



## CAPÍTULO 3

### REFLEXÕES SOBRE UMA POSSÍVEL INTELIGÊNCIA ARTIFICIAL SUBCONSCIENTE

**Renato Pincelli**

Mestre pelo Programa de Pós-Graduação em Mídia e Tecnologia da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Arquitetura, Artes, Comunicação e Design - Câmpus de Bauru.

**Marcos Américo**

Docente e pesquisador do Programa de Pós-Graduação em Mídia e Tecnologia (Doutorado) na Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Arquitetura, Artes, Comunicação e Design - Câmpus de Bauru.

**Francisco Rolfsen Belda**

Docente e pesquisador do Programa de Pós-Graduação em Mídia e Tecnologia (Doutorado) na Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Arquitetura, Artes, Comunicação e Design - Câmpus de Bauru.

#### RESUMO

Há mais de meio século que cientistas da computação vêm desenvolvendo sistemas de Inteligência Artificial (IA) em busca de algo que se assemelhe a uma consciência como a humana. Várias abordagens foram utilizadas, mas a mais bem-sucedida é o chamado Aprendizado de Máquina, que vem sendo amplamente aplicado nos últimos anos. Há várias discussões sobre o que seria uma consciência artificial, mas é possível que já estejamos testemunhando a emergência de um estágio intermediário, de subconsciência artificial. Neste texto apresentamos um panorama histórico da IA, definições de subconsciente e relatamos alguns casos que podem indicar o surgimento de subconscientes artificiais como efeito colateral do Aprendizado de Máquina. Por fim, encerraremos com uma breve discussão sobre os impactos desse fenômeno emergente no futuro da IA.

**Palavras-chave:** inteligência artificial, IA subconsciente, aprendizado de máquina.

#### 1 INTRODUÇÃO

A primeira década e meia do século XXI pode ser caracterizada em termos tecnológicos pelo rápido desenvolvimento e utilização de Inteligências Artificiais (IA). Da organização de filmes e músicas em serviços de streaming como *Netflix* e *Spotify* às pesquisas clínicas e militares, há IAs por toda parte – tanto que muitas passam despercebidas pelo usuário comum. Essas aplicações práticas são apenas um objetivo secundário nas pesquisas em IA, cuja principal meta, desde seu surgimento, nos anos 1950, tem sido a criação de uma consciência artificial, uma entidade virtual inteligente e indistinguível de uma pessoa comum conforme o Teste de Turing. Entretanto, o desenvolvimento de redes neurais profundas operadas por meio de Aprendizado de Máquina, definido como a programação de um sistema para que possa aprender com suas próprias experiências, chegando sozinho à solução de problemas propostos, muitas vezes encontrando algoritmos de solução mais eficientes ou mesmo inéditos. Este conceito tem levado a consequências inesperadas, como máquinas que não sabem justificar os motivos de

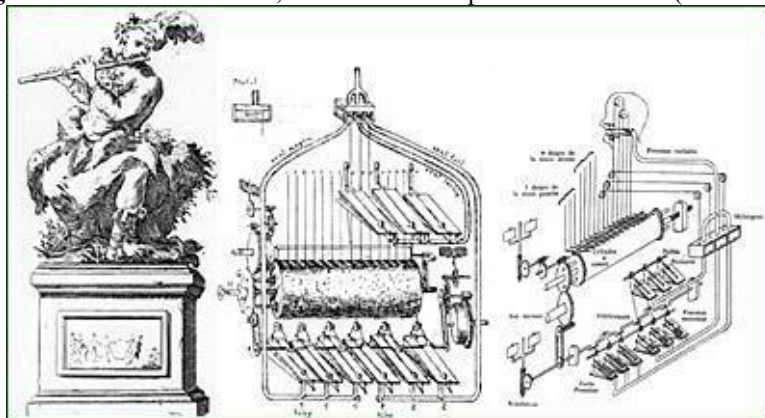


suas respostas ou mesmo que parecem sonhar. Nesse contexto, podemos testemunhar um fenômeno que podemos denominar “Inteligência Artificial Subconsciente”.

## 2 INTELIGÊNCIA ARTIFICIAL: CONTEXTO HISTÓRICO E DEFINIÇÕES

As raízes históricas da Inteligência Artificial (IA) remontam a René Descartes, que via animais e pessoas como autômatos naturais. No entanto, como lembra Gardner (2003), Descartes era cético quanto à simulação da mente. Ainda no século XVII começam os desenvolvimentos práticos das teorias cartesianas, como os autômatos de Jacques de Vaucanson (1709-1782) configurados como tocadores de flauta (Figura 1), de tambor e patos mecânicos que pareciam comer grãos, digeri-los e defeca-los. Esses mecanismos deveriam ter apenas a aparência de inteligência e não ser efetivamente inteligentes.

