



O DESIGN NA ERA DA INTELIGÊNCIA ARTIFICIAL: REFLEXÕES SOBRE A ÉTICA DESDE O DESIGN E A IMPORTÂNCIA DO CONTEXTO NESTE PROCESSO

CAROLINA VAITIEKUNAS PIZARRO

DOI 10.52050/9788579176753.11



A Inteligência Artificial tem permeado a sociedade de maneira irreversível impactando a vida das pessoas geralmente por meio de projetos de design. O impacto social destas produções tem levantado preocupações em diversos órgãos, instituições e governos, os quais concentram esforços na busca por debater para poder regular a IA visando salvaguardar os direitos dos seres humanos. Foi nesta perspectiva que o termo Ethics by design – ou Ética desde o Design em tradução livre – surgiu.

(...)

IMAGEM DE ABERTURA

Image by rawpixel.com on Freepik. Disponível em: "https://www.freepik.com/free-photo/technology-human-touch-background-modern-remake-creation-adam_17851045.htm#query=robot%20hand&position=0&from_view=search&track=ais">Image by rawpixel.com"

O presente capítulo trata de oferecer ao leitor um breve panorama sobre a IA e os desafios de ordem ética que esta tecnologia encerra, direcionando a reflexão para uma abordagem da Ética desde o Design que inclui a IA – proposta por Brey e Dainow (2023). A partir do entendimento da abordagem e considerando que o contexto é fator orientador de projetos em design, a autora deste capítulo defende que, diante dos impactos sociais da IA, o papel do design e do designer não estará mais restrito ao processo de projeto e à análise de contexto comumente realizada na fase de descoberta e pesquisa do projeto. Na era da IA e a partir dos impactos sociais dela, o papel do designer estará diretamente relacionado ao monitoramento e análise dos variados contextos gerais, interpretando-os de maneira cada vez mais aprofundada, compreendendo suas particularidades e subjetividades, atuando como uma força estratégica preditiva que informa o projeto de maneira constante, favorecendo a iteração e, portanto, colaborando para o desenvolvimento de soluções mais éticas.

A sociedade atual está envolvida em soluções baseadas em Inteligência Artificial (IA), as quais são incorporadas de maneira cada vez mais veloz ao cotidiano das pessoas e, frequentemente, por meio de um projeto de design.

O uso de tecnologias baseadas em IA permeia produtos, sistemas e serviços impactando diretamente a sociedade, com previsão de crescimento exponencial nos próximos anos conforme se ampliam as pesquisas, a capacidade computacional, o volume de dados disponíveis e as oportunidades que se abrem em consequência destes desenvolvimentos. A aceleração verificada na contemporaneidade no campo da IA configura o que Suleyman e Bhaskar (2023) denominam de “a nova onda” tecnológica que já faz parte do tecido social:

De modo contínuo, a IA deixa o reino das demonstrações e entra no mundo real. Em alguns anos, as IAs serão capazes de conversar, raciocinar e mesmo agir no mesmo mundo que nós. Seus sistemas sensoriais serão tão bons quanto os nossos. Isso não equivale a superinteligência [...], mas prevê sistemas incrivelmente

poderosos. Significa que a IA se tornará uma parte inextricável do tecido social (Suleyman; Bhaskar, 2023, local. 1123).

Para o design, enquanto área responsável por projetos voltados aos seres humanos, está posto o desafio de projetar as soluções que estabelecem o contato entre as pessoas e a IA, uma tecnologia cuja expansão se dá em um cenário marcado pela complexidade e no qual o design não se ocupa somente do projeto de produtos gráficos, de uso ou digitais, como também – e cada vez mais – do projeto de serviços e daqueles voltados a projetos para as experiências dos usuários (Boechat, 2019; Cerejo; Carvalhais, 2019).

Urge então, pensar a prática do design considerando não somente a presença da IA como tecnologia integrante das soluções projetuais no presente, como também no futuro, que ainda não se desenhou, mas que certamente apresentará uma IA mais robusta e cada vez mais acurada.

Muitas das soluções em IA que beneficiam os seres humanos atualmente permitem inferir que o cenário para o futuro é promissor e trará benefícios variados, contudo, é também verdade que trará múltiplos desafios, principalmente do ponto de vista ético e dos impactos sociais dessas soluções.

Os efeitos da IA no processo de design e na vida das pessoas é um debate do qual os designers não podem se isentar, sendo este um tópico que demanda agora – e demandará no futuro – a produção constante de pesquisas para compreender contextos, explorar cenários e buscar por diferentes possibilidades de interação e integração entre os campos do design e da IA, de modo que as soluções resultantes dessas pesquisas sejam benéficas, éticas e seguras.

Visando colaborar com investigações sobre os possíveis caminhos para projetos de design que integrem a IA de maneira ética, o presente capítulo no formato de ensaio busca trazer mais reflexões sobre o fazer do design inserido na era da IA, uma tecnologia cujas implicações influenciam – e

influenciarão – o processo de projeto, o próprio projeto, o seu uso e, especialmente, o contexto para o qual se projeta.

Considerando os aspectos éticos no cenário de desenvolvimento da IA, um termo que vem ganhando destaque é o Ethics by Design ou Ética desde o Design.

O foco das reflexões apresentadas a seguir é discutir o modelo proposto por Brey e Dainow (2023) de Ética desde o Design para a IA a partir da perspectiva do design, visando ponderar sobre a prática projetual que se desenvolve em estreita colaboração com diversas áreas – incluindo a computação – em e para contextos complexos como esse, no qual a IA se expande atualmente.

