

Desafios para o Design na era da Inteligência Artificial: reflexões e proposições para favorecer a ética no projeto

Carolina Vaitiekunas Pizarro^(*)

Resumen: Os desenvolvimentos em Inteligência Artificial (IA) tem impactado diferentes áreas de conhecimento, entre elas o design, com reflexos nos âmbitos social, científico, tecnológico, cultural e econômico. Pesquisadores destacam que a sociedade contemporânea vive um capitalismo de vigilância no qual o volume de dados gerados e disponíveis, além de matéria prima com valor comercial, converteu-se também em um novo meio de colonização. Tal contexto deflagra novas problemáticas a serem consideradas em projetos de design, uma área que já lidava com questões éticas relacionadas ao uso de padrões enganosos em produtos digitais. Portanto, o propósito desse estudo foi com base na literatura, discutir tais cenários e refletir a partir dos princípios éticos para uso da IA recomendados pela UNESCO, quais os possíveis caminhos pelos quais o design enquanto campo poderia colaborar. Sem a pretensão de esgotar o assunto ou definir práticas, a intenção foi contribuir com um debate que tende a se ampliar cada vez mais também no design.

Palavras chave: design – inteligência artificial – impactos sociais – ética

[Resumos em inglês e espanhol na página 59]

^(*) Doutora em Design - UNESP; Mestra em Design - UNESP e Graduada em Desenho Industrial – UNESP. Atualmente é Professora Assistente Doutora no Departamento de Design - FAAC/UNESP.

1. Introdução

Na atualidade, a presença de soluções projetuais para produtos, sistemas ou serviços baseados em Inteligência Artificial – IA – tem se ampliado de maneira acelerada nos diversos âmbitos da vida humana e impactando a sociedade de diferentes maneiras.

Enquanto campo do conhecimento, para Russel e Norvig (2013) a IA é um campo universal que busca não somente a compreensão da inteligência, mas também a sua construção. Os autores identificam que, historicamente, quatro estratégias para o estudo da IA têm sido seguidas a partir de métodos diferentes, as quais resultam em estudos e proposições diferentes. Estas quatro estratégias trazem oito definições do que é a IA dentro de seus espectros de ação, como mostra o Quadro 1:

Pensando como um humano	Pensando racionalmente
<p>“O novo e interessante esforço para fazer os computadores pensarem (...) máquinas com mentes, no sentido total e literal.” (Haugeland, 1985).</p> <p>“[Automatização de] atividades que associamos ao pensamento humano, atividades como a tomada de decisões, a resolução de problemas, o aprendizado...” (Bellman, 1978).</p>	<p>“O estudo das faculdades mentais pelo uso de modelos computacionais.” (Charniak e McDermott, 1985).</p> <p>“O estudo das computações que tornam possível perceber, raciocinar e agir.” (Winston, 1992).</p>
Agindo como seres humanos	Agindo racionalmente
<p>“A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.” (Kurzweil, 1990).</p> <p>“O estudo de como os computadores podem fazer tarefas que hoje são melhor desempenhadas pelas pessoas.” (Rich and Knight, 1991).</p>	<p>“Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (Poole et al., 1998).</p> <p>“Al... está relacionada a um desempenho inteligente de artefatos.” (Nilsson, 1998).</p>

Quadro 1. Definições de IA divididas em quatro estratégias de pesquisa e desenvolvimento. **Fonte:** Russel e Norvig (2013) adaptado pela autora.

Independente de qual dos âmbitos esteja sendo considerado, o fato é que a IA está cada vez mais presente no dia a dia das pessoas e, conforme Santaella (2021, p.8) tem invadido de forma crescente todas as atividades e aspectos da vida humana, incluindo as práticas culturais. Para a autora, conjuntamente à escalada da tecnologia ampliam-se também suas instâncias e os dilemas inerentes à estas.

As instâncias e dilemas que a IA traz à discussão, estão especialmente ligados ao impacto social destas produções. Para Cozman e Neri (2021, p.29) embora definir a IA ainda seja um desafio visto que a definição de “inteligência” é também fluida, atualmente é inegável que ela é parte do mundo real e de fato influencia a sociedade. Segundo os autores, a relação entre ambas – IA e sociedade – configura mais um eixo essencial no estudo de inteligências artificiais. Considerando o impacto social das produções, fica evidente o papel do design enquanto campo do conhecimento responsável por criar as interfaces

entre tecnologia e humanidade no cenário complexo atual (Bonsiepe, 2015; Denis, 2012; Norman, 2010).

O avanço da IA impacta os processos e práticas do design (Cerejo; Carvalhais, 2019; Satterfield; Abel, 2020; Verganti; Vendraminelli; Iansiti, 2020) trazendo oportunidades, mas também problemáticas a serem consideradas. Nesse sentido, as preocupações de ordem ética se fazem presentes como parte integrante de desenvolvimentos na área de origem da IA – a Computação –, e também na área de Design, que lida não somente com o desafio do projeto adequado para comportar as tecnologias inteligentes, como também, com as consequências do uso destas pelos seres humanos (Coeckelbergh, 2020; Boechat, 2019; Dignum, 2019).

Na perspectiva do uso, cabe destacar que o campo do design em sua dimensão digital já lida com questões éticas ligadas ao emprego, por parte das empresas, dos denominados padrões enganosos do design – antigos *dark patterns*, atualmente renomeados *deceptive patterns* – os quais manipulam ou coagem as pessoas usuárias a fazerem escolhas que não são de seu interesse (Sinders, 2022; Narayanan; Mathur; Chetty; Kshirsagar, 2020).