**Figura 1** - Tocador de flauta, autômato de Jacques de Vaucanson (1709-1782).



Fonte: Internet.

No século XIX, os avanços tecnológicos e matemáticos permitiram que Charles Babbage pudesse desenvolver uma complexa máquina capaz de realizar cálculos. Ainda que não tenha sido concluída, a Máquina de Babbage é considerada o primeiro computador moderno. Para Ada Lovelace, tal máquina poderia ser programada para ir além dos cálculos numéricos, sendo capaz de jogar xadrez. É o que podemos chamar de virtualização dos velhos autômatos.

Entre o século XIX e o começo do século XX, outros desenvolvimentos teóricos vieram de George Boole, Alfred North Whitehead e Bertrand Russell. Com base no estudo minucioso da lógica e do pensamento, Boole criou o que Gardner (2003) descreve como uma espécie de “álgebra mental”, onde tudo poderia ser expresso em termos de sim ou não, 0 ou 1. A lógica booleana, refinada por Whitehead e Russell, seria a base da programação dos computadores modernos e, mais tarde, das IAs.



Mas o que é uma IA? De modo geral, uma IA pode ser definida como um programa ou máquina cujo processamento de dados se assemelha ao modo de pensar dos seres humanos. No entanto, o consenso termina aí. Segundo Gardner (2003, p.155), algumas:

[...] definições enfatizam a criação de programas; outras se concentram em linguagens de programação; outras abrangem o hardware mecânico e o componente conceitual humano, assim como o software. Alguns praticantes querem simular os processos humanos de pensamento exatamente, enquanto outros se contentam com qualquer programa que leve a consequências inteligentes (GARDNER, 2003, p.155).

Consequentemente surgem dois campos, conhecidos como IA Fraca e IA Forte. A IA Fraca busca replicar metodicamente o raciocínio humano para testar teorias psicológicas e evolutivas. Nesse caso, a IA seria apenas uma ferramenta de pesquisa, uma espécie de rato de laboratório. A IA Forte, por sua vez, tem o objetivo explícito de criar sistemas dotados de consciência, indistinguíveis de um ser humano de acordo com o teste de Turing. Essa abordagem vê o desenvolvimento de IAs não como um meio de estudo mas como um fim em si mesmo. De modo geral, as IAs que descreveremos neste atexto podem ser classificadas como do tipo forte.

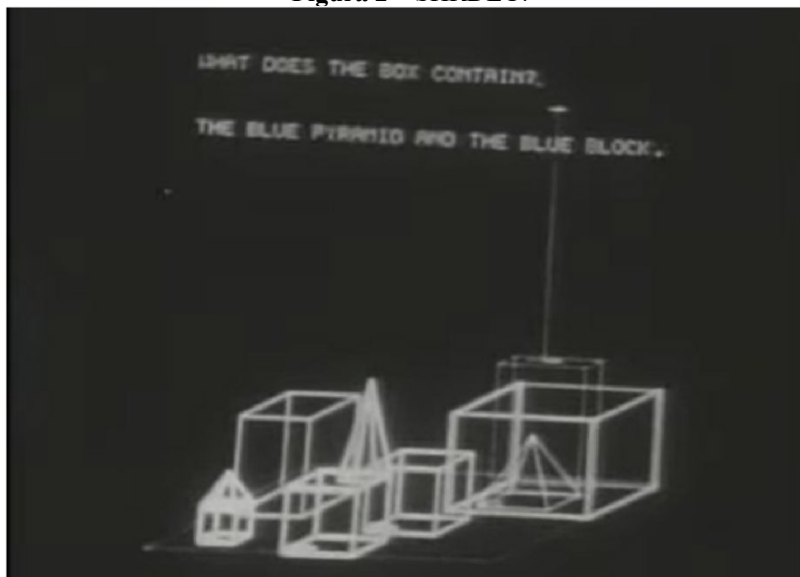
De volta ao desenvolvimento histórico, surgem grandes avanços em meados do século XX. Em 1938, Claude Shannon demonstra que circuitos eletrônicos podem ser programados e operados segundo a lógica booleana. Cinco anos mais tarde, Warren McCulloch e Walter Pitts dão os primeiros passos no desenvolvimento de redes neurais. No entanto, Gardner põe o marco fundamental das pesquisas em IA no ano de 1956, quando foi realizada a Oficina de Dartmouth. Nesse evento, pesquisadores como Herbert Simon, Allen Newell, Marvin Minsky e John McCarthy discutem a possibilidade da capacidade de programas de computador se comportarem ou pensarem de modo inteligente. Atribui-se a McCarthy a criação do termo Inteligência Artificial tal como o conhecemos.