Nesse sentido, entendendo-se que o design nasce de um contexto diretamente afetado pelos impactos sociais da IA, defende-se neste ensaio que o papel do design e do designer na era da IA estará, para além do projeto, voltado para o monitoramento e análise de contextos de modo a detectar

impactos sociais possíveis ou já identificados antes, durante e após o uso das soluções em IA. Assim, designers atuarão não somente como projetistas de produtos, sistemas e serviços que integrem a IA, mas analistas dos impactos sociais que se desenrolam em consequência do projeto. Tal papel será, portanto, o responsável por informar o projeto e suas iterações, deflagrando os ajustes necessários para desenvolvimentos éticos.

Importa destacar que não se pretende esgotar o assunto ou preconizar o uso de um único modo de pensar o design e a IA, uma vez que as relações atuais e vindouras entre ambos os campos tendem a provocar novas inquietações e, portanto, novas indagações e reflexões, gerando também novos métodos de pesquisa e de projeto.

O que se busca neste texto é fomentar o debate e explorar os caminhos sobre a prática de projeto nesse cenário, com vistas a ampliar as perspectivas e as possibilidades de ação dos designers e de interação e integração entre os campos do design e da IA.

A INTELIGÊNCIA ARTIFICIAL E SEUS DESAFIOS

A Inteligência Artificial constitui um vasto campo de conhecimento e pesquisa com origem na área da Computação, mas ao longo do tempo também se tornou tema de investigações em diferentes áreas tais como neurociência, linguística, matemática, filosofia entre outras.

Na atualidade, definir o que é a IA constitui um desafio, uma vez que a própria definição de inteligência é fluida (Cozman; Neri, 2021, p. 29). Além disso, existem diferentes definições dependendo das especificidades de cada campo no qual a IA está inserida, sendo a diversidade de pesquisas envolvendo esse tema, um fator que dificulta descrever com precisão o estado da arte da área (Kaufman, 2018; Santaella, 2023).

Dois dos principais autores que são referências no campo, Russel e Norvig (2013) identificam a IA como um campo universal, o qual tem como

objetivo não somente a compreensão da inteligência, mas também a sua construção e destacam que ao longo da história da área foram se desenhando quatro estratégias para o estudo da IA, as quais têm sido seguidas por diferentes pessoas, partindo de métodos diferentes e resultando em estudos e proposições também distintos.

Essas quatro estratégias trazem oito definições do que é a IA divididas em dois espectros: um referente a processos mentais e outro a comportamentos. O Quadro 1 mostra tais definições:

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 um ser humano	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>“AI... está relacionada a um desempenho inteligente de artefatos” (Nilsson, 1998).</p>

QUADRO 1

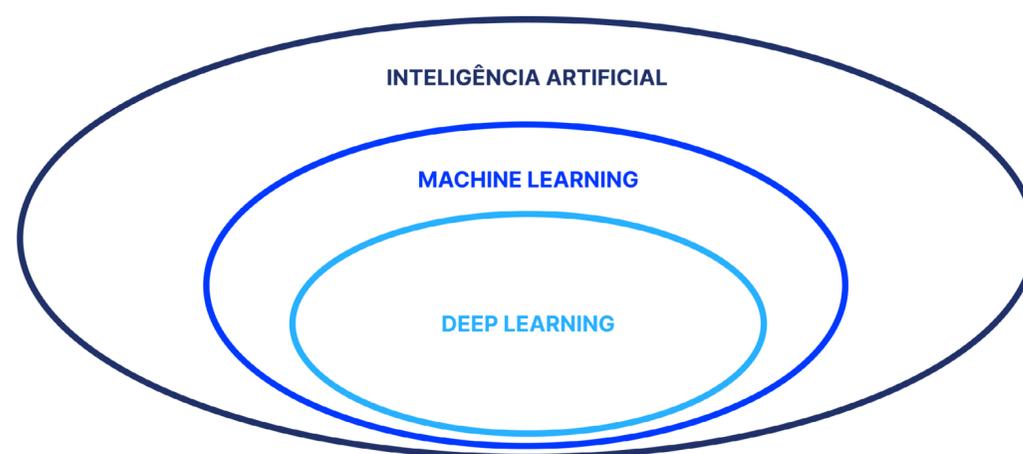
Definições de IA divididas em quatro estratégias de pesquisa e desenvolvimento. Fonte: Russel e Norvig (2013), tradução nossa, quadro elaborado pela autora.

Cada uma dessas estratégias pode favorecer o desenvolvimento de diferentes soluções em termos de produtos, sistemas ou serviços. Independente de qual seja escolhida como caminho para um determinado desenvolvimento, muitas delas já se encontram materializadas em soluções de IA que estão cada vez mais presentes na vida cotidiana das pessoas.

Em termos de estrutura, conforme apresentado na Figura 1, Taulli oferece uma visão dos elementos que compõem a IA da seguinte maneira:

FIGURA 1

Elementos que compõem a IA de acordo com Taulli. Fonte: Taulli (2020, p. 35), adaptado pela autora.



A figura evidencia a IA enquanto campo do conhecimento amplo, dedicado a investigar como os computadores podem aprender a partir do processamento de dados e com o emprego de algoritmos sofisticados.