Tais dinâmicas se desdobram em um cenário tido na atualidade como a base de uma nova ordem econômica, o Capitalismo de Vigilância (Zuboff, 2019; Morozov, 2018) no qual as grandes empresas de tecnologia se valem do amplo volume de dados captados na sociedade como matéria-prima para práticas comerciais, mercantilizando e controlando comportamentos, modificando atitudes, práticas e hábitos. A captação e o uso de dados, inclusive, além de ser a base desse novo Capitalismo, ainda serve a um propósito colonizador (Souza; Ribeiro; Nogueira, 2022; Cassino, 2021) que se utiliza das mesmas práticas predatórias do colonialismo histórico, mas que se dão atualmente com base no uso dos dados dos sujeitos, apropriando-se assim da vida humana uma vez que as pessoas não sabem quais de seus dados são coletados, como são utilizados ou quais as fontes coletoras.

Para que se possa operacionalizar esta infinidade de dados – denominados *Big Data* – é necessário o uso de algoritmos, sequência de ações programadas destinadas a tratar, categorizar e filtrar os dados captados. No contexto do aprendizado de máquina é a Inteligência Artificial que constrói seus próprios algoritmos com base em um grande volume de dados – chamados “input” – e na indicação humana dos resultados que se quer e o que não se quer – chamados “outputs”. Assim, o modelo é treinado para que ele aprenda qual o melhor caminho, visando o resultado que se espera. Neste processo, dependendo de como o modelo de IA foi treinado e a partir de quais dados, é possível verificar o surgimento de vieses algorítmicos, os quais refletem nos resultados de interpretação dos dados os preconceitos existentes na sociedade (Cozman; Kaufman, 2022; Silva, 2022; Simões-Gomes; Roberto; Mendonça, 2020).

Neste sentido, são impostos novos desafios para o design enquanto área responsável por humanizar as tecnologias: um Capitalismo de Vigilância que favorece um novo Colonialismo e no qual as pessoas usuárias estão sujeitas a lidar não somente com os padrões enganosos do design, mas também com produtos, sistemas e serviços que podem carregar preconceitos variados. O presente artigo tem propósito discutir tais cenários com base na literatura e refletir, a partir dos princípios éticos para uso da IA recomendados pela UNESCO, quais os possíveis caminhos pelos quais o design enquanto campo pode colaborar.

2. Desafio contemporâneo para o design: os padrões enganosos em design digital

O Design em sua vertente digital trabalha desenvolvendo soluções interativas para aplicativos, websites entre outras soluções na forma de interfaces digitais para variados sistemas e serviços impulsionando negócios no ambiente online.

A busca por fidelizar usuários e converter vendas ao longo do tempo, fez com que no desenvolvimento de soluções digitais, muitos negócios adotassem o emprego de padrões de design que atuam de maneira enganosa, favorecendo a conversão/venda.

Esta prática ganhou um nome e em 2010 o termo “padrões obscuros” de design – ou *dark patterns* – foi cunhado por Harry Brignull, introduzindo-o pela primeira no site *darkpatterns.org*, hoje denominado *www.deceptive.design*. Os padrões originalmente chamados de “*dark patterns*”, na atualidade não são mais assim referenciados por se tratar de um termo racista. Hoje utiliza-se “*deceptive patterns*” – padrões enganosos – em substituição (Sinders, 2022 – tradução nossa).

Em geral, estes elementos de design de interface podem “[...] enganar ou induzir os usuários a comportamentos que muitas vezes beneficiam o lado que implementa o design para o usuário final” (Owens et al., 2022, p. 64 – tradução nossa).

Os autores seguem exemplificando que esses padrões se revelam em práticas como: prover um serviço que facilita a assinatura do usuário com uma única interação, mas dificulta o cancelamento, ou ainda, um site que pode facilitar o consentimento para a coleta de dados das pessoas, mas torna difícil reverter esse consentimento. De todo modo, segundo Owens *et al.* quer o resultado de manipulação seja intencional por parte dos designers, ou seja efeito de um design deficiente – resultado de hábitos, suposições erradas ou prioridades de um designer –, ou ainda de outras restrições da interface, esses tipos de padrões de design “[...] dificultam que os usuários tomem e implementem as decisões que eles fariam em resposta a uma abordagem mais neutra ou de design centrado no usuário - impactando a privacidade, segurança, finanças, autonomia dos usuários, e mais” (Owens et al., 2022, p. 64 – tradução nossa).

Os antecedentes dos padrões enganosos têm décadas de idade. Embora a consciência pública sobre padrões enganosos seja relativamente nova, o próprio fenômeno se desenvolveu gradualmente[...]. A história também ajuda a explicar o que há de novo nos padrões enganosos. Não se trata apenas de um design artificioso ou de um varejo enganoso online. Em vez disso, o design foi transformado em arma usando pesquisa comportamental para servir aos objetivos da economia de vigilância (Narayanan et al. 2020, p.79 – tradução nossa).

Para Nakanayan *et al.* os padrões enganosos permitem aos designers extrair três recursos principais dos usuários: dinheiro, dados e atenção, três fatores essenciais para o funcionamento das dinâmicas instauradas pelo Capitalismo de vigilância em um contexto fortemente orientado pelo Colonialismo de dados.

3. Capitalismo de vigilância, Colonialismo de dados e Viés algorítmico

A escala com que os algoritmos da IA tratam dados, constitui a matéria prima principal para o avanço da IA – e obviamente, na mesma medida, este avanço resulta em impactos sociais benéficos e também problemáticos.

A enorme quantidade de dados disponíveis em rede e utilizados pela IA para diferentes fins são captados pelas empresas de tecnologia e mercantilizados criando assim, uma nova organização para o sistema econômico Capitalista, o que Shoshana Zuboff (2021) denominou como sendo o Capitalismo de Vigilância, o qual define como “[...] Uma nova ordem econômica que reivindica a experiência humana como matéria-prima gratuita para práticas comerciais dissimuladas de extração, previsão e vendas” (Zuboff, 2021, local. 13). Neste sentido, a autora enfatiza que embora alguns destes dados sejam realmente utilizados no aprimoramento de produtos e serviços digitais, o excedente é declarado como *superávit comportamental do proprietário*, sendo então utilizado para manufaturar previsões sobre o comportamento das pessoas.