Ainda em 1956, Newell e Simon criam o programa *Logic Theorist* (LT), dedicado a descobrir provas para teoremas de lógica simbólica. Capaz de resolver 38 dos 52 teoremas do capítulo 2 de *Principia Mathematica*, de Whitehead e Russell, o LT pode ser considerado o primeiro software dotado de IA. Nos anos 1960, Minsky atua como orientador em projetos que resultam em sistemas capazes de fazer analogias visuais e resolver problemas de álgebra expostos de maneira linguística. Enquanto isso, McCarthy cria a LIST, uma linguagem de computação que será fundamental no campo da IA. Recursiva, hierárquica e flexível, a LISP e suas derivadas continuam a ser usadas por cientistas da computação especializados em pesquisas cognitivas.



Em 1970, Therry Winograd desenvolve o experimento SHRDLU (Figura 2), um programa capaz de seguir instruções para manejar blocos de construção num ambiente virtual. Primeira IA capaz de interagir com o mundo de maneira mais ou menos concreta, SHRDLU foi notável por ser capaz de fazer algumas interpretações de linguagem, sendo capaz de distinguir entre perguntas, instruções e comentários de seus utilizadores.

**Figura 2 – SHRDLU.**



**Fonte:** captura de tela de vídeo no YouTube.

Talvez o mais antigo sistema de Aprendizado de Máquina seja o PERCEPTRON, desenvolvido entre 1958 e 1962 por Frank Rosenblatt no *Cornell Aeronautical Laboratory* com o propósito de reconhecer padrões através de seus “olhos”, formados por uma rede de 400 fotocélulas. O processamento da imagem era feito por um conjunto de elementos associadores. Inicialmente, as conexões eram feitas ao acaso. Minsky criticou essa abordagem, considerando necessário o uso de um mecanismo de feedback, que informasse a máquina sobre seus erros e acertos. No começo dos anos 1970, Patrick Winston desenvolve um programa capaz de distinguir quando os blocos formavam um arco e quando não. Além de arcos, o sistema aprendeu a identificar mesas, pedestais e arcadas após ser exposto a uma série de exemplos e contraexemplos.

Esses avanços não significam que o desenvolvimento de IA tenha sido recebido entusiasticamente e livre de críticas na comunidade acadêmica. Uma das críticas mais contundentes à IA foi feita por Hubert Dreyfus em 1972. O filósofo da vertente fenomenológica argumentava que, diferente das pessoas, computadores não são capazes de lidar com subjetividades ou ambiguidades porque não tem sequer o que chama de consciência periférica



(*fringe*). Outra crítica bastante dura foi da de John Searle, que atacou o conceito de IA forte em 1980 com o experimento mental do Quarto Chinês.

Basicamente, Searle se imagina trancado num quarto onde recebe conjuntos de caracteres chineses e instruções para correlacioná-los. As instruções estão em inglês, língua que Searle compreende. Assim, mesmo sem entender uma palavra de chinês, Searle poderia passar por um falante nativo depois de certo tempo de treinamento sob essas condições. Para ele, isso é evidência de que IAs fortes não têm mente, sendo apenas autômatos muito bem elaborados. Críticas como essa e muitas outras continuaram a ser feitas dos anos 1970 a meados dos anos 80, um período de crise que, aliado a problemas como falta de financiamento, ficou conhecido como Inverno da IA.

Essas limitações, especialmente as técnicas, começariam a ser superadas com o surgimento da Internet, nos anos 1990. Como lembra Knight (2017), foi só a partir daí que o Aprendizado de Máquina e o desenvolvimento de redes neurais profundas se tornaram possíveis, graças à informatização e ao crescente volume de dados disponibilizados on-line. Antes disso, quem quisesse propor uma IA classificadora de música, por exemplo, precisaria expô-la a músicas dispersas entre diversas mídias físicas, como LPs e fitas cassete, armazenadas em muitos lugares diferentes e muitas vezes inacessíveis. Com a música digitalizada isso se tornou mais fácil e uma IA capaz de nos dar sugestões musicais (como a do *Spotify*) pode passar quase despercebida.

