

DOI: 10.20911/21799024v16n1p74/2025

Os limites da Inteligência artificial à luz da Fenomenologia contemporânea: Inteligência Situada e Mundo da Vida

The limits of Artificial Intelligence in light of contemporary Phenomenology: Situated Intelligence and the Lifeworld

Alejandro José Castejón Gonzalez¹

Resumo: O presente artigo tem como objetivo central afirmar uma distinção essencial entre a inteligência humana e a inteligência artificial, por meio da argumentação de Hubert Dreyfus desenvolvida na sua obra *o que os computadores não podem fazer* e relacionando tal argumentação com o conceito husserliano de *lebenswelt* (*mundo da vida*). Como ponto de partida, será feito um esclarecimento do que se entende por inteligência artificial e as distinções entre inteligência artificial forte e fraca. Em continuidade, será apresentada a postura e a concepção predominante entre os cientistas da IA, que relacionam o funcionamento da inteligência artificial com a mente humana, será contextualizada o surgimento desta postura junto com os principais pressupostos. Finalmente, será exposta a crítica dreyfusiana e a influência desta crítica pelo filósofo Edmund Husserl, expondo principalmente o conceito de inteligência situada e mundo da vida.

Palavras-chave: Inteligência artificial, situação, relevância, contexto, mundo da vida.

1 Bacharel em filosofia pela Faculdade Jesuíta de Filosofia e Teologia.

Abstract: The main objective of this article is to assert a fundamental distinction between human intelligence and artificial intelligence, based on Hubert Dreyfus's argument developed in his work *What Computers Can't Do*, and relating this argument to the Husserlian concept of *Lebenswelt* (lifeworld). As a starting point, the article will clarify what is meant by artificial intelligence and the distinctions between strong and weak AI. Following this, it will present the prevailing stance and conception among AI scientists, who relate the functioning of artificial intelligence to the human mind, and it will contextualize the emergence of this stance along with its main assumptions. Finally, Dreyfus's critique will be presented, along with the influence of the philosopher Edmund Husserl on this critique, especially highlighting the concepts of situated intelligence and the lifeworld.

Keywords: Artificial intelligence, situation, relevance, context, lifeworld.

Introdução

A inteligência artificial se apresenta hoje como o grande fenômeno da contemporaneidade. Discussões acontecem na vida cotidiana, nas faculdades, nos ambientes científicos. As posturas nestas discussões são as mais variadas, sendo algumas otimistas, outras pessimistas e outras até apocalípticas. Mas um fato é indiscutível: a IA é um fenômeno que interessa a todos, afeta e afetará a todos. Um aspecto é igualmente muito discutido, no entanto, muitas vezes, com pouco rigor e precisão: a relação entre a mente humana e a chamada inteligência artificial.

A Inteligência artificial e a ciência cognitiva trabalharam e desenvolveram suas pesquisas numa íntima relação durante várias décadas. Tal como Daniel Andler afirmou estavam "unidos como gêmeos siameses" (2015, p.343). Estas duas ciências se apoiavam em pressupostos sobre a mente humana e em princípios metodológicos fundamentais para trabalhar. A fenomenologia, por outro lado, durante muito tempo tomou distância desta junção, pois considerava alguns pressupostos e métodos pouco viáveis e acertados. Desta forma, as intervenções da fenomenologia foram de "críticas frontais ao projeto da IA/Ciecog" (ANDLER, 2015, p.343). Na atualidade, muitas destas críticas feitas pela abordagem fenomenológica foram assumindo um espaço importante e interessante no ambiente científico, o impacto destas críticas provocou um distanciamento significativo entre a IA e a ciência cognitiva (Ciecog) e a IA acabou se rendendo diante de algumas críticas e considerações da fenomenologia (ANDLER, 2015, p.343). As últimas décadas foram testemunhas de algumas transformações metodológicas: alguns cientistas cognitivos assumiram como objetivo a reconciliação com a Ciecog considerando ainda algumas ideais defendidas pela fenomenologia. Por outro lado, alguns filósofos da mente e cientistas cognitivos numa linha mais analítica, pouco familiarizados e também pouco simpatizantes com a fenomenologia, têm buscando outras soluções à algumas críticas feitas

pela abordagem fenomenológica que se apresentam como incontornáveis. Neste sentido, o atual campo de discussão apresenta uma inadequação da antiga distinção entre filosofia analítica e continental. Hubert Dreyfus é um exemplo desta inadequação, sendo ele um filósofo estudioso da IA, fenomenólogo e ao mesmo tempo crítico e contrário a conceitos aprioristas e dualista. Dreyfus desenvolveu suas críticas à ciência da IA “no próprio campo dela, como um filósofo da ciência naturalista que entende a ciência e a filosofia como continuas” (ANDLER, 2015, p.344).