Estes dados excedentes alimentam um outro mercado voltado ao que Zuboff denomina ser o *mercado de comportamentos futuros* o qual, em um cenário competitivo, passa a ser valioso para as empresas anteciparem seus desenvolvimentos com base na previsão do comportamento das pessoas usuárias, do que elas querem ou precisam. É a partir desta dinâmica que o Capitalismo de Vigilância gera uma nova espécie de poder que a autora denomina *Instrumentalismo*, sintetizando-o da seguinte maneira:

O poder instrumentário conhece e molda o comportamento humano em prol das finalidades de terceiros. Em vez de armamentos e exércitos, ele faz valer sua vontade através do meio automatizado de uma arquitetura computacional cada vez mais ubíqua composta de dispositivos, coisas e espaços “inteligentes” conectados em rede (Zuboff, 2021, local. 21-22).

Zuboff, contudo, destaca que o Capitalismo de Vigilância triunfa, entre outros fatores, por não ter precedentes e por isso confunde a compreensão das circunstâncias por parte das pessoas. Cabe destacar, contudo, que a tecnologia empregada não é o problema uma vez que no Capitalismo de Vigilância não é a tecnologia em si, mas sim a lógica que permeia a tecnologia e a direciona à ação (2021, local.31).

Esta lógica de dominação por meio dos dados favorece o que Nick Couldry e Ulises A. Mejias (2019) denominam colonialismo de dados. A partir da reflexão destes autores, Cassino (2021) destaca que a partir do que ocorreu com o colonialismo tradicional é possível compreender o que ocorre hoje com o avanço das Tecnologias de Informação e de Comunicação (TIC), resultando em um possível novo formato de colonialismo o qual é baseado em dados.

O colonialismo tradicional atuou de modo a dominar as estruturas políticas, militares, sociais e econômicas de um outro povo, ocorrendo o fenômeno em toda a América, a partir do século XVI, e durante os séculos XVIII, XIX e XX, na África, Ásia e Oceania,

engendrado pelas nações europeias. O fim do colonialismo, contudo, não significa o fim da colonialidade como destacado por Souza, Ribeiro e Nogueira (2022):

O colonialismo se finda quando a sociedade subalternizada se torna independente politicamente de seu colonizador. Entretanto, a colonialidade é composta pelos padrões culturais, epistemológicos, subjetivos e sociais, como as relações de trabalho, as formas de produzir conhecimentos, e as determinações de raça e gênero, por exemplo, que permaneceram de herança do período colonial, e que, atualmente, ainda constituem as relações de poder, os embates políticos e sociais presentes em todas as ex-colônias (Souza; Ribeiro; Nogueira 2022, p.90).

Para Couldry e Mejias (2019, p. 75 – tradução nossa) “[...] O legado do colonialismo explica não só a dívida e a pobreza, mas também outras questões importantes do nosso tempo, como o racismo, a migração, a falta de educação oportunidades e terrorismo”. Como resultado, há uma diferença entre as oportunidades que se abrem em termos de avanços sociais e tecnológicos para a nações mais ricas – colonizadoras – quando comparadas às nações exploradas, que permanecem à margem destes desenvolvimentos – colonizadas. Tem-se então, que a colonialidade segue influenciando diferentes âmbitos destas sociedades e na contemporaneidade, exploram-nas de outra maneira: também no cenário digital. Neste formato, o colonialismo de dados:

[...] combinaria as mesmas práticas predatórias do colonialismo histórico com a quantificação abstrata de métodos computacionais. Trata-se de um novo tipo de apropriação no qual as pessoas ou as coisas passam a fazer parte de infraestruturas de conexão informacionais. A apropriação da vida humana (por meio da captura em massa de dados) passa a ser central (Cassino, 2021, p. 27-28).

O autor enfatiza que um dos efeitos mais marcantes sobre as pessoas que são colonizadas neste formato é que estas passam a ficar atados a julgamentos que são feitos com base em seus próprios dados sendo que não sabem, na realidade, quais de seus dados são coletados, por quais fontes e como são utilizados. Sabe-se, contudo, que as informações pessoais capturadas são a chave para as novas formas de geração de valor e, portanto, com potencial mercadológico.

Tal cenário se torna ainda mais problemático uma vez que evidencia que, como resultado, as nações mais ricas se tornaram as principais produtoras das tecnologias, ao passo que as nações mais pobres se tornam as usuárias das tecnologias providas pelas primeiras, o que perpetua as disparidades de acesso e benefícios.

Para Couldry e Mejias (2019) a vida humana está sendo colonizada por dados e precisa ser descolonizada. Os autores (2021, p. XIX – tradução nossa) destacam que o resultado do processo de colonização “[...] degrada a vida, primeiro por expô-la continuamente ao monitoramento e vigilância (através do qual dados são extraídos) e, em segundo lugar, tornando a vida humana uma entrada direta para a produção capitalista.” Para os autores

a decolonialidade proposta como base para esta descolonização alinha-se à perspectiva decolonial de Aníbal Quijano e o grupo conhecido como *Modernidad/Colonialidad ou Proyecto M/C* que segundo Souza, Ribeiro e Nogueira (2022):

[...] a qual objetiva fazer um estudo crítico das amarras renascentes da colonização presentes na América Latina, como o processo de racialização dos povos não-europeus, visando construir novas bases de pensamento e conhecimentos que tenham como protagonistas as comunidades subalternizadas e suas epistemologias (Souza; Ribeiro; Nogueira 2022, p.89).