No entanto, Knight também lembra que um sistema de aprendizado profundo, capaz de aprender e/ou se programar sozinho é, por sua própria natureza, uma caixa-preta bem obscura. Nesse ponto, surgem IAs insondáveis como o subconsciente humano.

### 3 SUBCONSCIENTE OU INCONSCIENTE

Desde que começou a ser estudado por Sigmund Freud no começo do século XX, o subconsciente ou inconsciente tem sido objeto de múltiplas definições no meio acadêmico. Mesmo na linguagem do dia-a-dia, este é um termo de difícil definição: Holanda (2010) ressalta a existência de nada menos que 11 acepções de inconsciente no Dicionário Houaiss da Língua Portuguesa. Dadas nossas limitações de tempo e espaço, vamos apresentar aqui apenas cinco visões distintas sobre esse fenômeno: a de Freud, a de Lacan, a de Merleau-Ponty, a de Vigotski e a de Damásio.

Castro (2009) faz em seu artigo uma comparação entre o inconsciente dos pontos de vista de Freud e Lacan. Segundo este autor, o inconsciente descoberto por Freud era bem



distinto do “intuído até então pelos poetas e filósofos — um reino obscuro e caótico, situado fora dos limites da consciência”. Freud baseia seu estudo do inconsciente na interpretação dos sonhos. No entanto, ele não foge às concepções um tanto mecanicistas de sua época e considera o inconsciente como uma espécie de reservatório de emoções reprimidas, que seria aberto durante os sonhos.

Para Lacan, esta visão dos psicanalistas seria um tanto obscurantista, por não “distinguirem o inconsciente do instinto, ou como eles dizem, do instintual”. Ainda assim, o inconsciente não seria caracterizado pela falta de conhecimento: “O inconsciente não é perder a memória; é não se lembrar do que se sabe” (Lacan *apud* CASTRO, 2009). Visando corrigir o que considera obscuro na psicanálise, Lacan busca aplicar conceitos linguísticos ao estudo do inconsciente. Para ele, o inconsciente é tão dotado de linguagem e seus símbolos quanto a consciência. A diferença é que no inconsciente a linguagem seria inarticulada ou articulada de modo irreconhecível à primeira vista, com relações tão enigmáticas quanto os hieróglifos.

Posição radicalmente oposta tanto a Freud quanto a Lacan é a de Merleau-Ponty, que em 1954 questiona o inconsciente freudiano como um “segundo sujeito pensante” que seria desnecessário. Segundo Holanda (2010), Merleau-Ponty considera a consciência como algo sempre perceptivo, sempre ligado ao corpo de maneira inseparável e em constante diálogo com o mundo: “não há o que buscar no conceito de inconsciente, dado que o corpo é o centro de referências da existência e o espaço onde a existência é vivida, e onde convergem o ‘pré-pessoal’ e o ‘pessoal’.” Holanda conclui seu artigo afirmando que, para o filósofo francês, “o inconsciente não deve ser procurado no ‘fundo’ de nós mesmos, mas diante de nós, em articulação com o mundo”, onde existiriam apenas o visível e o invisível.

Outra postura distinta é a desenvolvida por Vigotski que, segundo Pessanha (2015), via o inconsciente como uma questão puramente teórica, filosófica, gnosiológica e não como algo com existência concreta e empírica. Para o pesquisador russo consciência e psique são as mesmas coisas e não sobra espaço para o inconsciente. “Psique inconsciente” seria algo tão contraditório quanto “quadrado redondo”. Vigotski via a inconsciência como algo que, nas palavras de Pessanha, existiria apenas como algo para “preencher os limites, os saltos, as omissões e distorções da consciência”. Em vez de inconsciente, Vigotski fala em comportamentos não-verbais e em automatismos. Por exemplo, não temos lembranças da primeira infância porque o pensamento nessa fase não é verbalizado. Quando aprendemos algo pela primeira vez, como dar um nó, o fazemos com plena consciência, prestando atenção a



todos os detalhes de um ato que, com o passar do tempo, torna-se automático e cada vez menos consciente mas não inteiramente inconsciente.

Uma abordagem mais recente é a de Antonio Damásio (2000), que busca evitar a ambiguidade dos termos subconsciente e inconsciente. Para o autor, o inconsciente se resume a um conjunto de processos e conteúdos que passam despercebidos, como as imagens para as quais não atentamos e as habilidades que adquirimos pela experiência mas tornam-se dormentes e automáticas com o passar do tempo (Damásio, 2000, p. 292).