Dentro da IA existe a subárea de Machine Learning (ML), responsável pelo aprendizado de máquina, que ocorre quando um computador, a partir de técnicas estatísticas sofisticadas, é capaz de aprender sem ser explicitamente programado (*Ibid.*, p. 93).

Ainda dentro do escopo do Machine Learning existe a subárea de Deep Learning a qual, segundo o autor, é um sistema que permite o processamento de enormes quantidades de dados identificando relacionamentos e padrões que os seres humanos poderiam não conseguir identificar. Ainda de acordo com Taulli (2020, p. 98), a palavra “deep” se refere ao número de camadas ocultas na rede neural – modelo de IA que imita

o modo de operar do cérebro humano com diversas camadas de análise – as quais fornecem grande parte do poder de aprendizagem.

Todas essas técnicas compõem atualmente uma variedade de produtos, especialmente os digitais, nos quais a dinâmica de uso envolve diretamente a captação de dados, uma vez que, para que as pessoas possam fazer uso dos benefícios que o produto oferece, elas cedem seus dados e informações de navegação, alimentando os processos de treinamento da IA. Como destacado por Kaufman:

Parte do sucesso da Netflix está em seu sistema de personalização, em que algoritmos analisam as preferências do usuário e de grupos de usuários com preferências semelhantes e, com base nelas, sugerem filmes e séries. Acessamos sistemas inteligentes para programar o itinerário com o Waze, pesquisar no Google e receber do Spotify recomendações de músicas [...] O varejo físico incorpora as “vantagens” do varejo online por meio de sensores que permitem identificar por onde o cliente circulou nas lojas, a trajetória do seu olhar nas prateleiras, por quantas vezes e por quanto tempo. São os

*algoritmos de IA que transformam em informação útil a imensidão de dados gerados pelas movimentações digitais (“rastros digitais”)
(Kaufman, 2018, local. 25).*

A autora destaca que os benefícios entregues por tais soluções são inegáveis, entretanto, uma série de riscos fazem parte dos efeitos colaterais da IA, os quais constituem verdadeiros desafios a serem equacionados tais como “[...] (a) o viés nos processos de decisão automatizados, (b) a invasão da privacidade e as novas formas de controle, e (c) a personalização dos acessos e pesquisas online” (*Ibid.*, local. 25).

Nesse contexto, destacam-se os riscos de manipulação a que podem estar sendo submetidos os cidadãos, especialmente os não produtores de tecnologia situados no Sul Global, um cenário que favorece o que Nick Couldry e Ulises A. Mejias (2019) denominam Colonialismo de Dados, o qual, segundo sintetizado por Cassino:

[...] 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).

Esse processo cria um abismo entre as oportunidades geradas nos âmbitos sociais e tecnológicos para as nações colonizadoras em comparação às nações colonizadas, as quais ficam à margem desses desenvolvimentos. Assim, a colonialidade por dados influencia tais sociedades explorando-as também no cenário digital.

Além do colonialismo de dados, o viés algorítmico surge como um grave efeito colateral a ser equacionado. O surgimento de vieses ocorre quando o modelo de IA é treinado com um grande volume de dados que podem ser tendenciosas, não

representam a realidade ou sofrerem influência da visão de mundo de quem desenvolve a tecnologia (Cozman; Kaufman, 2022; Silva, 2022).

Tais fatores pautam discussões importantes, pois, à medida que a IA avança enquanto tecnologia com potencial para promover benefícios diversos, ela se encontra também inserida em um contexto de mercado que visa o lucro e cujas regras são ditadas pelas grandes empresas de tecnologia do Norte Global, também chamadas Big Techs.

Empresas como Google, Amazon, Meta, entre outras, operam captando, processando e analisando dados em volumes colossais, tendo “[...] não só o poder de extrair, mercantilizar e controlar comportamentos, mas também de produzir novos mercados, por sua capacidade de predição analítica e da modificação de atitudes, práticas e hábitos” (Cassino, 2021, p. 28).

Muitas destas empresas superam em valor de mercado o PIB de diversos países – entre eles o Brasil – e no momento atuam sem qualquer

regulação. De acordo com Daron Acemoglu e Simon Johnson (2023) a Microsoft e a Alphabet/Google assumiram a liderança em desenvolvimentos relacionados à IA. A Microsoft alcançou o feito com estabelecimento do ChatGPT (Generative Pre-trained Transformer) da Open AI, um modelo de linguagem natural pertencente à família das IAs Generativas – capazes de gerar novas informações de maneira original – que utiliza inteligência artificial para formular textos e responder a perguntas variadas feitas pelos usuários.

Já a Alphabet/Google se destacou com o lançamento de uma série de ferramentas baseadas em IA que dão suporte para e-mail, planilhas e elaboração de todo tipo de texto, todos estes produtos que o Google oferece.

Os autores ainda enfatizam que, com base nos fatos da história humana, a distância aberta pelas duas empresas em relação aos concorrentes constitui um importante alerta de que, quando a

distribuição de informação é deixada nas mãos de poucos, tem-se como resultado provável a opressão política e econômica:

O fato de estas empresas estarem tentando ultrapassar umas às outras, na ausência de salvaguardas impostas externamente, deveria nos dar ainda mais motivos de preocupação, dado o potencial da IA causar grandes danos ao emprego, à privacidade e à segurança cibernética (Acemoglu; Johnson, 2023).

