adequada adaptação dinâmica e contínua ao meio.

Nesse contexto, com o objetivo de favorecer a construção do conhecimento, é proposta a apresentação de características funcionais dessas porções do SN, destacando suas contribuições.

Uma característica fundamental do SN para uma adaptação dinâmica dos seres vivos ao ambiente é a capacidade de sensibilização de receptores, estruturas componentes do SNP capazes de registrar variações específicas para diferentes estímulos externos (luz, sons, tato, sabores) e internos (pressão e gases sanguíneos, osmolaridade do sangue, dos tecidos orgânicos, fome, sede, entre outros). As alterações estruturais produzidas pelos estímulos nesses receptores são transformadas em potencial elétrico (transdução) que segue através de nervos até a porção central do SN^{15,16}.

Esses estímulos elétricos são conduzidos por trinta e um pares de nervos espinhais que fazem sinapse na medula espinhal e outros estímulos são conduzidos por doze pares de nervos cranianos ao tronco encefálico; a exceção está nos nervos olfatório e óptico (respectivamente primeiro e segundos pares de nervos cranianos) que fazem sinapse com o encéfalo¹⁶.

No tronco encefálico, essas aferências (*inputs*) são endereçadas a um conjunto de núcleos (corpos de neurônios no interior do tecido neural) dispostos em colunas envolvendo todos os seus segmentos (bulbo, ponte e mesencéfalo), relacionados a funções viscerais que, em conjunto, formam a Substância Reticular Ascendente (SRA)^{15,17}.

Essa estrutura desempenha participação determinante no mecanismo da consciência, considerando sua característica funcional de encaminhar estímulos excitatórios ao tálamo (porção do diencefalo subdividida em núcleos), que, dentre outras funções, está relacionado com a motricidade, visão, audição e ativação de regiões corticais que sinalizam as nossas emoções^{15,16}.

Além dessas participações funcionais, o tálamo recebe todos os estímulos sensitivos, exceto o olfato, e é responsável pela ativação de praticamente todas as áreas corticais funcionais primárias, distribuídas pelos diferentes lobos do telencéfalo como o córtex pré-frontal, responsável pelas habilidades executivas, planejamento, além de ativação de outras áreas corticais, a motora, no lobo frontal, a somatossensorial, no

lobo parietal, a visual, localizada no lobo occipital e auditiva, no lobo temporal¹⁸.

Essa última função talâmica justifica o seu envolvimento fisiopatológico em quadros neurológicos graves relacionados ao trauma, acidente vascular encefálico ou neoplasias. Nessas situações, a ausência de conexão com a SRA causa perda da consciência/coma e a presença de posturas patológicas, como a rigidez de decorticação e descerebração, considerando a localização da lesão em relação ao núcleo rubro do mesencéfalo^{19,20}.

O estado vegetativo é uma condição clínica frequentemente descrita como “vigília sem consciência”. Os pacientes acometidos abrem os olhos, frequentemente se movem de forma espontânea e exibem ciclos de sono e vigília, embora permaneçam totalmente não responsivos a qualquer forma de estímulo externo ou estimulação (além de simples reflexos). Com base nisso, presume-se que eles não têm qualquer *consciência*, incluindo quem são, onde estão e a situação em que se encontram²¹.

Outras contribuições cerebrais que merecem destaque por comporem a condição de percebermos o mundo de maneira consciente e a partir dessa perspectiva, fazermos escolhas e tomadas de decisões, correspondem à memória e ao sistema límbico como responsáveis pela modulação das emoções.

A possibilidade de a consciência ser “a função biológica crítica que nos permite saber que estamos sentindo tristeza ou alegria, sofrimento ou prazer, vergonha ou orgulho, pesar

por um amor que se foi ou por uma vida que se perdeu”⁵. E, a fim de que o nível da consciência englobe todas as nossas experiências, faz-se necessária a ação de neurônios que irão, em nível fisiológico e biomolecular, proporcionar a existência de uma *mente consciente*.