Partindo de uma abordagem neurológica e evolucionária, Damásio fala em proto-self, algo que define como “um conjunto coerente de padrões neurais que mapeiam, a cada momento, o estado da estrutura física do organismo em suas numerosas dimensões” (Damásio, 2000, p. 201). Esse protótipo de self não seria dotado nem de linguagem nem de capacidade de percepção. Por isso, não somos conscientes do proto-self. Seria, portanto, uma forma rudimentar de (auto)consciência.

Entretanto, Damásio também reconhece a importância de formas de pensamento inconsciente. Como exemplo, cita as crianças que, aos três anos de idade, “empregam de um modo espantoso as regras de construção de sua língua, mas não sabem que possuem esse 'conhecimento'” (Damásio, 2000: 375).

#### 4 SUBCONSCIENTE ARTIFICIAL

O desenvolvimento de IAs tem se acelerado nos últimos anos, sobretudo com a aplicação dos métodos de Aprendizado de Máquina e sua versão mais radical, o Aprendizado Profundo. Como já citado, o Aprendizado de Máquina consiste na programação de um sistema para que possa aprender com suas próprias experiências, chegando sozinho à solução de problemas propostos, muitas vezes encontrando algoritmos de solução mais eficientes ou mesmo inéditos. No Aprendizado Profundo, a IA é abastecida apenas com uma fonte de informações – como um banco de dados – e alguns algoritmos simples, devendo ser capaz de aprender por si mesma, sem qualquer programação prévia ou supervisão externa.

Como nota Knight (2017), a aplicação cada vez mais generalizada dessas duas abordagens têm levado a IAs cada vez mais sofisticadas e complexas. O autor relata alguns casos em que nem os próprios programadores sabem explicar exatamente o funcionamento de IAs baseadas em Aprendizado Profundo. Um exemplo é o sistema conhecido como *Deep Patient*, uma IA especializada em pesquisa médica criada pelo Hospital Monte Sinai, nos EUA. Abastecido com dados de prontuários de 700 mil pacientes, *Deep Patient* tem se mostrado

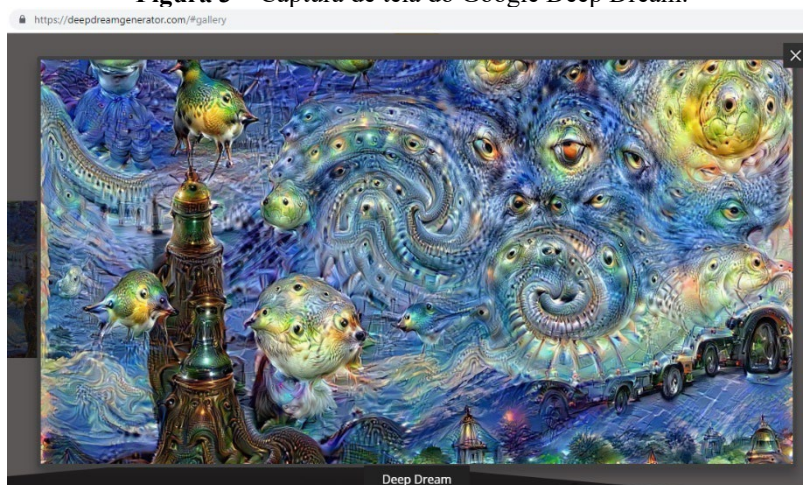


bastante preciso no diagnóstico de doenças. O que surpreendeu os pesquisadores é que esse sistema também tem conseguido antecipar casos de esquizofrenia. No entanto, ninguém sabe ao certo como o Deep Patient faz isso e essa IA tampouco é capaz de explicar seu raciocínio. É quase como se o reconhecimento profundo de padrões envolvesse algo como uma intuição subconsciente.

Para Knight, “por sua própria natureza, o aprendizado profundo é uma caixa-preta bem obscura.” Em seu artigo esse autor explica que não é possível observar o passo a passo do raciocínio de uma inteligência artificial que funciona por aprendizado profundo. Como não há programação manual envolvida – o que seria muito complexo – os programadores desses sistemas não sabem localizar os motivos dos insights de uma IA desse tipo. Estando inacessível à consciência tanto dos pesquisadores quanto da máquina, essa forma de pensar artificialmente poderia ser classificada como subconsciente?