Este artigo assume como principal objetivo afirmar uma distinção essencial entre a inteligência humana e a inteligência artificial, por meio da argumentação de Hubert Dreyfus desenvolvida na sua obra *o que os computadores não podem fazer* e, ao mesmo tempo, relacionando tal argumentação com o conceito husserliano de *lebenswelt* (*mundo da vida*). Inicialmente, será apresentado o que se entende por inteligência artificial e as distinções entre inteligência artificial forte e fraca. No segundo momento, será apresentada a postura e a concepção predominante entre os cientistas da IA, que relacionam o funcionamento da inteligência artificial com a mente humana, será contextualizado o surgimento desta postura junto com os principais pressupostos. Finalmente, será exposta a crítica de Hubert Dreyfus e a influência desta crítica pelo filósofo Edmund Husserl, expondo o conceito de inteligência situada, e mundo da vida.

1. Esclarecimentos conceituais sobre Inteligência Artificial

Antes de aprofundar nas discussões principais deste artigo, é preciso uma contextualização do problema e um esclarecimento de conceitos essenciais para assim adentrar adequadamente, e na medida do possível, nas discussões referentes a este campo. Esta precaução é necessária para a discussão, pois é possível perceber claramente diversos impedimentos e imprudências nestas discussões. Como ponto de partida, é evidente que as discussões sobre inteligência artificial têm ocupado um grande espaço nas últimas décadas nas diversas disciplinas, ciências e na vida cotidiana. Com efeito, tal fenômeno chegou para afetar e modificar todos os campos da vida humana. Em segundo lugar, o fato da inteligência artificial ser um fenômeno que produz impacto em diversos aspectos da vida humana, traz como consequência a intervenção de diversas ciências e disciplinas no campo da discussão. O estudo da inteligência artificial enquanto tal assume um caráter interdisciplinar, estando ligada a ciências cognitivas, psicologia, ciência de dados, neurociência, computação, entre outras (COECKE-LBERGH, 2023, p.45) e pelo seu impacto, outras disciplinas se introduzem nas discussões, entre elas, o direito, a filosofia, a sociologia, a antropologia, as ciências políticas, a educação, entre outras. A principal problemática são as discordâncias conceituais e epistemológicas que variam em cada uma dessas ciências e disciplinas. De um lado, temos o otimismo das ciências da computação e das ciências cognitivas em relação ao futuro da IA, ao mesmo tempo que se revela, em alguns casos, o desconhecimento ou até o desinteresse das implicações éticas, filosóficas, jurídicas, antropológicas e sociais que tal otimismo implica. De

outro lado, temos as críticas ferozes dos filósofos, sociólogos, entre outros e, ao mesmo tempo, grande desconhecimento do funcionamento essencial da IA.

Hubert Dreyfus se apresenta como uma saída adequada para este impasse. Dreyfus, ademais de filósofo, dedicou-se durante muito tempo ao estudo da ciência e da tecnologia, trabalho pelo qual recebe um importante reconhecimento no espaço acadêmico. De fato, é conhecido pelo trabalho de interconexão da interpretação fenomenológica, principalmente de Husserl, Heidegger e Merleau-Ponty, com questões científicas e tecnológicas, especificamente na inteligência artificial. Desta forma, Dreyfus foi um filósofo que se inseriu na discussão, tendo amplos conhecimentos nos vários campos disciplinares essenciais para dar aportes e críticas. Antes de aprofundar nas argumentações dreyfusianas, é preciso esclarecer o que se entende hoje por inteligência artificial e as distinções entre inteligência artificial forte e fraca e inteligência artificial enquanto ciência e enquanto tecnologia.

Inteligência artificial é definida como “inteligência exibida ou simulada por código (algoritmos) ou máquinas” (COECKELBERGH, 2023, p.43). Sendo ainda fraca esta definição, principalmente em termos filosóficos, Mark Coeckelbergh apresenta outra definição desenvolvida por Philip Jansen: inteligência artificial é “a ciência e a engenharia de máquinas com capacidades que são consideradas inteligentes pelo padrão de inteligência humana” (2023, p.43). Com efeito, outras definições foram desenvolvidas e podem ser até mais fortes e completas. O mesmo Coeckelbergh, no seu livro *Ética na inteligência artificial*, apresenta outras posturas e as problematiza, no entanto, em todas parece aparecer uma constante conceitual: IA como *simulação da inteligência*. Nesta constante conceitual aparece, de início, várias problemáticas: o que é inteligência? Inteligência é um aspecto exclusivamente humano? Se este conceito não é exclusivo do ser humano, qual inteligência pretende-se simular? É possível imitar a inteligência humana? Além disso, para imitar é preciso uma concepção de inteligência humana, qual conceito de inteligência humana é utilizado pelos cientistas da IA? A concepção de inteligência desenvolvida por estes cientistas se apresenta como suficientemente abrangente? Se é possível imitar a inteligência humana, será possível pensar que a IA superará a inteligência humana? (*superinteligência e transhumanismo*) (COECKELBERGH, 2023, p.13). Todas estas questões e ainda outras não mencionadas aparecem constantemente nos debates e são respondidas das mais diversas formas. Hunter e Shane, por exemplo, catalogam, num pequeno ensaio, setenta definições diferentes sobre inteligência artificial presentes em diferentes autores (2007, p.1)