A decolonização é necessária para que se possa ler criticamente as circunstâncias que cercam os seres humanos na atualidade, em um contexto no qual se desenrola a apropriação de recursos sociais por meio dos dados, operando de modo semelhante à expropriação de recursos naturais. Segundo Couldry e Mejias:

As relações com os dados são novos tipos de relações humanas que dão às corporações uma visão abrangente da nossa sociabilidade, permitindo que a vida humana se torne um insumo ou recurso para o capitalismo. Neste esquema neocolonial, a colônia não é uma localização geográfica, mas uma “realidade melhorada” na qual conduzimos nossas interações sociais sob condições de extração contínua de dados. Os recursos que estão sendo colonizados são as associações, normas, códigos, conhecimentos e significados que nos ajudam a manter conexões sociais, os processos humanos e materiais que constituem a atividade econômica, e o espaço do sujeito a partir do qual enfrentamos o mundo social (Couldry; Mejias, 2019, p.85 – tradução nossa).

Um efeito preocupante do amplo uso de dados, neste cenário é a existência do viés algorítmico. No âmbito da área de Computação, os algoritmos são “[...] uma sequência de instruções a serem cumpridas em uma determinada ordem, ou “procedimentos codificados que, com base em cálculos específicos, transformam dados em resultados desejados” (Gillespie, 2018 Apud Simões-Gomes; Roberto; Mendonça, 2020 P.141). Tais algoritmos são os responsáveis por tratar, categorizar e filtrar os dados com os quais a IA trabalha, mas nem sempre este processo traz somente os resultados desejados. Ao operacionalizar um grande volume de dados no retorno da IA podem surgir resultados enviesados – ou resultados discriminatórios por gênero, raça, etnia, entre outros (Kaufman, 2022, local.11) – que perpetuam preconceitos:

Em geral, atribui-se vieses integralmente às bases de dados tendenciosas. Porém, vieses podem emergir antes da coleta de dados em função das decisões tomadas pelos desenvolvedores (os atributos e variáveis contemplados no modelo, inclusive, determinam a seleção dos dados). No caso de viés associado aos dados, existem duas principais origens: os dados coletados não representam a composição proporcional do universo objeto em questão, ou os

dados refletem os preconceitos existentes na sociedade (Cozman; Kaufman, 2022, p. 198).

Segundo Kaufman (2022, local.13), distorções deste tipo são comuns nos sistemas de IA e remetem à importância da diversidade na formação das equipes que elaboram esses sistemas, uma vez que os desenvolvedores tem o foco na funcionalidade da solução e por vezes não contemplam os impactos éticos e sociais desses desenvolvimentos. Dessa forma, o viés algorítmico se estabelece – juntamente aos padrões enganosos – como mais um desafio importante para a área de design lidar na prática de projetos que envolvam a IA, dentro do contexto de Colonialismo de dados inserido no modo de operar do Capitalismo de vigilância.

Considerando estes desafios, torna-se necessário refletir sobre como a IA potencializa cada um deles, tanto no âmbito dos desenvolvimentos, quanto no âmbito do uso de produtos, sistemas e serviços. Lidar de maneira coerente e justa com estes pontos só é possível a partir de uma perspectiva em que a ética seja central no processo.

Embora o avanço da IA traga importantes e significativas oportunidades para diversas áreas do conhecimento humano, é preciso também considerar que, além dos benefícios, se formam também – como visto anteriormente – panoramas problemáticos em termos de impacto social.

Considerando os enfrentamentos que possam ser necessários no futuro a UNESCO elaborou um documento com recomendações sobre o uso ético de Inteligência Artificial.

4. Recomendações sobre a ética da Inteligência artificial – UNESCO e como o design pode colaborar

Os fatores até aqui discutidos evidenciam o quanto a prática do design é influenciada por tomadas de decisão mercadológicas que muitas vezes colocam em segundo plano as pessoas usuárias.

No contexto de aceleração da IA e com a inevitável incorporação desta tecnologia em projetos de design, tais problemáticas – e outras que podem vir a surgir em decorrência do avanço da IA – ganham nova dimensão, daí a necessidade da regulação.

Esforços tem sido feitos por diferentes nações na busca por regulamentar a IA, inclusive no Brasil por meio do Projeto de lei 2338/23 que dispõe sobre o uso da Inteligência Artificial (Brasil, 2023) e que tramita atualmente no Senado Federal brasileiro. Cabe destacar, contudo, que especialistas destacam que a urgência em aprovar uma regulação logo pode prejudicar o processo, que necessita de um tempo de maturação para que o debate seja mais amplo como destaca Dora Kaufman em reportagem da Revista FAPESP “A regulamentação prematura pode restringir a inovação e não proteger a sociedade [...] O processo é tão importante quanto o resultado final” (Schmidt, 2023).

Compreendendo a urgência da pauta, até que os Estados – entre eles o Brasil – desenvolvam suas regulamentações sobre a aplicação ética da IA, a UNESCO se antecipou e elaborou um documento intitulado “Recomendações sobre a ética da Inteligência Artificial” (UNESCO, 2022).

O documento considera as diferentes dimensões nas quais o uso da IA pode levantar questões éticas. Para contribuir com o debate e favorecer ações neste âmbito, a publicação destaca dez princípios orientadores que podem ser utilizados como referência pelos Estados no desenvolvimento de estruturas ou mecanismos para lidar com desenvolvimentos em IA de maneira ética visando salvaguardar os seres humanos e seus direitos ao interagirem com produções que envolvam IA.