Nesse sentido, para Santaella “[...] emerge a necessidade de enfrentamento de dilemas relativos a acesso a dados, viés algorítmico, ética e transparência, e responsabilidade legal por decisões resultantes de IA” (Santaella, 2021, p. 8). Atualmente diversos países têm dedicado esforços para a discussão e regulação da IA, entre eles o Brasil, por meio do Projeto de Lei 2338/23 o qual dispõe sobre o uso da Inteligência Artificial (Brasil, 2023) e tramita no Senado Federal do país. Cabe destacar, contudo, que na visão de especialistas a urgência em aprovar uma regulação pode prejudicar a qualidade do documento e o próprio processo.

Dora Kaufman aponta, em reportagem da Revista FAPESP, ser necessário um tempo de maturação para que o debate seja mais amplo: “a regulamentação prematura pode restringir a inovação e não proteger a sociedade [...] O processo é tão importante quanto o resultado final” (Schmidt, 2023).

Fica evidente então, que parte importante da busca por garantir desenvolvimentos éticos e benéficos envolvendo a IA não diz respeito à velocidade com que se postulam regras ou restrições, mas antes, passa pela necessidade de um debate permanente e amplo para que as propostas de regulação estejam adequadas às particularidades e características das diferentes sociedades, inclusive para fomentar a inovação em IA dentro dos diferentes contextos que as constituem.

ÉTICA E IA: CONTRIBUIÇÕES DA ONU E DA UNESCO PARA A EDUCAÇÃO E A PESQUISA

As preocupações com os desdobramentos da IA e os impactos sociais decorrentes de seu uso em escala massiva – e que ocorrem quando as empresas liberam versões de produtos para o uso do público em geral – fizeram com que a UNESCO se dedicasse a discutir o tema, especialmente considerando a urgência de que os diversos atores envolvidos na criação de soluções baseadas em IA priorizem os aspectos éticos no processo de desenvolvimento destes projetos.

Para discutir tais questões, em 2022 a UNESCO elaborou um documento intitulado “Recomendações sobre a ética da Inteligência Artificial” (UNESCO, 2022). Este documento disponível em diferentes idiomas leva em consideração as várias dimensões nas quais o uso da IA pode levantar questões éticas. Nesse contexto, visando

favorecer ações que possam lidar com esta complexidade, a publicação destaca quatro valores e dez princípios orientadores que podem ser utilizados pelos países como referência na concepção de estruturas ou mecanismos para lidar com desenvolvimentos em IA de maneira ética. Os valores são: Respeito, proteção e promoção dos direitos humanos, das liberdades fundamentais e da dignidade humana; Prosperidade ambiental e ecossistêmica; Garantir diversidade e inclusão; e Viver em sociedades pacíficas, justas e interconectadas. De acordo com a publicação (UNESCO, 2022, p. 18), estes valores devem ser respeitados e promovidos por todos os atores envolvidos no ciclo de vida dos sistemas de IA funcionando também como ideais motivadores para a formulação de medidas nos âmbitos político e jurídico.

Os dez princípios são: Proporcionalidade e não causar dano; Segurança e proteção; Justiça e não discriminação; Sustentabilidade; Direito à privacidade e proteção de dados; Supervisão humana e determinação; Transparência e explicabilidade; Responsabilidade e prestação de contas;

Conscientização e alfabetização; Governança e colaboração adaptáveis e com múltiplas partes interessadas. Ainda de acordo com a publicação a confiabilidade e a integridade desses sistemas são essenciais para garantir que as tecnologias de IA trabalhem para o bem da humanidade.

A confiabilidade, inclusive, é um fator que tem estado presente nas notícias sobre a IA e o seu uso, principalmente levando em consideração a ampliação do lançamento para público em geral de ferramentas baseadas em IA generativa – GenAI – e sua ampla aceitação e uso por parte das pessoas. A IA Generativa compreende uma tecnologia de IA que:

[...] gera conteúdo automaticamente em resposta a solicitações escritas em interfaces conversacionais de linguagem natural. Em vez de simplesmente fazer curadoria páginas da web existentes, aproveitando o conteúdo existente, a GenAI realmente produz novos conteúdos [...] textos escritos em linguagem natural, imagens (incluindo fotografias e pinturas digitais e desenhos animados), vídeos, músicas e código de software. A GenAI é treinada usando dados

coletados de páginas da web, conversas em mídias sociais e outras mídias on-line. Ela gera seu conteúdo analisando estatisticamente as distribuições de palavras, pixels ou outros elementos nos dados que ingeriu e identifica e repete padrões comuns (por exemplo, quais palavras normalmente seguem quais outras palavras). Embora a GenAI possa produzir novos conteúdos, não pode gerar novas ideias ou soluções para os desafios do mundo real, pois, não entende, objetos ou relações sociais do mundo real que sustentam a linguagem (UNESCO, 2023, p. 8, tradução nossa).