Um outro aspecto importante de ser ressaltado são algumas distinções essenciais na compreensão da IA. A primeira distinção é inteligência artificial fraca e inteligência artificial forte. Antes, é preciso esclarecer que tais distinções estão na atualidade no campo de debate e discussão. Coeckelbergh, por exemplo, relaciona o conceito de IA fraca com o conceito de IA restrita e o conceito de IA forte com o conceito de IA geral. Para este filósofo a inteligência artificial fraca ou restrita, possui um campo de domínio e estudo mais restritivo (2023, p.45). Ela possui finalidades mais específicas, exemplos como máquinas que jogam

xadrez, que escolhem imagens, o reconhecimento de voz, chatbot, as *Alexas*, sistemas de recomendação, etc. Em outras palavras a IA fraca é projetada para fazer tarefas limitadas a partir de um conjunto de regras predefinidas (MARTINS, 2023, p.38). Este tipo de inteligência artificial se encontra muito presente na atualidade e a projeção de seu uso para o futuro gera grandes otimismo. A inteligência artificial forte ou geral pretende levar a termo qualquer tarefa cognitiva humana possível (COECKELBERGH, 2023, p.45). Esta última se apresenta para muitos como ambiciosa. Este tipo de inteligência apresenta a proposta de fazer que máquinas pensem, resolvam problemas, tal como um ser humano. Este tipo de inteligência ainda se apresenta com um objetivo a ser alcançado (MARTINS, 2023, p.38). Esta distinção apresenta na contemporaneidade discussões e é alvo de muitas críticas. Por uma parte alguns, assumem estas distinções num sentido mais prático, considerando a IA fraca como aquela que representa uma ferramenta para o estudo das faculdades cognitivas e a IA forte como aquela que se engaja na missão de criar uma mente. No entanto, esta distinção não deveria ser relacionada com os conceitos de IA geral e restrita. Para muitos, tal relação apresenta problemáticas filosóficas nas quais se evidenciam ambiguidades conceituais, pois uma IA forte no sentido prático de IA geral seria fraca pois seria apenas um modelo da mente e não uma mente e viceversa.

A seguinte distinção importante é da inteligência artificial enquanto tecnologia e enquanto ciência. IA enquanto tecnologia, tal como a própria expressão apresenta, pretende ser uma ferramenta para as diversas atividades humanas. Esta se apresenta como uma tecnologia para propósitos úteis e práticos. Por outro lado, a IA como ciência estuda o fenômeno da inteligência, da mente e do cérebro (COECKELBERGH, 2023, p.45). Esta última trabalha a partir de interligações fortes entre a ciência da computação e as ciências cognitivas (psicologia, neurociência, filosofia da mente). Entre suas principais finalidades se encontra a imitação da inteligência humana em entes artificiais e a compreensão da mente humana por meio de modelos computacionais.

2. Inteligência Artificial e a mente humana

Vários filósofos desenvolveram críticas às concepções de inteligência artificial forte e de inteligência artificial como ciência. A maioria destas críticas não apontam para a possibilidade ou não do alcance projetado das capacidades destes entes artificiais, mas a um pressuposto muito fraco que se encontra no fundamento desta ciência: a compreensão de inteligência humana a partir de modelos computacionais. Este artigo pretende somar-se a estas críticas por meio da argumentação do filósofo Hubert Dreyfus e finalmente apontar para um elemento dos últimos escritos do pai da fenomenologia Edmund Husserl: *o mundo da vida*.

Dreyfus desenvolve esta discussão e suas críticas numa obra intitulada *What computers Can't do*. Nesta obra, o filósofo aponta que a IA, na realidade, não estaria dando passos significativos na criação de artefatos inteligentes e este fato se devia a pressuposições equivocadas sobre a mente, que os especia-

listas da IA ignoravam (ANDLER, 2012, p.343-344). A argumentação dreyfusiana é desenvolvida por Daniel Andler da seguinte forma, lemos:

A IA pretende explicar a inteligência humana construindo máquinas inteligentes e, para esse fim, faz uso de um conjunto de hipóteses (H) referentes à natureza essencial da mente. Mas os pressupostos em H estão equivocados e, além disso, a mente exibe um conjunto de propriedades (P) que parece explicar a inteligência ou ser constitutivo dela que a IA ignora. A falsidade de (H) e a realidade e a importância de (P) resultam em parte de considerações independentes, em parte de uma inferência à melhor explicação (2012, p.344).

Este capítulo seguirá o fio condutor desta síntese do argumento de Dreyfus, desenvolvida por Daniel Andler. Como ponto de partida, será apresentada (H), isto é, as Hipóteses dos cientistas da IA, a contextualização desta postura e as suas concepções e pressupostos. Em continuidade será analisada (P), isto é, as propriedades ignoradas que a IA ignora, a postura de Hubert Dreyfus. Dito de outra forma, (H) como a postura da maioria dos cientistas da IA e (P) como a crítica e a posição de Dreyfus.