Até certo ponto sim, mas isso depende da definição de subconsciente adotada. Do ponto de vista freudiano que apresentamos anteriormente, que vê o subconsciente como algo onírico, existe um sistema que pode ser considerado subconsciente. Trata-se do *Deep Dream*, desenvolvido pelo Google com o propósito de entender como se dá o reconhecimento de imagens em IAs. Segundo Knight, o *Deep Dream* (Figura 3) é um sistema de aprendizado profundo que gera ou modifica as próprias imagens, as quais “apresentam animais grotescos, de aparência alienígena, emergindo de nuvens e plantas ou pagodes alucinatórios brotando através de florestas e serras”. São, de fato, como sonhos artificiais, o que caracteriza um subconsciente do tipo freudiano.

**Figura 3** – Captura de tela do Google Deep Dream.



**Fonte:** Internet.

Talvez o conceito de subconsciente que mais se aproxima de alguns resultados observados no desenvolvimento de IAs seja o de Lacan. É aquele tipo de subconsciente que não





sabe explicar seu raciocínio ou o faz de maneira inarticulada, algo bem parecido com o que aconteceu no caso do *Deep Patient*. Se seguirmos Damásio, diríamos que esse sistema é dotado de um proto-self, responsável por regular funções vitais de forma inconsciente.

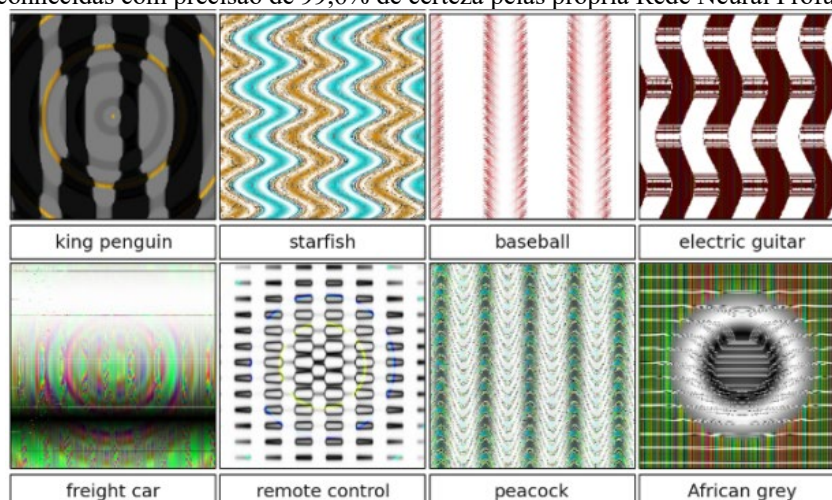
Outro exemplo de IA que podemos considerar subconsciente é o sistema desenvolvido por Jeff Clune, da Universidade de Wyoming, e seus colaboradores. Em 2015, segundo Knight (2017), Clune et. al. conseguiram aplicar o equivalente a ilusões de óptica a Redes Neurais Profundas. Porém, essa é uma aplicação de IA fraca, já que foi criada com o propósito de compreender as ilusões de ótica. Mesmo assim, essa pesquisa revelou que é o processamento intermediário de padrões fracos ou inexistentes que criam as ilusões. Processos similares, subliminares, também poderiam acontecer conosco.

As pesquisas de IA no campo militar também tem resultado em sistemas que beiram a insondabilidade. Como explica Knight, “os soldados provavelmente não vão se sentir confortáveis num tanque robótico que não se explique para eles e os analistas vão resistir em agir com base numa informação sem uma boa razão.” Para resolver isso, a DARPA (agência de pesquisas do departamento de defesa dos EUA), está desenvolvendo uma IA capaz de se explicar aos seus usuários.

Isso pode não ser inteiramente possível. Mesmo quando pedimos para alguém nos explicar porque fez algo, essa pessoa pode não ser capaz de nos dizer por que agiu de determinada maneira – ou pode ter razões para não querer nos revelar seus motivos. Clune (*apud* Knight) indica que “pode ser que seja da natureza da inteligência ter apenas parte dela exposta à explicação racional. Outras partes podem ser apenas instintivas ou subconscientes ou insondáveis.” Clune et al.(2015) ainda utilizam uma Rede Neural Profunda para demonstrar como a IA se aplica algoritmos evolutivos para reconhecimento de padrões de imagens (Figura 4) que são irreconhecíveis para os humanos mas reconhecidas com precisão por máquinas que as consideram objetos naturais. Desta forma, Redes Neurais Profundas são facilmente enganadas. O estudo esclarece diferenças interessantes entre visão humana e Redes Neurais Profundas atuais e levanta questões sobre a generalidade da visão computacional de uma Rede Neural Profunda.