As tecnologias específicas por trás do GenAI fazem parte da família de Machine Learning, anteriormente abordada, a qual utiliza algoritmos para melhorar o desempenho a partir do processamento e análise de dados. O tipo de Machine Learning adotado nas GenAI é conhecido como Redes Neurais Artificiais (RNAs), as quais são inspiradas no funcionamento do cérebro humano e sinapses dos neurônios. Enquanto a IA geradora de texto utiliza o Transformador Generativo Pré-treinado (GPT), a geradora de imagens utiliza Redes Generativas Adversariais (GANs). Segundo Santaella:

Os avanços responsáveis pelo sucesso da IA são devidos à aprendizagem de máquina, em especial à técnica de aprendizagem profunda (deep learning) que, por meio das redes neurais artificiais (RNAs), imita até certo ponto o funcionamento das redes neuronais do nosso cérebro [...] A modalidade generativa foi introduzida na década de 1960 em chatbots. Mas foi apenas em 2014, com a introdução de redes adversariais generativas, ou GANs — um tipo de algoritmo de aprendizado de máquina — que a IAG foi capaz de criar imagens, vídeos e áudios. As GANs usam duas redes neurais que competem entre si para realizar previsões mais precisas, colocando uma contra a outra (portanto, "adversarial") de modo a gerar novas instâncias de dados sintéticos que podem passar por dados reais. Elas apresentam uma estrutura cooperativa de jogo de soma zero para aprender. Essa invenção abriu dois caminhos opostos. De um lado, serviu, por exemplo, para melhorar a dublagem de filmes e produzir rico conteúdo educacional. De outro, deu origem a Deep Fakes, isto é, imagens ou vídeos falsificados digitalmente (Santaella, 2023, local. 17-18).

Considerando este cenário desafiador, no ano de 2023 a UNESCO publicou um outro documento intitulado “Orientação para IA generativa em educação e pesquisa” no qual examina as controvérsias e questões éticas, bem como os riscos levantados por todos os sistemas GenAI e suas implicações para a educação. Entre as controvérsias (UNESCO, 2023, p. 14-17) a publicação destaca: a ampliação da pobreza digital, uma vez que o acesso a dados tornou-se essencial para o desenvolvimento econômico dos países e para a criação de oportunidades nesse ambiente, as nações que não possuem tal acesso são deixadas à margem; a ultrapassagem da regulação federal por parte das empresas que acessam dados em âmbito nacional, mas não permitem revisão acadêmica independente; o uso de conteúdo sem consentimento, identificado como um risco pelo fato de as GenAI serem treinadas com dados disponíveis on-line; o uso de modelos que não explicáveis ou totalmente transparentes; o aumento de conteúdos produzidos pelas GenAI poluindo a internet; a falta de entendimento do mundo real,

uma vez que as GenAI trazem conteúdos errados e falsos que podem ser prejudiciais em diferentes âmbitos da sociedade; a redução da diversidade de opiniões, uma vez que as tecnologias são criadas por países em sua maioria do Norte Global podendo marginalizar ainda mais vozes já marginalizadas; e, por fim, a criação de deepfakes – vídeos e imagens falsos que parecem reais – mais perigosos dos que os já vistos na atualidade, pois a tendência é que as IA que geram imagens, vídeos e áudios se tornem cada vez mais acuradas com a evolução destes modelos.

Para regular o uso da IA generativa na educação e na pesquisa, o documento da UNESCO (2023, p. 18-20, tradução nossa) destaca que a abordagem deve ser centrada no ser humano e traz sete passos para essa regulação, a saber: Endossar ou desenvolver os Regulamentos Gerais de Proteção de Dados internacionais ou regionais; Adotar/revisar e financiar estratégias em IA envolvendo o governo como um todo; Solidificar e implementar regulamentos sobre a ética da IA; Ajustar ou aplicar as leis existentes sobre direitos autorais para regular

o conteúdo gerado por IA; Elaborar quadros regulamentares sobre a IA generativa; Desenvolver capacidade para o uso adequado da GenAI em educação e pesquisa; e Refletir sobre as implicações a longo prazo da GenAI para educação e pesquisa.

Os desafios não são poucos diante de uma tecnologia que se desenvolve em uma velocidade na qual muitas vezes governos e instituições não conseguem acompanhar, mas precisarão desenvolver dispositivos regulatórios para tal.

Entendendo a urgência da questão, no dia 26 de outubro de 2023, o secretário geral da ONU António Guterres anunciou a criação de um novo órgão consultivo de alto nível, o qual atuará discutindo os riscos e benefícios associados à governança da inteligência artificial (IA) em âmbito global, o High-level Advisory Body on Artificial Intelligence. O órgão será um ponto de partida para uma:

[...] conversa global, multidisciplinar e multissetorial sobre a governança da IA para que os seus benefícios para a humanidade – toda a

humanidade – sejam maximizados, e os riscos contidos e diminuídos (ONU, 2023, p. 2, tradução nossa).

No início do comunicado de lançamento do órgão consultivo, Guterres afirma ainda ser extraordinário o potencial da IA para solucionar diversos desafios da humanidade, mas que isso depende de as tecnologias de IA serem aproveitadas de forma responsável e também serem tornadas acessíveis a todos, o que inclui os países em desenvolvimento, que são os que mais precisam de suporte e auxílio. E enfatiza:

Do jeito que as coisas estão, a experiência em IA está concentrada em um punhado de empresas e países. Isso poderia aprofundar as desigualdades globais e transformar as divisões digitais em abismos. Os potenciais danos da IA estendem-se a sérias preocupações sobre a desinformação; a consolidação do preconceito e da discriminação; vigilância e invasão de privacidade; fraude e outras violações dos direitos humanos (ONU, 2023, p. 2, tradução nossa).

Ainda no mesmo texto, Guterres afirma que o uso malicioso da IA poderia “minar a confiança nas instituições, enfraquecer a coesão social e ameaçar a própria democracia”.