Nesse contexto, como apresentado no modelo de armazenamentos múltiplos²², os registros das experiências vividas internas e externas vividas, funcionariam, inicialmente, como aquisição e podem, na dependência dos significados, após a etapa de retenção temporária, ser consolidados, chegar à retenção duradoura ou esquecimento, e em outras situações, ser evocados, determinando potencialmente, mudanças no comportamento²³.

Quanto às emoções, em 1878, o médico, cirurgião, anatomista e antropólogo francês, Paul Pierre Broca descreve o grande lobo límbico, que recebeu essa denominação por sua topografia, na qual um conjunto de segmentos encefálicos circundavam estruturas cerebrais centrais. Em 1937, o médico e neuroanatomista americano James Papez propõe que um grupo de estruturas corticais, subcorticais e diencefálicas estariam relacionadas, a partir de uma circuitaria neural específica, à participação nas emoções. Porém, apenas em 1952, Paul D. MacLean adota o termo sistema límbico, relacionando-o ao controle das emoções e destacando, como porção central, o hipotálamo que, além de participar dessa sinalização, constitui o componente central do sistema nervoso autônomo e da regulação neuroendócrina através do eixo hipotálamo-hipófise^{24,25}.

A participação das emoções representa a condição de oferecer um viés afetivo, comportamental e, por isso, compõe o significado para as vivências nos processos cognitivos e do aprendizado em associação à memória na *composição da consciência*²⁶.

Outro fator importante é que, envolvidos pelas emoções, neurônios do hipotálamo, prosencéfalo basal e tronco cerebral liberam neurotransmissores em várias porções mais rostrais do cérebro e, assim, transformam temporariamente o modo de funcionamento de muitos circuitos neurais⁵. Há uma sensação de prazer ou desconforto, e até mesmo de aceleração ou desaceleração dos processos neurais como consequências típicas do aumento ou da diminuição na liberação dessas substâncias, sugerindo que o ser humano é um sistema biológico complexo, dinâmico e autorregulador, capaz de se adaptar de acordo com o meio em que está inserido⁵.

Torna-se válido também ressaltar a participação da linguagem na evolução e no funcionamento da mente consciente. A linguagem é o que mais radicalmente distingue a cognição humana da de outros animais, não representando apenas um sistema de comunicação duplamente articulado, mas também uma *faculdade cognitiva* que permite que os humanos expandam sua criatividade e conhecimento de forma *quase* infinita¹⁰.

Essa complexa e imbricada integração funcional entre diferentes segmentos e áreas corticais do SNC permite, a partir da análise neurobiológica, favorecer a compreensão

sobre os mecanismos neurais que subsidiam o entendimento sobre esse *intrigante* fenômeno, a consciência²⁷.

Consciência e Tecnologia

*"O pensamento consciente do homem não se apresenta como algo algorítmico."*²⁸

A consciência desempenha papéis importantes nos debates em torno do problema mente-corpo, nas controvérsias a respeito dos tipos de IA, bem como no campo da bioética. No entanto, apesar dessa enorme contribuição, a consciência não se encontra inserida nas discussões atuais sobre aspectos éticos da IA e da robótica²⁹.

Há de ser considerado que, apesar de estarem muito correlacionados, os conceitos de "*consciência*" e "*inteligência*" são distintos. Os conceitos de consciência já abordados neste trabalho, estão estritamente relacionados à emoção e à evocação de memórias. Contudo, a definição de inteligência, assim como o da consciência, também possui *inúmeros* significados e sentidos, podendo ser sintetizada como a capacidade de raciocínio, de planejamento, de resolução de problemas e de aprendizado a partir de experiências.

Com base nesses conceitos, pode-se implicar que os refinamentos no aprendizado de máquina, baseados na neurobiologia, permitiram que as redes neurais artificiais se aproximassem ou, ocasionalmente, ultrapassassem a inteligência humana³⁰.