Como já foi especificado, a IA, ao mesmo tempo que pretendia criar artefatos inteligentes, isto é, máquinas que pensam como qualquer humano pensaria, pretendia explicar a inteligência humana, construindo estas máquinas. O percorrido desta concepção é contextualizado pelo próprio Dreyfus da seguinte forma:

Desde que os gregos inventaram a lógica e a geometria, a ideia de que todo raciocínio poderia ser reduzido a algum tipo de cálculo — de modo que todos os argumentos pudessem ser expostos de uma vez por todas — tem fascinado a maioria dos pensadores radicais da tradição ocidental. (DREYFUS, 1975, p.17).

O filósofo aponta para as raízes dos pressupostos presentes na compreensão da IA já na filosofia antiga. Nesta aparece uma certa continuidade e presença constante destes pressupostos, que influenciaram o pensamento ocidental. Este conceito pode ser definido como a formalização do conhecimento aspecto predominante no espaço filosófico. Esta ideia ainda germinante dos modelos computacionais e da futura IA, amadureceria durante a história com alguns aportes filosóficos e fatos histórico-científicos que consolidariam esta forma de pensar a mente humana. Hobbes, por exemplo, afirmaria que o pensamento e o raciocínio não é nada além do que um ato de calcular; Leibniz pensou ter encontrado uma álgebra ou uma linguagem simbólica universal na qual poderia atribuir, enumerar e caracterizar todos os objetos. George Boole, com sua álgebra de tipo binária, conceberia o pensamento como simplesmente manipulação de símbolos (DREYFUS, 1975, p. 19. 21). Estes e outros pensadores contribuíram na futura concepção de inteligência artificial.

Ao mesmo tempo que se desenvolviam estas teorias, o campo prático progredia influenciando-se destes teóricos. Charles Babbage, por exemplo, projetaria um *motor analítico* que deveria funcionar como um computador digital moderno, operando com símbolos abstratos e produzindo operações e decisões

lógicas. Um passo importante neste percorrido para o surgimento da IA foi o trabalho de Alan Turing quem, ao mesmo tempo que tentava aprender as capacidades destas máquinas, aprofundava na pesquisa e no estudo da natureza da própria razão. Com efeito, foi ele quem suscitou a pergunta: podem as máquinas pensar? E pretendeu respondê-la por meio de um teste que ele próprio chamou de *Jogo da imitação*. (DREYFUS, 1975, p.23). Turing extrairia conclusões cheias de otimismo que influenciariam os trabalhos subsequentes:

Podemos esperar que as máquinas acabem competindo com os homens em todos os campos puramente intelectuais. Mas quais são os melhores para começar? Até mesmo essa é uma decisão difícil. Muitas pessoas acham que uma atividade muito abstrata, como jogar xadrez, seria a melhor. Também é possível afirmar que é melhor fornecer à máquina os melhores órgãos sensoriais que o dinheiro pode comprar e depois ensiná-la a entender e falar inglês. Esse processo poderia seguir o ensino normal de uma criança. As coisas seriam apontadas e nomeadas, etc. Novamente, não sei qual é a resposta certa, mas acho que ambas as abordagens devem ser tentadas. (TURING, 1950, p.460).

Esta conclusão de Alan Turing influenciou e impulsionou as pesquisas práticas que deram os principais avanços no surgimento do que hoje se entende por IA. Claude E Shannon e Allen Newel se focariam em máquinas que jogassem xadrez com jogadas fortes e na velocidade de um ser humano. As afirmações destes cientistas aparecem como essenciais no caminho da IA: “Com um mínimo conseguido no caminho da complexidade, entramos, por fim, na arena das jogadas humanas” (DREYFUS, 1975, p.26). Esta afirmação levaria a Allen Newel e Hebert Simon, por meio da análise dos procedimentos que os estudantes executavam para resolver problemas lógicos, a decidirem simular a inteligência prática. Em outra linha, mas, igualmente influenciada por Turing, Anthony Oettinger trabalhava na busca de máquinas que traduzissem diversas línguas (DREYFUS, 1975, p.26).

Já em 1950 aparece a proposta clara da “ciência produzir máquinas que fazem coisas que exigiriam inteligência, se feita pelos homens” (TADDEY; DA COSTA; FURTADO, 2029, p.98). A proposta segundo Churchland é criar dispositivos puramente físicos que possuam inteligência genuína. Haugeland afirma uma psicologia como engenharia reversa. Ele afirmaria que, por meio da percepção da capacidade dos computadores realizar tarefas inteligentes, seria necessário afirmá-las como inteligentes, para assim, finalmente, concluir que a mente humana funcionaria como um computador (TADDEY; DA COSTA; FURTADO, 2029, p.98). Gardner explicita e radicaliza esta concepção afirmando que a inteligência artificial se insere no hexágono das ciências cognitivas que buscam explicar a mente e a inteligência humana: filosofia, psicologia, neurociência linguística e antropologia (TADDEY; DA COSTA; FURTADO, 2029, p.98).