Tais dificuldades trazem novos contornos para um cenário já considerado complexo por Denis (2012) que destaca ser o design um campo de muitas possibilidades uma vez que a área “[...] tende a se ampliar à medida que o sistema se torna mais complexo e à medida que aumenta, por conseguinte, o número de instâncias de inter-relação entre suas partes” (Denis, 2012, p. 234).

O autor defende que a grande importância do design na contemporaneidade reside na capacidade que a área tem de “[...] construir pontes e forjar relações num mundo cada vez mais esfacelado pela especialização e fragmentação de saberes” (*Ibid.*, p. 234) e enfatiza que a crescente complexidade do cenário atual traz como desafio aos designers repensar velhos conceitos buscando novas respostas ou, ao menos, formular novas perguntas mais precisas e eficazes.

Concordando com Denis, é defendido neste ensaio que o design apresenta condições para formular, de modo transdisciplinar, as perguntas necessárias do ponto de vista dos aspectos éticos e impactos sociais no desenvolvimento de soluções em IA, uma vez que a área parte da lente de projeto, cujo foco sempre é o ser humano.

ETHICS BY DESIGN OU ÉTICA DESDE O DESIGN

A aceleração da IA – e o seu uso cada vez mais frequente pelas pessoas – não só impacta como modifica o contexto social para o qual se projeta, tornando-o cada vez mais complexo e oscilante, sendo um desafio tanto para os profissionais envolvidos no desenvolvimento de produtos, sistemas e serviços (entre eles o designer), quanto para o público usuário.

Para Cozman e Neri (2021, p. 29) a relação IA e sociedade configura mais um eixo essencial no estudo de Inteligências Artificiais. Nesse cenário,

deve-se ter em conta que as pessoas usuárias dessas produções podem não ter os recursos necessários ou mesmo o letramento digital suficiente para compreender as dimensões daquilo que estão acessando, bem como os benefícios e os riscos associados de modo que possam tomar decisões informadas de interação a partir de uma perspectiva crítica de uso. Para Xavier o letramento digital é:

[...] o domínio pelo indivíduo de funções e ações necessárias à utilização eficiente e rápida de equipamentos dotados de tecnologia digital, tais como computadores pessoais, telefones celulares, caixas-eletrônicos de banco, tocadores e gravadores digitais, manuseio de filmadoras e afins (Xavier, 2011, p. 6).

Esse letramento, inclusive, será imperativo para um uso consciente das produções digitais atuais e as que ainda virão, englobando cada vez mais técnicas de IA. De acordo com Kaufman, o uso da tecnologia implica externalidades positivas e negativas:

São múltiplas as externalidades negativas dos sistemas atuais de IA, éticas e sociais. O desafio é como mitigá-las preservando as externalidades positivas intrínsecas aos modelos de negócio baseados em dados (data-driven models). Esse cenário recomenda manter o foco na busca por soluções de curto-médio prazo, deixando o longo prazo para a esfera da ficção científica (Kaufman, 2021a, p. 163).

Entendendo as vulnerabilidades humanas frente aos avanços das tecnologias inteligentes e suas externalidades é que surgiu a abordagem denominada Ethics by Design – ou Ética desde o Design. Segundo Brey e Dainow (2023), trata-se de:

[...] uma abordagem para incluir de forma sistemática e abrangente considerações éticas no processo de design e desenvolvimento de novos sistemas e dispositivos tecnológicos. Embora a abordagem possa ser aplicada a qualquer tecnologia, historicamente o seu foco tem sido na concepção de sistemas de IA (Brey; Dainow, 2023, tradução nossa).

Os autores destacam que, os estudos neste campo, surgiram a partir de pesquisas europeias formuladas na área de ciência da computação envolvendo ética e inovação, notadamente da metade para o final da década de 2010, sendo referências os estudos de Dignum et al. (2018) e d’Aquin et al. (2018).

Tal abordagem ganhou mais impulso na Europa por meio de uma resolução do Parlamento Europeu do ano de 2019, a qual identificava a necessidade da existência de uma política abrangente de IA e que os modelos implantados por meio dessa tecnologia deveriam prever a ética desde a sua concepção. Brey e Dainow (2023) identificam ainda, que embora importantes estudos tenham sido desenvolvidos trazendo princípios, definições e requisitos, nenhum apresentou uma abordagem completa que pudesse ser utilizada em processos de design reais, por esse motivo os autores assumiram a tarefa de desenvolver uma abordagem completa:

No período 2019-2021 [...] numa colaboração entre dois projetos financiados pela UE com foco na ética da IA: os projetos SHERPA e SIENNA [...] Assumimos a responsabilidade pela versão final da abordagem SIENNA/SHERPA Ethics by Design, uma vez que somos os autores do relatório final sobre Ethics by Design e também trabalhamos com a Comissão Europeia para desenvolver o documento de orientação sobre ética desde a concepção para financiamento europeu (Brey; Dainow, 2023, tradução nossa).

A abordagem criada pelos autores denomina-se “Ethics by Design for AI” ou “EbD-AI”, sendo o equivalente em português em tradução livre “Ética desde o Design para IA” ou “EdD-IA”. Os autores entendem que a abordagem proposta baseia-se na “[...] convicção de que a tecnologia não é neutra e que os valores podem, até certo ponto, ser incorporados no processo de design” (*Ibid.*, tradução nossa). Eles destacam que a premissa-chave é que as escolhas de design não são moralmente neutras, podendo ter impactos éticos

significativos. Mesmo entendendo que muitas vezes os impactos podem se originar do contexto ou de como as tecnologias são utilizadas, o design também tem seu papel, pois dependendo de como uma função é projetada, por exemplo, se ela viola a privacidade – o que independe de onde, quando e como ela é utilizada – o prejuízo já está posto.