Existem duas maneiras principais de pensar sobre a utilização e o poder atual da IA. A primeira é a hipótese da “IA fraca”, a qual afirma que uma máquina executando um programa é, no máximo, apenas capaz de simular o comportamento e a consciência humana real. IA, como a usada em diagnósticos médicos e certas intervenções, são exemplos de IA fraca porque se concentram em uma tarefa limitada. Como tal, esse tipo de IA justifica as afirmações de que o *programa* em execução é, no máximo, uma simulação de um processo cognitivo, mas não é um processo cognitivo em si. Em contraponto, a segunda forma, a “IA forte”, ainda não *concebida*, afirma que um programa em execução em uma máquina é na verdade uma mente e que não há diferença significativa entre um fragmento de *software* emulando exatamente as ações do cérebro e as ações de um ser humano³¹.

Até o momento, os sistemas de IA têm seguido os aspectos racionais de inteligência. Isso é importante para conferir eficiência diante do objetivo que é necessário alcançar, com a vantagem de não possuir *distratores*, como acontece com os seres humanos frente a variações emocionais, dor ou cansaço. Além disso, a IA pode combinar grandes variedades de algoritmos de maneira rápida e interativa, permitindo que o software aprenda com os padrões identificados no processo. A conexão cérebro-computador permite estimar a atividade cerebral, seguindo o processamento e criação de canais de comunicação com o meio ambiente. Assim, as máquinas tornam-se independentes, alcançando o seu próprio aprendizado e a resolução de problemas para aumento do desempenho^{32,33}.

A busca de resposta para os fenômenos da consciência, através das tecnologias informacionais, pode também ser motivada pela sua vindoura serventia nas ciências da saúde humana. Achados neurocientíficos apontam que as interfaces cérebro-computador, baseadas em ressonância magnética funcional, eletroencefalografia e/ou potenciais relacionados com eventos específicos, também podem ser uma estratégia viável para detectar vida consciente “não-descoberta” em pacientes que permanecem em um estado alterado de consciência. Tal detecção ocorre também quando esses pacientes saem de um período de coma secundário a uma lesão cerebral¹³.

O questionamento ecoante é: as máquinas serão, um dia, conscientes? Poderemos contar com uma *consciência-artificial* e com os andróides de *Westworld* com toda a sua senciência? Seria interessante para o ser humano que, um dia, as máquinas possam ter sentimentos e ter emoções? O desenvolvimento da IA possibilitará, no futuro, o desenvolvimento ou a simulação computacional da consciência?

Como o cérebro gerencia o significado é ainda desconhecido, mas é claro que o problema vai além do uso da linguagem e além dos humanos. Para desenvolver uma teoria do significado, mas deve ser conhecido sobre como os neurônios codificam e transformam os sinais sensoriais, sobre a base neural de memória, aprendizagem e emoção e sobre a interação dessas capacidades com o sistema motor³⁴.

Ao pensar em consciência artificial, enfrenta-se vários problemas. Fundamentalmente, existe a dificuldade de explicar a consciência em si e como a subjetividade pode emergir da matéria, sendo definida por David Chalmes, já citado anteriormente neste texto, como o *problema difícil da consciência*. Além disso, a compreensão da consciência humana é moldada por nossa própria experiência fenomenal e estritamente relacionada ao *self*. Há de ser considerado que o que se sabe sobre a consciência humana da perspectiva da primeira pessoa e a consciência artificial só será acessível da perspectiva da terceira pessoa, ou seja, por um observador externo. Relacionado a isso está a questão de como será possível saber se uma máquina tem ou não “consciência”^{11,29}.

A IA é frequentemente expandida pelo estudo e simulação da função neuronal especificamente e também promovida pela neurociência em geral. De forma recíproca, o progresso da pesquisa do cérebro também é inspirado pelos desenvolvimentos em IA. Um tema importante e um salto para a neurociência durante as últimas décadas, é o estudo de neurônios-espelho no cérebro de primatas, assim como as suas relações com a cognição e aprendizado em humanos. Com isso, a possibilidade de produzir neurônios espelho artificiais é um desafio adicional às fronteiras da *consciência-artificial* e esse conceito está sendo aplicado à IA para aumentar a sua sofisticação e sutileza³⁵.

E quando as máquinas atingirem o estado consciente: elas poderão ter direitos e considerações éticas? O campo da robótica deve lutar com essas perguntas: É ético criar e

continuar a usar robôs com consciência da maneira como eram usados os robôs originalmente projetados para as nossas necessidades? Os robôs merecem ser protegidos? Esse debate pode desafiar os limites da moralidade humana e polariza visões da sociedade sobre a possibilidade de os robôs conscientes serem objetos ou entidades vivas³⁶.