Este percorrido no qual a inteligência artificial surgiria, apresentaria conclusões bem claras da concepção de inteligência humana e a possibilidade da simulação e imitação desta. Hubert Dreyfus aponta para cinco teses presentes

no desenvolvimento da IA que se apresentam como pressupostos que geram todo um otimismo no desenvolvimento desta ciência:

Presumir a tese biológica de que em algum nível de operação [...] o cérebro processe informações em operações distintas por intermédio de um agente biológico que liga ou desliga os interruptores. [...] A tese psicológica de que a mente pode ser interpretada como um mecanismo que opera a base de bits de informação segundo regras preestabelecidas [...]. Um pressuposto epistemológico de que todo conhecimento pode ser formalizado, isto é, que tudo o que pode ser compreendido pode ser expresso em termos de relações lógicas [...]. Finalmente, [...] que toda informação pertinente sobre o mundo, tudo aquilo essencial à produção do comportamento inteligente deve em princípio ser analisável como um conjunto de elementos determinados de situação livre. Este é o pressuposto ontológico que o que existe é um conjunto de fatos cada um logicamente independente de todos os outros (DREYFUS, 1975, p. 120-121).

Essas teses se sintetizam na ideia de que a mente humana é concebida, nas várias etapas, como semelhante a um computador genérico e não especializado; isto é, a mente operaria com um tipo de informação que poderia ser expressar por meio de símbolos estruturados segundo o formalismo lógico. Além disso, este formalismo lógico implica a compreensão do mundo como conjunto de fatos independentes e que são conectados por relações lógicas. As ideias anteriormente apresentadas englobam as hipóteses (H) presentes e fundantes da concepção da IA enquanto ciência e na compreensão da natureza essencial da mente humana. Estas hipóteses são criticadas por Dreyfus e afirmadas por ele como equivocadas. As cinco teses anteriormente apresentadas são, nas palavras do filósofo, indevidamente otimistas (1975, p.119)

3. Situação, Contexto, Relevância e Mundo da Vida

No último tópico deste artigo se pretende desenvolver as propriedades (P) que explicam a inteligência humana e sua constituição, mas que são ignoradas pela IA. Dreyfus introduz no debate sobre inteligência artificial, dimensões ignoradas para a compreensão da inteligência humana. Estas dimensões foram desenvolvidas pelo filósofo pela grande influência que recebeu da fenomenologia contemporânea, principalmente de filósofos como Edmund Husserl, Martin Heidegger e Merleau-Ponty. Estas temáticas permanecem fortemente no debate ainda na atualidade e consistem principalmente na ideia de que a inteligência e a mente humana atuam essencialmente por meio de três aspectos: corporeidade, engajamento e contexto (ANDLER, 2012, p.344). Estes três aspectos se apresentaram como problemáticas no avanço do projeto da IA enquanto ciência. Na *corporeidade* e no *engajamento* as principais posturas se dividem e as partes possuem argumentos muito fortes e sólidos. No *contexto* ainda tem se tornado problemática a refutação por parte dos cientistas otimista da IA e é justamente este aspecto que será explorado nos seguintes parágrafos.

O ser humano está situado num contexto cheio de modificações e escolhas indefinidamente multiplicáveis e é justamente nesta realidade que a mente está inserida e responde continuamente a tais estímulos. A formalização da inteligência tem como desafio se confrontar com o fato da mente humana se adequar constante e indefinidamente em contextos ou situações mundanas, nas quais, de forma quase espontânea, esta mente faz um processo de escolha do que é relevante ou não para continuar seus processos. No entanto, tal desafio não se apresenta como fácil para os cientistas da IA, pois a compreensão da formalização da inteligência compreende que “o pensamento resulta de nada além de uma realização de cálculos a partir de representações simbólicas construídas de pedaços elementares de informação independentes do contexto” (ANDLER, 2012, p.345). Dreyfus inicia sua argumentação em relação à inteligência situada da seguinte forma:

Ao discutirmos a solução de problemas e a tradução da linguagem nos deparamos com a ameaça de uma regressão de regras para a determinação da relevância e do significado. Igualmente, ao iniciarmos o processo de aprendizagem, algo deve ser conhecido antes que quaisquer regras sejam ensinadas ou aplicadas. Em ambos os casos, descobrimos que, se não existem fatos com significado imutável, então somente com o auxílio do contexto poderemos impedir essa regressão. Devemos agora nos voltar diretamente para uma descrição da situação ou contexto, a fim de podermos explicar, da melhor maneira possível, o modo singular pelo qual os seres humanos estão “inseridos no mundo” e a função especial que esse mundo desempenha, tornando possível um comportamento ordenado, mas, ao mesmo tempo, não normativo (1975, p. 234).

Nesta citação, o filósofo aponta para algo que já anteriormente é conhecido pelos seres humanos antes de qualquer regra ensinada e aplicada no momento da aprendizagem. Esse algo não é outra coisa que o contexto e a situação na qual os seres humanos pensam e sem a qual é impossível levar adiante o ato de pensar, pelo menos na forma em que os humanos fazem isso. De fato, estamos inseridos no mundo, nossa consciência tende para ele de forma intencional e é por esta inserção e intencionalidade que o mundo não aparece ao homem como um objeto estranho. Contrariamente, o ser humano inteligente se entende, age, pensa, sente e movimenta sua inteligência no mundo. Para Dreyfus o fato de afirmar que o comportamento humano pode ser regular não implica a afirmação de ele ser normativo e governado por regras formalizáveis (DREYFUS, 1975, p.235).