A partir desse cenário, para operacionalizar projetos de IA com a ética sendo considerada desde o design a proposta de Brey e Dainow (2023) é fornecer tarefas tangíveis específicas aos engenheiros e profissionais envolvidos no processo.

Cabe destacar, contudo, que não há uma especificação de como tais tarefas devem ser realizadas, uma vez que tais fatores dependem de qual é a aplicação sendo desenvolvida, bem como da organização que está coordenando o projeto. Os autores enfatizam ainda que sua proposta funciona bem para questões éticas rotineiras, entretanto, entendendo que podem ocorrer questões éticas especiais eles sugerem uma abordagem reflexiva

e deliberativa recomendando, inclusive, que um profissional especialista em ética componha a equipe de projeto para avaliar riscos, identificar e refletir sobre tais questões junto à equipe.

O foco da abordagem apresentada por Brey e Dainow (2023), segundo os autores, está em instanciar os valores morais fundamentais necessários para que um projeto atenda aos padrões éticos indo da privacidade à justiça destacando que, assim como alguns termos éticos, o significado exato destes valores dentro de uma aplicação irá variar, ou mesmo ser discutível e caberá aos envolvidos adotarem uma postura coerente sendo que a EdD-IA não identifica como isso deve ser feito, mas sim o que deve ser feito. Em seu modelo:

Os requisitos aplicar-se-ão não só ao próprio sistema de IA, mas também a alguns dos processos e ferramentas utilizados no seu desenvolvimento. Primeiro, os valores e princípios morais fundamentais de natureza geral são convertidos em requisitos éticos para o sistema de IA específico a ser desenvolvido, depois é determinado como construir o sistema de uma

forma que os instancie. Assim, os requisitos éticos são traduzidos em tarefas concretas, objetivos, ferramentas, funções, restrições e assim por diante (Brey; Dainow, 2023, tradução nossa).

Para identificar quais seriam os valores morais fundamentais de natureza geral a serem convertidos em requisitos éticos, os autores analisaram diferentes publicações elaboradas por inúmeras organizações internacionais e governos nacionais – num total de 16 documentos – para os quais propuseram diretrizes éticas para IA.

Como resultado, identificaram um consenso global sobre quais valores morais são especificados como fundamentais no desenvolvimento e uso da IA, a saber: liberdade, privacidade, justiça, transparência, responsabilidade e bem-estar (dos indivíduos, da sociedade e do meio ambiente).

Deste ponto em diante, os autores endossam esses seis valores dentro da abordagem EdD-IA e identificam os principais requisitos de projeto que

deles decorrem. As informações especificadas sobre cada valor e seus requisitos estão presentes no Quadro 2, a seguir:

QUADRO 2

Valores da Ética desde o Design para a IA, por Brey e Dainow (2023). Fonte: Brey e Dainow, 2023, tradução nossa, quadro elaborado pela autora.

Agência humana

A agência humana abrange três valores que são reconhecidos como direitos humanos: liberdade, autonomia e dignidade. O respeito pela agência humana envolve o respeito pelo direito do ser humano de ter os seus próprios pensamentos, tomar as suas próprias decisões e realizar as suas próprias ações.

Requisitos de concepção

1. Um sistema de IA não deve ser concebido ou utilizado de uma forma que prive as pessoas da capacidade de tomar decisões que deveriam ser capazes de tomar por si próprias.
2. Uma IA não deve ser concebida ou utilizada de uma forma que resulte na redução das liberdades humanas básicas, incluindo a liberdade de circulação, reunião, expressão e informação.
3. Os sistemas de IA não devem ser concebidos ou utilizados para subordinar, coagir, enganar, manipular, objetificar ou desumanizar pessoas
4. Os sistemas de IA não devem ser concebidos ou utilizados para criar dependência do sistema ou dos serviços que presta.
5. As aplicações de IA devem ser concebidas para proporcionar aos operadores do sistema e (tanto quanto possível) aos utilizadores finais a capacidade de controlar, dirigir e intervir nas operações do sistema
6. Os sistemas de IA nunca devem tomar a decisão final sobre questões importantes de natureza pessoal, moral ou política. Eles podem recomendar, mas a decisão final deve sempre ser tomada por um ser humano.

Privacidade e governança de dados

Este princípio afirma o direito à privacidade de todos os seres humanos e afirma a importância da governação de dados, incluindo medidas para apoiar a qualidade e precisão dos dados, o acesso aos dados e outros direitos de dados, como a propriedade. Os requisitos aplicam-se tanto aos dados utilizados durante a concepção do sistema de IA como aos dados por ele gerados quando em utilização.

Requisitos de concepção

1. Os direitos dos titulares dos dados devem ser respeitados durante o tratamento de dados pessoais.
2. Para garantir a responsabilização, quando os dados pessoais são processados por uma IA, devem existir formas de demonstrar como esta garante a legalidade, a justiça e a transparência desse processamento de dados.
3. Devem ser implementadas medidas para salvaguardar os direitos dos titulares dos dados através de medidas técnicas, como a anonimização, bem como através de medidas organizacionais, como sistemas de controlo de acesso.
4. Sempre que relevante, os sistemas de IA devem apoiar o direito de um indivíduo retirar o consentimento para a utilização dos seus dados pessoais.
5. Os dados devem ser adquiridos, armazenados e processados de uma forma que possa ser auditada por seres humanos.