O debate biotecnológico evidenciado na primeira temporada da série também pode amparar a formação de valores e promover a discussão e o entendimento técnico-científico, principalmente em adultos jovens, que são o público-alvo do seriado.

CONSIDERAÇÕES FINAIS

As atuais perspectivas de evolução da robótica e da IA demandam robusta reflexão ética, a qual tem como pano de fundo a questão kantiana “O que é o homem?”. Trata-se de pensar o quanto andróides dotados de IA seriam próximos aos humanos, pois, é possível que ambos, por sua semelhança – quiçá lembrando Philip K. Dick –, preencham com *ovelhas elétricas* os próprios sonhos.

REFERÊNCIAS

1. Lawrence DR, Palacios-González C, Harris J. Artificial Intelligence. The Shylock Syndrome. *Camb Q Healthc Ethics* 2016;25:250-61. <http://dx.doi.org/10.1017/S0963180115000559>
2. Pugh J. *Autonomy, Rationality, and Contemporary Bioethics*. Oxford: Oxford University Press; 2020. <http://dx.doi.org/10.1093/oso/9780198858584.001.0001>

3. Esperidião-Antonio V, Majeski-Colombo M, Toledo-Monteverde D, Moraes-Martins G, Fernandes JJ, Assis MB, *et al.* Neurobiologia das emoções. *Rev Psiquiatr Clín* 2008;35:55-65. <http://dx.doi.org/10.1590/S0101-60832008000200003>
4. Adams B, Petruccione F. Quantum effects in the brain: A review. *AVS Quantum Sci* 2020;2:022901. <http://dx.doi.org/10.1116/1.5135170>
5. Damasio A. O mistério da consciência. São Paulo: Companhia das Letras; 2015.
6. Volkova PS. Harmonization of Consciousness: To the Question of Art Therapy. *Media Music J* 2017;8. http://mediamusicjournal.com/Issues/8_3.html
7. Carvalho-Nascimento EC, Silva E, Siqueira-Batista R. The "Use" of Sex Robots: A Bioethical Issue. *ABR* 2018;10:231-40. <http://dx.doi.org/10.1007/s41649-018-0061-0>
8. Tourinho CDC. A Filosofia da Mente Hoje. *In: Esperidião V, Siqueira-Batista R (Org.). Neurociências: diálogos e interseções.* Rio de Janeiro: Editora RUBIO, 2012, p.530-45.
9. Churchland P. *Neurophilosophy: Toward a Unified Science of the Mind/Brain.* Cambridge: The MIT Press; 1986.
10. Fabbro F, Cantone D, Feruglio S, Crescentini C. Origin and evolution of human consciousness. *Prog Brain Res* 2019;317-43. <http://dx.doi.org/10.1016/bs.pbr.2019.03.031>
11. Chalmers D. The Hard Problem of Consciousness. *In: Schneider S, Velmans M (eds). The Blackwell Companion to Consciousness.* New Jersey: John Wiley & Sons; 2017; p.32-42. <http://dx.doi.org/10.1002/9781119132363.ch3>
12. Bayne T, Hohwy J, Owen AM. Are There Levels of Consciousness? *Trends Cogn Sci* 2016;20:405-13. <http://dx.doi.org/10.1016/j.tics.2016.03.009>
13. Robles del Olmo B, García-Collado D. Retos éticos del hallazgo de conciencia encubierta con neuroimagen en estados vegetativos. *Med Clín* 2016;146:218-22.
<http://dx.doi.org/10.1016/j.medcli.2015.07.011>
14. Fazekas P, Overgaard M. Perceptual consciousness and cognitive access: an introduction. *Phil Trans R Soc B* 2018;373:20170340. <http://dx.doi.org/10.1098/rstb.2017.0340>
15. Frith CD. The neural basis of consciousness. *Psychol Med* 2019;51:1-13. <http://dx.doi.org/10.1017/S0033291719002204>
16. Kandel ER, Schwartz JH, Jessell TM. *Princípios de Neurociências. O sistema somatossensorial: receptores e vias centrais.* São Paulo: Editora AMGH; 2014.