O fato do ser humano estar inserido no mundo implica estar inserido numa situação que pode ser determinada como pragmática-holística e interessada (TADDEY; DA COSTA; FURTADO, 2029, p.109-110). Esta natureza da inteligência de situar-se de forma pragmática e interessada caminha na contramão da proposta predominante da IA, que entende a inteligência a partir de uma natureza atômica, formalista e desinteressada. Esta última forma de compreender a inteligência conduz ao problema da organização incontável de fatos, que também é chamado do problema da relevância.

Hubert Dreyfus apresenta a problemática de como se determinam quais fatos são relevantes ou irrelevantes, essenciais ou não essenciais num processo inteligente. Com efeito, esta problemática tem sido uma dificuldade nos procedimentos da inteligência artificial. No contexto humano, por exemplo, a relevância nunca aparece de modo fixo, mas sempre em função de objetivos humanos; isto leva a conclusão que todos os fatos podem ser relevantes em determinadas situações. Em outras palavras, a relevância é situacional (DREYFUS, 1975, p.236). Afirmar a relevância de forma situacional e não fixo e afirmar que todos os fatos são relevantes conduz a uma série de questionamentos e problemáticas na tentativa de simular a inteligência humana.

O caráter situacional da relevância opera em dois sentidos: em uma situação particular qualquer, um número indefinido de fatos pode ser relevante, e um número indefinidamente grande ser irrelevante. Entretanto, como um computador não está colocado no contexto de uma situação, ele precisa tratar todos os fatos como sendo possivelmente relevantes em todas as situações, o que deixa os pesquisadores da IA num dilema: ou armazenar e posicionar uma infinidade de fatos, ou excluir alguns possivelmente relevantes do campo de cálculos do computador (DREYFUS, 1975, p.236).

Diante dos fatos indefinidamente multiplicáveis e diante da impossibilidade dos computadores estarem num contexto ou situação, os cientistas da IA se encontram diante de dois caminhos pouco factíveis: ou armazenar neles uma infinidade de fatos ou excluir fatos possivelmente relevantes. Ambos caminhos apresentam um dilema. Por um lado, assumir como tarefa introduzir inúmeros fatos nestes computadores se apresenta como um caminho infinito e sempre insuficiente, além disso, estas máquinas correriam o risco de uma *paralisia* diante da análise infinita de fatos. Por outro lado, a exclusão de fatos possivelmente relevantes, determinaria estes computadores a atividades limitadas e os distanciaria completa e definitivamente de uma inteligência situada, desta forma, a IA estaria condenada a ser exclusivamente IA restrita. Além disso, a extração do contexto conduziria ao fato das informações perderem o seu significado. Desta forma, estas máquinas não poderiam efetivamente fazer o que nós fazemos quando pensamos, pois, neste contexto, pensar se apresenta como saber o que deixar de fora diante de infinitas possibilidades de fatos relevantes. Sendo assim, as máquinas não poderiam pensar. A inteligência artificial precisaria “ter acesso a todos os fatos banais que um ser humano conhece sobre o modo como o mundo habitual funciona em circunstâncias habituais” (ANDLER, 2012, p.346) e além disso, “a qualquer momento, qualquer coisa pode vir a se tornar relevante em qualquer situação” (ANDLER, 2012, p.347).

Desta problemática apresentada por Hubert Dreyfus é possível extrair uma conclusão com uma postura bem definida. Se afirmamos a inteligência como corporificada, engajada e situada, se estes elementos são essenciais para definir o que é pensar para um humano, se as máquinas chamadas de inteligentes não podem desenvolver uma atividade inteligente situada (como foi argumentado), logo podemos responder à pergunta de Huber Dreyfus: *O que os computadores não podem fazer?* Seguindo o percurso desta argumentação podemos afirmar

que elas podem fazer muitas coisas, excluindo pelo menos uma: elas não podem pensar, pelo menos como os seres humanos pensam. Não se pretende negar o fato de que estas máquinas possam simular algumas operações que o ser humano faz utilizando suas faculdades inteligentes, mas o ato de pensar no rigor do termo exposto não pode ser reduzido a estas operações. Igualmente não se pretende dar predições do futuro deste campo pois ele para todos é incerto e obscuro.