Entendendo que a ética para o desenvolvimento da IA deve ser também a base dos projetos em design que a integrem, o presente estudo propõe, com base nos dez princípios elencados pela UNESCO, estratégias projetuais que podem colaborar com produtos, sistemas e serviços baseados em IA visando produções éticas.

Sem a pretensão de esgotar o assunto, o que se propõe nesta pesquisa é que a prática do design ofereça recursos para que as pessoas usuárias se sintam informadas e seguras no uso de soluções de design que envolvam tecnologias inteligentes. As propostas aqui elencadas não são definitivas ou mesmo imutáveis – e certamente não serão as únicas – uma vez que tanto a área de design, quanto a da IA, se transformam em resposta às mudanças e demandas constantes nos âmbitos sociais, culturais, ambientais e econômicos. Também os perfis dos produtos, sistemas e serviços projetados terão diferentes recursos e restrições, assim, as estratégias adotadas para um produto podem diferir substancialmente de outro. Em qualquer que seja o caso, as estratégias aqui apresentadas buscam ser o mais abrangentes possível podendo ser aplicadas no processo de design – como parâmetros norteadores – e dentro dos produtos prontos – como recursos.

Isto posto, a seguir, estão elencados cada um dos dez princípios apresentados pela UNESCO juntamente com as estratégias propostas pela pesquisadora.

O primeiro princípio é o de *Proporcionalidade e não causar dano* e para ele o documento destaca o cuidado para que nenhum dos processos relacionados ao ciclo de vida dos sistemas de IA exceda o necessário para atingir objetivos ou metas legítimas e aponta que estes devem ser adequados a cada contexto. É essencial também a implementação de procedimentos para avaliação de risco e a adoção de medidas para impedir a ocorrência de danos tanto aos seres humanos, quanto ao meio ambiente (UNESCO, 2022, p.20). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Captar somente dados necessários; Informar riscos; Destacar direitos; Auxiliar denúncias; Garantir o anonimato sempre que possível; Permitir que usuários retirem o consentimento de uso e Permitir que usuários solicitem o apagamento de dados a qualquer momento.

O segundo princípio é *Segurança e proteção*, e nele o documento destaca a necessidade de garantir proteção e segurança no uso de sistemas da IA sendo que os danos indesejados, tal qual os riscos de proteção e as vulnerabilidades a ataques que configuram riscos de segurança devem ser evitados e abordados, impedidos e eliminados durante o ciclo de vida de tais sistemas. A publicação destaca também que a elaboração de marcos sustentáveis de proteção à privacidade dos dados de acesso são ações que permitirão que melhores modelos de IA sejam desenvolvidos permitindo o desenvolvimento de soluções em IA mais seguros e protegidos (UNESCO, 2022, p.20). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Captar somente dados necessários; Informar riscos; Destacar direitos; Oferecer

recursos de proteção facilitados; Ensinar boas práticas de segurança ao usuário; Permitir que usuários retirem o consentimento de uso e Permitir que usuários solicitem o apagamento de dados a qualquer momento.

O terceiro princípio é o de *Justiça e não discriminação* e enfatiza o cuidado necessário para a promoção da justiça social, equidade e não discriminação de qualquer tipo, respeitando o direito internacional e a partir de uma abordagem inclusiva com foco em assegurar que os benefícios das tecnologias de IA estejam disponíveis e acessíveis a todos. Tais ações, para serem adequadas, devem levar em consideração as especificidades dos diferentes grupos que possam ser desfavorecidos ou marginalizados atuando também contra a discriminação e a determinação algorítmica tendenciosa (UNESCO, 2022, p.20-21). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Captar somente dados necessários; Informar riscos; Destacar direitos; Auxiliar denúncias; Garantindo o anonimato quando do tratamento dos dados; Permitir que usuários retirem o consentimento de uso e Permitir que usuários solicitem o apagamento de dados a qualquer momento.

O quarto princípio é o de *Sustentabilidade*, e enfatiza que para que seja possível alcançar sociedades mais sustentáveis é necessário que sejam estabelecidas ações no âmbito tecnológico, e isto envolve realizar avaliações contínuas dos impactos das tecnologias de IA nos âmbitos humanos, sociais, culturais, econômicos e ambientais. Tais ações devem constantes e envolver o pleno conhecimento das implicações que estas tecnologias trazem para a sustentabilidade, constituindo um conjunto de objetivos a serem monitorados e em constante evolução (UNESCO, 2022, p.21). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Captar somente dados necessários; Informar riscos; Destacar direitos; Auxiliar denúncias; Garantir que os produtos desenvolvidos auxiliem diferentes realidades; Garantir que os produtos desenvolvidos promovam a emancipação e Promovendo benefícios adequados para realidades desiguais.

O quinto princípio é *Direito à privacidade e proteção de dados*, no qual é destacado que a privacidade configura um direito essencial para proteger a dignidade, a autonomia e a capacidade de ação humanas e deve ser respeitada, protegida e promovida ao longo do ciclo de vida dos sistemas de IA. Assim, os dados captados devem ser coletados, utilizados, compartilhados, arquivados e apagados em consonância respeitando o direito internacional de modo a estarem alinhados com os valores e princípios estabelecidos na Recomendação da UNESCO discutida no presente artigo, respeitando também os marcos jurídicos nacionais, regionais e internacionais relevantes (UNESCO, 2022, p.22). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Captar somente dados necessários; Informar riscos; Destacar direitos; Auxiliar denúncias; Garantir o atendimento às Leis de proteção de dados dentro do produto; Reportar erros do sistema para usuário e entidades legais; Permitir que usuários retirem o consentimento de uso e Permitir que usuários solicitem o apagamento de dados a qualquer momento.