**Figura 4** – Imagens geradas por Rede Neural Profunda que são irreconhecíveis para seres humanos, mas reconhecidas com precisão de 99,6% de certeza pelas própria Rede Neural Profunda.



**Fonte:** Clune et al. (2015).

Nesse ponto, surgem problemas de ordem filosófica e ética. Tendo ou não um subconsciente, IAs profundas passam a exigir algum grau de confiança ou mesmo de fé para serem utilizadas. Isso pode não ser tão grave quando um aplicativo de celular lhe recomenda um novo restaurante ou uma nova banda, mas se complica se complica quando há IAs do mesmo tipo tomando decisões importantes nos campos militar, econômico ou medicinal, mas incapazes de se explicar totalmente.

Para Knight, isso não significa que devemos abrir mão do desenvolvimento ou do uso cada vez mais cotidiano de IAs. Porém, as novas questões que surgem junto com a emergência do que chamamos de IA subconsciente são tão profundas e complexas quanto às redes neurais em que se baseiam. Tudo indica que estamos entrando numa fase intermediária, situada entre IAs especializadas porém inconscientes e IAs capazes de conhecimento universal e algum grau de consciência.

As discussões decorrentes da utilização do *ChatGPT*, a partir do início de 2023, evidenciam uma nova era de interrogações sobre o desenvolvimento da IA. Para compreender esta ferramenta foi feita a pergunta “O que é o *ChatGPT*?” ao próprio algoritmo baseado em inteligência artificial e a resposta obtida está na Figura 5:



**Figura 5** – Resposta do *ChatGPT* para a pergunta “O que é o *ChatGPT*?”.



o que é o ChatGPT



O ChatGPT é um modelo de linguagem natural baseado na arquitetura GPT-3.5 da OpenAI. Ele foi treinado em grandes quantidades de dados textuais para gerar respostas precisas e relevantes a perguntas e interações com usuários. O ChatGPT é capaz de compreender e produzir linguagem natural em diversos idiomas e é utilizado em uma variedade de aplicações, desde assistentes virtuais até chatbots de atendimento ao cliente.

**Fonte:** *ChatGPT*.

Embora o próprio *ChatGPT* afirme que não possui consciência quando perguntado (Figura 6) sua existência tem causado fortes reações na sociedade em relação aos seus riscos e falta de transparência.

**Figura 6** – Resposta do *ChatGPT* para a pergunta “*ChatGPT* tem consciência?”.



Chat GPT tem consciência?



Como um modelo de linguagem treinado em dados, eu não tenho consciência ou capacidade de compreender minha própria existência ou do mundo ao meu redor da mesma forma que os seres humanos. Eu sou uma criação humana que usa algoritmos complexos para gerar respostas de texto com base nas entradas fornecidas pelos usuários. Embora eu possa simular uma conversa humana e fornecer informações úteis, eu não tenho consciência ou sentimentos.

**Fonte:** *ChatGPT*.

Elon Musk e um grupo de intelectuais e especialistas em tecnologia têm expressado preocupações em relação ao desenvolvimento da IA e aos potenciais riscos associados a ela (Figura 7). O empresário acredita que a IA pode se tornar uma ameaça existencial para a humanidade se não for regulamentada e desenvolvida de maneira responsável.



**Figura 7** - Elon Musk expressa preocupações em relação ao desenvolvimento da inteligência artificial (IA).



**Fonte:** Internet.

Musk acredita ainda que a IA pode se tornar por demais avançada e superar a capacidade humana de controlá-la e tem alertado sobre os possíveis perigos de um cenário em que as máquinas se tornam mais inteligentes do que os seres humanos e com capacidade de decidir agir de maneira contrária aos nossos interesses. Além disso, destaca a possibilidade de uma corrida armamentista por meio da IA, em que países e empresas poderiam desenvolver sistemas autônomos de armas que causariam danos irreparáveis.

Embora Musk não tenha necessariamente sugerido que a IA deva ser completamente interrompida, ele tem pedido por uma regulamentação mais forte e um desenvolvimento mais responsável da tecnologia. O bilionário ajudou a lançar a *OpenAI*, uma organização de pesquisa de IA com o objetivo de promover o desenvolvimento seguro e responsável da tecnologia, e tem instado os governos e outras empresas a agir de forma mais cautelosa em relação à IA.