17. Augustine JR. Chapter 9: The Reticular Formation. *In: Human Neuroanatomy*. 2nd ed. New Jersey: John Wiley & Sons. 2016; p.141-53.
18. Koch C, Massimini M, Boly M, Tononi G. Neural correlates of consciousness: progress and problems. *Nat Rev Neurosci* 2016;17:307-21. <http://dx.doi.org/10.1038/nrn.2016.22>
19. Whitney E, Alastrá AJ. *Neuroanatomy, Decerebrate Rigidity*. Treasure Island: StatPearls Publishing; 2020. <https://pubmed.ncbi.nlm.nih.gov/31613467/>
20. Iwaczuk W, Guniczak P. Neurophysiological foundations of sleep, arousal, awareness and consciousness phenomena. Part 1. *Anaesthesiol Intensive Ther* 2015;47:162-7. <http://dx.doi.org/10.5603/AIT.2015.0015>
21. Owen AM. The Search for Consciousness. *Neuron* 2019;102:526-8. <http://dx.doi.org/10.1016/j.neuron.2019.03.024>
22. Waugh NC, Norman DA. Primary memory. *Psychol Rev* 1965;72:89-104. <http://dx.doi.org/10.1037/h0021797>
23. Lent R. *Cem Bilhões de Neurônios*. São Paulo: Atheneu; 2010.
24. Torricco TJ, Abdijadid S. *Neuroanatomy, Limbic System*. Treasure Island: StatPearls Publishing; 2020. <https://www.ncbi.nlm.nih.gov/books/NBK538491>
25. Raj Mohan V, Mohandas E. The limbic system. *Indian J Psychiatr* 2007;49:132. <http://dx.doi.org/10.4103/0019-5545.33264>
26. Catani M, Dell'Acqua F, Thiebaut de Schotten M. A revised limbic system model for memory, emotion and behaviour. *Neurosci Biobehav Rev* 2013;37:1724-37. <http://dx.doi.org/10.1016/j.neubiorev.2013.07.001>
27. Sousa A. Towards an integrative theory of consciousness: Part 1 (Neurobiological and cognitive models). *Mens Sana Monogr* 2013;11:100. <http://dx.doi.org/10.4103/0973-1229.109335>
28. Penrose R. *The Emperor's New Mind - Concerning Computers, Minds and the Laws of Physics*. New York: Penguin; 1991.
29. Hildt E. Artificial Intelligence: Does Consciousness Matter? *Front Psychol* 2019;10:1535. <http://dx.doi.org/10.3389/fpsyg.2019.01535>
30. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, et al. Mastering the game of Go with deep neural networks and tree search. *Nature* 2016;529:484-9. <http://dx.doi.org/10.1038/nature16961>
31. Scerri M, Grech V. Artificial intelligence in medicine. *Early Human Development* 2020;145:105017. <http://dx.doi.org/10.1016/j.earlhumdev.2020.105017>

- 32.Klein E. Neuromodulation ethics: Preparing for brain-computer interface medicine. *In*: Illes J. Neuroethics: anticipating the future. Oxford: Oxford Scholarship Online. Oxford University Press; 2017. <http://dx.doi.org/10.1093/oso/9780198786832.003.0007>
- 33.Wolpaw J, Wolpaw EW. Brain-Computer Interfaces. Principles and Practice. Oxford: Oxford University Press; 2012. <http://dx.doi.org/10.1093/acprof:oso/9780195388855.001.0001>
- 34.Churchland P, Churchland PS. Could a Machine Think? *Sci Am* 1990;262:32-9. <https://www.jstor.org/stable/24996642>
- 35.Shapshak Paul. Artificial Intelligence and brain. *Bioinform* 2018;14:38-41. <http://dx.doi.org/10.6026/97320630014038>
- 36.Spatola N, Urbanska K. Conscious machines: Robot rights. *Science* 2018;359:400. <http://dx.doi.org/10.1126/science.aar5059>