O sexto princípio é *Supervisão humana e determinação* cujo foco é garantir que sempre seja possível atribuir responsabilidade ética e legal em qualquer estágio do ciclo de vida dos sistemas de IA. Neste caso, entende-se a supervisão como não somente a supervisão

humana individual, mas também a supervisão pública inclusiva desses sistemas e seus impactos. Este princípio destaca que as pessoas podem recorrer à IA por motivos de eficácia para alguma necessidade ou tarefa, mas a decisão de ceder o controle em contextos limitados continua sendo de responsabilidade humana sendo, portanto, também humana a prestação de contas. Tal princípio também destaca que decisões de vida e morte não devem ser transferidas a sistemas de IA (UNESCO, 2022, p.22). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Identificar para os usuários os responsáveis pelo sistema em uso; Auxiliar denúncias; Informar como o sistema é auditado ou supervisionado; Garantir que o usuário possa definir suas preferências de interação/uso; Reportar erros do sistema para usuário e entidades legais e Permitindo que usuários possam solicitar responsabilização por parte do fabricante.

O sétimo princípio é *Transparência e explicabilidade* item no qual o documento enfatiza a importância das soluções em IA serem transparentes e explicáveis para os envolvidos, sendo ambas essenciais para garantir o respeito, a proteção e a promoção dos direitos humanos, das liberdades fundamentais e dos princípios éticos. Por meio da transparência é possível permitir que as pessoas entendam o impacto e como cada estágio de um sistema é colocado em prática, considerando também o contexto e a sensibilidade do sistema de IA. Já a explicabilidade significa tornar inteligível e também fornecer informações sobre o resultado dos sistemas de IA. A explicabilidade está diretamente relacionada com a transparência, uma vez que os resultados devem ser compreensíveis e rastreáveis, conforme o contexto (UNESCO, 2022, p.22-23). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Identificar para os usuários os responsáveis pelo sistema em uso; Explicar como o sistema funciona de maneira fácil para o usuário; Explicar o impacto de cada escolha no uso; Garantir que o usuário possa definir suas preferências de interação/uso; Reportar erros do sistema para usuário e entidades legais e Permitir o rastreamento dos resultados do uso.

O oitavo princípio é *Responsabilidade e prestação de contas* no qual a publicação defende que os atores de IA e os Estados-membros devem assumir responsabilidades éticas e jurídicas conforme estabelecido pelo direito nacional e internacional. Assim, é necessário o desenvolvimento de mecanismos para supervisão, avaliação de impacto, auditoria e diligência, incluindo proteção de denunciantes, visando garantir que exista a responsabilização pelos sistemas de IA, bem como seu impacto social (UNESCO, 2022, p.23). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Identificar para os usuários os responsáveis pelo sistema em uso; Explicar como os usuários podem recorrer à justiça caso necessário; Auxiliar denúncias; Informar como o sistema é auditado ou supervisionado; Reportar erros do sistema para usuário e entidades legais e Permitir o rastreamento dos resultados do uso.

O nono princípio é *Conscientização e alfabetização* e nele são destacadas a importância e a necessidade da promoção da compreensão pública das tecnologias de IA, bem como do valor dos dados, fazendo isso por meio de educação aberta e acessível, do engajamento cívico, das habilidades digitais e também do treinamento em ética da IA. Também enfatiza a alfabetização midiática e informacional, além de treinamento que seja conduzido em

conjunto por diferentes atores como governos, organizações intergovernamentais, sociedade civil, universidades, meios de comunicação, líderes comunitários e setor privado. Ainda é enfatizado que a compreensão dos sistemas de IA devem ser baseadas em seu impacto nos direitos humanos, no meio ambiente e nos ecossistemas (UNESCO, 2022, p.23). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Conscientizar por meio do ensino na formação de novos profissionais; Conscientizar a sociedade por meio de cursos gratuitos oferecidos à diferentes públicos; Alfabetizar no uso dentro do próprio produto ao informar riscos e benefícios; Informar sobre iniciativas de conscientização e alfabetização externos dentro dos produtos; Colaborar na criação de políticas públicas problematizando a IA por meio do Design e Permitindo o apoio a públicos vulneráveis – não letrados.

O décimo princípio é *Governança e colaboração adaptáveis e com múltiplas partes interessadas*, este princípio estabelece que o direito internacional e a soberania nacional devem ser respeitados no uso de dados e que os Estados podem regulamentar tanto dados gerados dentro de seus territórios como também dados que os atravessam, além de poder tomar medidas para uma regulamentação de dados eficaz com base no respeito ao direito à privacidade e em conformidade com as leis internacionais e outras normas e padrões de direitos humanos (UNESCO, 2022, p.23). Para colaborar com o cumprimento deste princípio, as estratégias de design seriam: Desenvolver pesquisas que promovam o debate; Colaborar na criação de políticas públicas problematizando a IA por meio do Design; Advogar pelos direitos dos usuários independente dos atores envolvidos (governo/setor privado/academia/sociedade civil); Respeitando as legislações vigentes nos diferentes países; Reportar erros do sistema para usuário e entidades legais e Auxiliar no estabelecimento de diretrizes de regulação para produtos, sistemas e serviços baseados em IA sempre que necessário ou pertinente.

5. Considerações finais

A área do design é condicionada às mudanças de ordem social, cultural, ambiental e tecnológica. Neste sentido, as mudanças que surgem no horizonte para o design a partir da integração da Inteligência Artificial nos projetos permitirá não somente novas possibilidades e dimensões de interação, como trará também novos desafios perante àqueles que já atingem o design como a colonização de dados inserida em um contexto de capitalismo de vigilância, os vieses algorítmicos e o uso de padrões enganosos. Neste cenário, devem ser permanentes as reflexões críticas sobre como a IA potencializa cada um deles, tanto no âmbito dos desenvolvimentos, quanto no âmbito do uso de produtos, sistemas e serviços.