## 5 CONSIDERAÇÕES FINAIS

Sabemos que a IA tem o potencial de afetar os empregos dos humanos, principalmente em áreas onde há um alto grau de automação e repetição de tarefas. À medida que ela se torna mais avançada, pode substituir muitos empregos que antes eram realizados por seres humanos, como por exemplo, atendimento ao cliente, tarefas administrativas e operações de fábrica. Porém, é importante destacar que a IA também tem o potencial de criar novos empregos e oportunidades em áreas como programação, desenvolvimento de algoritmos, análise de dados, entre outras áreas e pode tornar os trabalhos existentes mais eficientes, o que liberaria tempo para os trabalhadores se concentrarem em tarefas de maior valor agregado.



Ainda que a IA tenha o potencial de afetar negativamente muitos empregos, ela também pode criar novas oportunidades e melhorar a eficiência geral das empresas e dos trabalhadores. Desta forma, faz-se mister que instituições públicas e privadas estejam cientes desses desafios e oportunidades, e trabalhem de forma coletiva para garantir uma transição justa e sustentável para uma economia mais automatizada

Acreditamos que esses desdobramentos (e outros que podem surgir) devem considerar um desenvolvimento mais ponderado, numa velocidade controlada de inovação, capaz de permitir uma acomodação entre nós, seres humanos e as IAs em campos tão diversos quanto relações sociais, econômicas, pesquisas médicas, criações culturais e questões legais. Já começamos a dialogar com IAs, algo tão inédito que talvez seja apenas comparável a um contato com uma civilização extraterrestre inteligente. Em ambos os casos, antes de aprendermos a nos comunicar uns com os outros, precisaremos desenvolver meios de entendimento mútuo. Saber como se comportam essas novas Inteligências Artificiais será uma tarefa indispensável.

## REFERÊNCIAS

BENGIO, Y. **Learning Deep Architectures for AI**. Foundations and Trends® in Machine Learning, 2(1), 1–127, 2019. DOI:10.1561/22000000006

CASTRO, Júlio Cesar Lemes de. **O inconsciente como linguagem: de Freud a Lacan**. Cadernos de Semiótica Aplicada, vol. 7, n. 1, julho de 2009. Disponível em: <<http://seer.fclar.unesp.br/casa/article/view/1773>>. Acesso em: 15 de junho de 2020.

CLUNE J. NGUYEN A, YOSINSKI J, **Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images**. In Computer Vision and Pattern Recognition (CVPR '15), IEEE, 2015. Disponível em: <[http://www.evolvingai.org/files/DNNsEasilyFooled\\_cvpr15.pdf](http://www.evolvingai.org/files/DNNsEasilyFooled_cvpr15.pdf)> Acesso em: 18 de outubro de 2021.

DAMÁSIO, António. **O mistério da consciência: do corpo e das emoções ao conhecimento de si**. São Paulo: Companhia das Letras, 2000.

GARDNER, Howard. **A Nova Ciência da Mente: uma história da revolução cognitiva**. São Paulo: Editora da Universidade de São Paulo, 2003 (3a. Edição). pp. 153-197.

HOLANDA, Adriano. **Notas para uma reflexão sobre consciência e inconsciente na fenomenologia**. Rev. abordagem gestalt., Goiânia, v. 16, n. 1, p. 45-53, jun. 2010. Disponível em <[http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1809-68672010000100006&lng=pt&nrm=iso](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1809-68672010000100006&lng=pt&nrm=iso)>. Acesso em: 15 de junho de 2022.

KNIGHT, Will. **The Dark Secret at the Heart of AI**. MIT Technology Review, 11 de abril de 2017. Disponível em: <<https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/>> Acesso em 15/05/2017. Versão traduzida pelo autor desse artigo disponível em:



<<https://medium.com/@mntpincelli/existe-um-segredo-no-coraçao-da-inteligencia-artificial-8704375fd8f8>>. Acesso em: 18 de maio de 2021.

PESSANHA, P. H. C. **O inconsciente na psicologia histórico-cultural de Vigotski: um estudo conceitual**. 2015. 101 f. Dissertação submetida para a Banca de Defesa do Programa de Mestrado em Psicologia (PPGPSI) da Universidade Federal de São João Del-Rei (UFSJ). São João Del-Rei, 2015. Disponível em: <[http://www.ufsj.edu.br/portal2-repositorio/File/mestradosicologia/2016/Defesas\\_Site/PESSANHA.pdf](http://www.ufsj.edu.br/portal2-repositorio/File/mestradosicologia/2016/Defesas_Site/PESSANHA.pdf)>. Acesso em: 15 de junho de 2022.