É preciso esclarecer que, embora a postura de Dreyfus apresente argumentações fortes e críticas ao projeto de IA enquanto ciência, há também outras posturas contrárias que entram no debate para problematizar e tentar fugir das dificuldades propostas pelos diversos críticos da IA. A argumentação de Dreyfus sobre a situação, o contexto e a relevância aparece, por exemplo, como uma resposta à postura de Daniel Dennet, quem afirma que é possível, a partir de uma descrição detalhada de um dado bruto, conseguir chegar a conceitos que os seres humanos revestem nas coisas. Dennet num pequeno ensaio chamado *Cognitive Wheels: The Frame Problem of AI* apresenta outra postura em relação ao problema da relevância. Para Dennet, muitas das críticas feitas por Dreyfus sobre as limitações da IA se encontram sistematicamente próximas do que ele chama de problema do enquadramento (DENNET, 1984, p.2). O problema do enquadramento faz referência a como fazer representações úteis de todas as informações que ser inteligente conquistou arduamente pela experiência (DENNET, 1984, p.6). A inteligência deve ser capaz de realizar estas representações num mundo em movimento, considerando de acordo com as situações o que é banal e o que não. Para este filósofo, um agente artificial com um conjunto de quadros e roteiros bem abastecidos, apropriadamente ligados entre eles e em contato com os estímulos do mundo por meio dos mecanismos perceptuais seria capaz de enfrentar o mundo a partir de um sistema elaborado de hábitos de atenção e tendência, para assim chegar a diversas conclusões em diversas circunstâncias (DENNET, 1984, p. 8).

Esta postura mais otimista, como foi explicitado, caminha na contramão da argumentação de Dreyfus. A relevância, a situação e o contexto são coisas inerentes ao ser humano e Dreyfus afirma que esta função não pode ser igualada por um computador, pois este trabalha unicamente com objetos já determinados (1975, p.241). A compreensão dreyfusiana sobre o ser humano como situado no mundo é sem dúvida nenhuma influência da fenomenologia contemporânea. Dreyfus recebeu grande influência de Heidegger, Merleau-Ponty e principalmente do pai da fenomenologia Edmund Husserl. Hubert quem influenciou numerosos cognitivistas e críticos da IA, trouxe a fenomenologia husserliana para este campo, apresentando Husserl como um protocognitivista (ESCUADERO, 2012, p.297).

A inteligência como situada e contextualizada pode ser explorada a partir do conceito husserliano de Mundo da vida (*lebenswelt*). O mundo da vida faz referência ao horizonte pré-científico da vida humana. É o espaço da *doxa*, do cotidiano e relativo aos fins e propósitos humanos (FERRER, 2012, p.15). Este *mundo*, além de ser inerente e essencial ao ser humano, se apresenta como o

subsolo de toda e qualquer ciência, pois é nela que todas se iniciam. Em contrapartida, Husserl apresenta uma problemática que se torna fundamental:

Devemos considerar agora como da maior importância uma substituição, que está em efetivação já em Galileu, do único mundo efetivo, o que é efetivamente devido à medida da percepção, do único mundo alguma vez experienciado e experienciável - o nosso mundo da vida cotidiano - pelo mundo matematicamente subtraído das idealidades. Esta substituição foi rapidamente transmitida aos seus continuadores, aos físicos de todos os séculos subsequentes (HUSSERL, 2012, p.38).

Esta citação da obra husserliana *A crise das ciências europeias* apresenta a problemática da colonização do mundo científico e matemático sobre o mundo da vida e o cotidiano experiencial. Colonização que transposta para atualidade e para a problemática da inteligência artificial pode ser entendida como o domínio do cálculo sobre o juízo, do algoritmo sobre a crítica e o discernimento. Husserl, além de argumentar esta substituição aponta para o déficit da desconsideração do mundo da vida pelo ambiente científico e, além disso, aponta para as problemáticas da pretensão da descontinuidade e até da ruptura do mundo da vida com o mundo científico, lemos:

Nunca se questiona cientificamente a maneira como o mundo da vida funciona em permanência como plano de fundo, como as suas múltiplas validades pré-lógicas são fundamentadoras para as verdades lógicas, as verdades teóricas. E talvez a cientificidade, que este mundo da vida, como tal e na sua universalidade, exige, seja uma cientificidade específica, justamente não lógico objetiva, e que, como a cientificidade fundamentadora última, o seu valor não seja o de uma cientificidade menor, mas superior (HUSSERL, 2012, p. 101).

Nesta citação Husserl questiona a irrelevância com a qual o mundo científico considera o mundo da vida. Este questionamento apontado por Husserl ressoa hoje nas argumentações de Dreyfus ao perceber como a IA desconsiderou elementos tão fundamentais da inteligência humana que hoje resultam problemáticos. O conceito de mundo da vida é entendido aqui como a explicação da natureza situada da inteligência exclusiva do ser humano, na qual por meio de contextos e situações é possível, de forma simples e sem problemas, se aderir ou renunciar a fatos relevantes. Esta capacidade natural humana é possível por meio do fato essencial de também os seres humanos estarem inseridos de forma especial no mundo). Tal como afirma Ricardo Mejia Fernandez, o mundo da vida recebe seu significado no nosso centro de experiência, o mundo da vida é o mundo de intenções e necessidades individuais, coletivas e sociais (2014, p.15-16).