Atenta para os diversos cenários que podem ser resultantes do uso da IA e seus impactos na sociedade, a UNESCO propôs recomendações sobre a ética da IA e a partir destas se pôde inferir no presente trabalho algumas ações por meio das quais o design pode colaborar, tanto no âmbito de projeto quanto no âmbito do uso.

Obviamente não se pretendeu esgotar o assunto, muito menos afirmar que estas sejam as únicas ações possíveis, uma vez que as áreas de design e da IA tendem a se ampliar em resposta às demandas sociais, culturais, ambientais e econômicas. Neste sentido, o perfil dos projetos futuros e suas dimensões trarão novos questionamentos e diferentes necessidades. Com base neste panorama incerto e entendendo ser o design a área responsável por humanizar as tecnologias, a autora do presente estudo defende que missão do design na era da IA será não somente a de projetar soluções que integram a IA e suas oportunidades, mas também terá um caráter vigilante sobre o impacto social destes desenvolvimentos.

A área de design continuará antecipando problemas para buscar soluções, mas agora em dimensão e velocidade maiores dado o fato de que a aceleração da IA certamente acelerará o design como nunca antes. Nesta perspectiva, não teremos somente *problemas de projeto que integrem a IA* do ponto de vista de necessidades dos usuários, mas teremos também *problemas de projeto como consequência do uso de designs que integrem a IA*. Em síntese: o design não só projetará para integração da IA, mas também para as consequências desta integração quando em uso pela sociedade; o design não só projetará para os usuários, mas também para lidar com o impacto social deste uso – que pode se tornar imprevisível. Mais do que projetar, portanto, na era da IA designers devem estar vigilantes e seu papel será o de questionar cada vez mais o contexto de desenvolvimento das soluções inteligentes projetadas.

Referências:

- Boechat, Cid. (2019). Inteligência Artificial, empatia e inclusão: um problema de design. *Ergodesign & HCI*, n especial, v. 7, ano 7. Disponível em: <http://periodicos.puc-rio.br/index.php/revistaergodesign-hci/article/view/1306/735>. Acesso em: 8 abr.2023.
- Bonsiepe, Gui. (2015). *Design do material ao digital*. São Paulo: Blucher.
- Brasil. Projeto de Lei nº 2338, de 2023. (2023) Dispõe sobre o uso da Inteligência Artificial. *Plenário do Senado Federal (Secretaria Legislativa do Senado Federal)*. Brasília, DF. Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233#tramitacao_10494842. Acesso em: 10 jul. 2023.
- Brignull, Harry. (2023). *What are deceptive patterns?* Disponível em: <https://www.deceptive.design/>. Acesso em: 10 ago. 2023.
- Cassino, João Francisco. (2021). O sul global e os desafios pós-coloniais na era digital. In: Silveira, Sérgio Amadeu; Souza, Joyce; Cassino, João Francisco. *Colonialismo de dados: Como opera a trincheira Algorítmica na guerra neoliberal*. São Paulo: Autonomia literária.
- Cerejo, Joana; Carvalhais, Miguel. (2019). Design After the Rise of AI-Driven Services: Learning from Literature Review. *DIGICOM 3rd International Conference on Digital Design & Communication*. Disponível em: https://www.researchgate.net/publication/337631988_Design_After_the_Rise_of_AI-Driven_Services_Learning_from_Literature_Review. Acesso em: 8 abr. 2021.
- Coeckelbergh, Mark. (2020). *AI ethics*. Description: Cambridge, MA : The MIT Press.

- Couldry, Nick.; Miejas, Ulisses. A. (2019). *Costs of connection – how data is colonizing human life and appropriating it for Capitalism*. Califórnia: Satanford University Press.
- Cozman, Fábio. G; Neri, Hugo. (2021) O que, afinal, é Inteligência Artificial? In: Cozman, Fabio G; Plonski, Guilherme Ary; Neri, Hugo. *Inteligência artificial: avanços e tendências*. São Paulo: Instituto de Estudos Avançados. Edição digital – e-Book.
- Cozman, Fabio G; Kaufman, Dora. (2022). Viés no aprendizado de máquina em sistemas de inteligência artificial: a diversidade de origens e os caminhos de mitigação. *Revista USP*, (135), 195-210. Disponível em: <https://www.revistas.usp.br/revusp/article/view/206235>. Acesso em: 11 ago.2023.
- Denis, Rafael C. (2012). *Design para um mundo complexo*. São Paulo: Cosac Naify.
- Dignum, V. (2019). Responsible Artificial Intelligence - How to Develop and Use AI in a Responsible Way. In: Petke, Justyna. *Artificial Intelligence: Foundations, Theory, and Algorithms*. Springer.
- Kaufman, Dora. (2022). *Desmistificando a Inteligência Artificial*. Belo Horizonte: Autêntica. Edição digital – e-Book.
- Morozov, Evgeny. (2018). *Big Tech: a ascensão dos dados e a morte da política*. traduzido por Claudio Marcondes. São Paulo: Ubu Editora.
- Narayanan, Arvind; Mathur, Arunesh; Chetty, Marshini; Kshirsagar, Mihir. (2020). Dark Patterns Past, Present, and Future: The Evolution of tricky user interfaces. *ACM Queue*. Volume 18. Issue 2 March-April. pp. 67–92. Disponível em: <https://doi.org/10.1145/3400899.3400901>. Acesso em: 30 ago.2023.
- Norman, Donald. (2010). *O design do futuro*. Rio de Janeiro: Rocco.
- Owens, Kentrell; Gunawan, Johanna; Choffnes, David; Emami-naeini, Pardis; Kohno, Tadayoshi; Roesner, Franziska. (2022). Exploring Deceptive Design Patterns in Voice Interfaces. *EuroUSEC '22: Proceedings of the 2022 European Symposium on Usable Security*. September. Pages 64–78. Disponível em: <https://doi.org/10.1145/3549015.3554213>. Acesso em: 4 ago. 2023.
- Russell, Stuart; Norvig, Peter. (2013). *Inteligência artificial*. 3 ed. tradução Regina Célia Simille. Rio de Janeiro: Elsevier.
- Santaella, Lucia. (2021). *Inteligência artificial e cultura: oportunidades e desafios para o sul global*. Centro UNESCO. CETIC.BR/NIC.BR. Disponível em: <https://cetic.br/pt/publicacao/inteligencia-artificial-e-cultura-oportunidades-e-desafios-para-o-sul-global/>. Acesso em: 02 jul. 2022.
- Satterfield, Debra., Abel, Troy D. (2020). AI Is the New UX: Emerging Research Innovations in AI, User Experience, and Design as They Apply to Industry, Business, and Education, and Ethics. In: Spohrer J., Leitner C. (eds) *Advances in the Human Side of Service Engineering. AHFE 2020. Advances in Intelligent Systems and Computing*, vol 1208. Springer, Cham. https://doi.org/10.1007/978-3-030-51057-2_26. Acesso em: 20 abr. 2021.
- Schmidt, Sarah. (2023). Os desafios para regulamentar o uso da inteligência artificial - Brasil, Canadá e países da Europa elaboram legislação para reduzir os riscos de emprego inadequado de programas e aplicativos dessa área. *Revista Pesquisa FAPESP*. Edição 331 set. Disponível em: <https://revistapesquisa.fapesp.br/os-desafios-para-regulamentar-o-uso-da-inteligencia-artificial/>. Acesso em: 10 set.2023.

- Silva, Tarcízio. (2022). *Racismo algorítmico: inteligência artificial e discriminação nas redes digitais*. São Paulo: Edições Sesc São Paulo.
- Simões-Gomes, Letícia; Roberto, Enrico; Mendonça, Jônatas. (2020). Viés algorítmico—um balanço provisório. *Estudos de Sociologia*, v. 25, n. 48 . Disponível em: <https://nev.prp.usp.br/publicacao/vies-algoritmico-um-balanco-provisorio/>. Acesso em: 28 ago.2023.
- Sinders, Caroline. (2022). *What's In a Name? Unpacking Dark Patterns versus Deceptive Design*. Medium. Disponível em: <https://medium.com/@carolinesinders/whats-in-a-nameunpacking-dark-patterns-versus-deceptive-design-e96068627ec4>. Acesso em: 28 ago.2023.
- Souza, Mariane Pizarro de.; Ribeiro, Paulo Rennes Marçal.; Nogueira, Claudete de Souza. (2022). Decolonialidade e educação antirracista: interseções e aproximações. *Revista Em Favor De Igualdade Racial*, 5(3), 88–98. <https://doi.org/10.29327/269579.5.3-8>. Disponível em: <https://periodicos.ufac.br/index.php/RFIR/article/view/6077>. Acesso em: 11 ago.2023.
- UNESCO. (2022). Organização das Nações Unidas para a Educação, a Ciência e a Cultura. *Recomendações sobre a ética da Inteligência Artificial*. Disponível em < https://unesdoc.unesco.org/ark:/48223/pf0000381137_por>. Acesso em 08 jun. 2022.
- Verganti, Roberto; Vendraminelli, Luca; Iansiti, Marco. (2020). Innovation and Design in the Age of Artificial Intelligence. *Journal Prod Innovation Management*, 37(3):212–227. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1111/jpim.12523>. Acesso em: 30 ago.2023.
- Zuboff, Shoshana. (2021). *A Era do capitalismo de vigilância: a luta por um futuro humano na nova fronteira do poder*. Rio de Janeiro: Intrínseca. Edição digital – e-Book.

Abstract: Developments in Artificial Intelligence (AI) have impacted different areas of knowledge, including design, with repercussions in the social, scientific, technological, cultural and economic spheres. Researchers highlight that contemporary society lives in surveillance capitalism in which the volume of data generated and available, in addition to raw material with commercial value, has also become a new means of colonization. This context triggers new issues to be considered in design projects, an area that already dealt with ethical issues related to the use of misleading patterns in digital products. Therefore, the purpose of this study was, based on the literature, to discuss such scenarios and reflect, based on the ethical principles for the use of AI recommended by UNESCO, on the possible paths through which design as a field could collaborate. Without the intention of exhausting the subject or defining practices, the intention was to contribute to a debate that tends to expand increasingly also in design.

Keywords: design – artificial intelligence – social impacts – ethics

Resumen: Los avances en Inteligencia Artificial (IA) han impactado diferentes áreas del conocimiento, incluido el diseño, con repercusiones en los ámbitos social, científico, tecnológico, cultural y económico. Los investigadores destacan que la sociedad contempo-

ránea vive en un capitalismo de vigilancia en el que el volumen de datos generados y disponibles, además de materia prima con valor comercial, se ha convertido también en un nuevo medio de colonización. Este contexto desencadena nuevas cuestiones a considerar en los proyectos de diseño, un área que ya abordaba cuestiones éticas relacionadas con el uso de patrones engañosos en productos digitales. Por lo tanto, el propósito de este estudio fue, a partir de la literatura, discutir tales escenarios y reflexionar, a partir de los principios éticos para el uso de la IA recomendados por la UNESCO, sobre los posibles caminos a través de los cuales el diseño como campo podría colaborar. Sin ánimo de agotar el tema ni definir prácticas, la intención fue contribuir a un debate que tiende a expandirse cada vez más también en el diseño.

Palabras clave: diseño – inteligencia artificial – impactos sociales – ética

[Las traducciones de los abstracts fueron supervisadas por su autor]