Considerações finais

Neste artigo foi desenvolvida uma argumentação na qual se apontou para uma distinção essencial e fundamental entre inteligência humana e inteligência

artificial. A crítica de Hubert Dreyfus aos cientistas otimistas da IA aponta justamente para uma compreensão errada da mente humana, especificamente, a compreensão de uma mente que relaciona fatos independentes do mundo por meio de operações lógicas, formalísticas e simbólicas. Ao mesmo tempo, esta crítica aponta para a desconsideração de elementos importantes e fenomenologicamente evidentes na natureza da mente humana; especificamente o *contexto*, a *situação* e a *relevância* com os quais a mente humana pensa. Estes elementos foram relacionados com o conceito de *mundo da vida* (*lebenswelt*) presente nos últimos escritos do filósofo Edmund Husserl, que justamente apontam para o subsolo fundamental no qual o ser humano pensa, age e vive: o mundo do cotidiano, da experiência imediata. A mente humana pensa a partir de um engajamento natural e espontâneo no mundo, este engajamento aponta para uma forma pré-científica e pré-lógica de pensar própria do ser humano. Husserl, ao mesmo tempo que apresenta este conceito, aponta para o problema da sociedade contemporânea ocidental da colonização do mundo científico, técnico e matemático sobre este mundo da vida. Esta colonização implica a desconsideração e desvalorização do mundo da vida na atualidade. Este objetivo foi desenvolvido a partir de um esclarecimento prévio de conceitos fundamentais no campo da inteligência artificial e a partir de uma contextualização de alguns pressupostos presentes nesta ciência.

A problemática e os impasses que foram apresentados no corpo do texto elucidam o dilema no qual a IA se encontra se assume com rigor determinados conceitos. Por uma parte ela pode escolher explicar a inteligência artificial desconsiderando elementos fundamentais da inteligência humana como engajamento, corporeidade e contexto e assim falhando no seu projeto de imitar e criar uma inteligência segundo o paradigma humano; por outro lado, pode assumir as propriedades essenciais da inteligência humana anteriormente mencionadas como fundamento conceitual e renunciar à concepção formalista, lógica e dualista na qual se pretendia levar adiante o projeto cientista da IA geral, mas ao mesmo tempo assumindo as problemáticas e dificuldades que tal projeto apresenta. Pode-se concluir este artigo afirmando que qualquer discurso ou pesquisa sobre a inteligência artificial devem sempre ser apresentados em aberto, pois mesmo as melhores argumentações e os melhores dados, esta ciência aparece como projetiva. Tal fato não proíbe induzir consequências éticas, políticas, epistemológicas e até científicas que os dados e as pesquisas atuais apontam. As posturas em relação ao tema da IA se dividem exponencialmente e das diversas formas. O principal dado que estes fatos apontam é a importância e a relevância que o fenômeno IA assume na contemporaneidade; realmente este fenômeno interessa e deve interessar a todos.

Referências

ANDLER, Daniel. Fenomenologia na inteligência artificial e na ciência cognitiva. In: DREYFUS, Hubert; WRATHALL, Mark (orgs). *Fenomenologia e existencialis-*

mo. Tradução: Cecília Camargo Bartalotti e Luciana Pudenzi, São Paulo: Edições Loyola, 2012. p.343-357.

COECKELBERGH. Ética na inteligência artificial. São Paulo / Rio de Janeiro: Ubu Editora / Editora PUC-Rio, 2023.

DENNET, Daniel. Cognitive Wheels: The Frame Problem of AI. In: *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. 1984. Disponível em: https://www.researchgate.net/publication/225070451_Cognitive_Wheels_The_Frame_Problem_of_AI.

DREYFUS, HUBERT. *O que os computadores não podem fazer: crítica da razão artificial*. Rio de Janeiro: 1975. A Casa do livro el dorado.

ESCUADERO, Jesús Adrián. Mente y cuerpo. Las aportaciones de la fenomenología genética de Husserl a las nuevas ciencias cognitivas: el caso de la neurofenomenología. *Thémata*. Barcelona, n. 46, 2012, p. 293-298.

FERRER, Diogo Falcão. Apresentação de Husserl. A crise das ciências Europeias e a fenomenologia transcendental: uma introdução à filosofia fenomenológica. 1 Ed. Rio de Janeiro: Forense Universitária, 2012.

HUNTTER, Marcus; LEG, Shane. A collection of definitions of intelligence. *Technical Report*. IDSIA -7-7. 2007, p.1-12. Disponível em: <https://arxiv.org/abs/0706.3639>.

HUSSERL, Edmund. *A crise das ciências europeias e a fenomenologia transcendental: uma introdução à filosofia fenomenológica*. Trad: Diogo Falcão Ferrer. 1 Ed. Rio de Janeiro: Forense Universitária, 2012.

MEJIA FERNANDEZ, Ricardo. Fenomenologia Husserliana e inteligência artificial en Hubert L, Dreyfus. Universidade de Salamanca (Trabajo fin de máster) 2014.

MOHANTY, J.N. Intencionalidade. In: DREYFUS, Hubert; WRATHALL, Mark (orgs). *Fenomenologia e existencialismo*. Tradução: Cecília Camargo Bartalotti e Luciana Pudenzi, São Paulo: Edições Loyola, 2012. p.75-81.

TADDEY, Paulo; DA COSTA, Arthur; FURTADO, Robson. A fenomenologia de Heidegger na crítica de Dreyfus à IA simbólica. *Revista natureza humana*, São Paulo, v. 22, n, 1, 2020, p.95-122.