Justiça

Este valor implica que as pessoas devem ter direitos e oportunidades iguais e não devem ser favorecidas ou desfavorecidas indevidamente. A justiça implica a ausência de qualquer forma de discriminação, bem como o apoio à diversidade e à inclusão.

Requisitos de concepção

1. Os sistemas de IA devem evitar distorções algorítmicas, incluindo distorções nos dados de entrada, na modelização e na concessão de algoritmos.
2. Os sistemas de IA devem, na medida do possível e relevante, ser universalmente acessíveis e oferecer as mesmas funcionalidades e benefícios aos utilizadores finais, independentemente das suas diferentes capacidades, crenças, preferências ou interesses.
3. Os sistemas de IA devem, na medida do possível, ser concebidos para evitar impactos sociais negativos nos grupos sociais, especialmente nos grupos sociais protegidos.

QUADRO 2

Valores da Ética desde o Design para a IA, por Brey e Dainow (2023). Fonte: Brey e Dainow, 2023, tradução nossa, quadro elaborado pela autora.

Bem-estar individual, social e ambiental

De acordo com este princípio, os sistemas de IA devem contribuir, e não prejudicar, o bem-estar individual, a qualidade e o funcionamento da sociedade e a qualidade do ambiente.

Requisitos de concepção

1. Os sistemas de IA devem ser seguros de utilizar e não devem ter propensão para prejudicar ou reduzir significativamente a saúde e o bem-estar físico ou psicológico de quaisquer partes interessadas (utilizadores, clientes, titulares de dados e outras partes afetadas).
2. O desenvolvimento da IA deve ter em conta os princípios da sustentabilidade ambiental, tanto no que diz respeito ao próprio sistema como à cadeia de abastecimento à qual está ligado.
3. Os sistemas de IA não devem ter um impacto negativo na qualidade da comunicação, da interação social, da informação, das relações sociais ou dos processos democráticos; por exemplo, amplificando notícias falsas ou segregando pessoas em bolhas de filtros.

Transparência

Isto refere-se à ideia de que a finalidade, os inputs e as operações das aplicações de IA devem ser conhecidos e compreensíveis para as suas partes interessadas. Isto é para que possam compreender como e com que finalidade estes sistemas funcionam e como são tomadas as suas decisões.

Requisitos de concepção

1. Deve ficar claro para os usuários que eles estão interagindo com um sistema de IA – especialmente para sistemas que simulam a comunicação humana, como chatbots.
2. A finalidade, as capacidades, as limitações, os benefícios e os riscos do sistema de IA e as decisões que toma devem ser comunicadas abertamente a todas as partes interessadas.
3. Os sistemas de IA devem ser construídos de modo que as pessoas possam auditar, questionar, contestar ou procurar alterar as suas atividades. Isto inclui processos organizacionais através dos quais os operadores podem receber e avaliar pedidos de terceiros.
4. Ao construir um sistema de IA, deve-se considerar quais medidas permitirão a rastreabilidade do sistema de IA durante todo o seu ciclo de vida, desde a concepção inicial até a avaliação pós-implantação.
5. Sempre que relevante, as decisões sobre IA devem ser explicáveis aos utilizadores. Sempre que possível, isto deve incluir as razões pelas quais o sistema tomou uma decisão específica. Reconhecemos que isso pode não ser possível com alguns sistemas. No entanto, o sistema (ou aqueles que o implementam) deve sempre ter um mecanismo para explicar qual foi a decisão e que dados foram utilizados para tomar essa decisão. A explicabilidade é especialmente importante para sistemas que tomam decisões ou executam ações pelas quais a responsabilização pode ser necessária, tais como decisões que podem causar danos ou restringir os direitos de um indivíduo.
6. Os processos de desenvolvimento de IA sempre envolvem a tomada de decisões sobre questões éticas, como como remover preconceitos de um conjunto de dados. A transparência exige que os processos e ferramentas de desenvolvimento registem estas decisões éticas de design para que seja possível compreender como as obrigações éticas foram cumpridas. Estas informações podem ser necessárias para auditorias, para contestar decisões tomadas pelo sistema ou para corrigir quaisquer questões éticas que surjam após a implantação.

Responsabilidade e supervisão

A responsabilização pelas aplicações de IA significa que os intervenientes envolvidos no seu desenvolvimento e operação assumem a responsabilidade pela forma como funcionam e pelas consequências daí resultantes. Supervisão humana significa que os humanos são capazes de compreender, supervisionar e controlar o design e a operação dos seus sistemas. A supervisão é uma condição para a responsabilização, uma vez que os intervenientes precisam dela para ter a informação e a influência necessárias para a responsabilização.

Requisitos de concepção

1. Os sistemas de IA devem permitir a supervisão humana relativamente aos seus ciclos de decisão e funcionamento, a menos que possam ser apresentadas razões convincentes que expliquem por que razão a supervisão não é necessária.
2. O processo de implantação de um sistema de IA deve incluir a avaliação de riscos. Os procedimentos de mitigação após a implantação devem estar em vigor a partir do momento em que o sistema é implantado.
3. Os sistemas de IA devem ser auditáveis por terceiros independentes. Os procedimentos e ferramentas disponíveis no âmbito da abordagem XAI [18] apoiam as melhores práticas neste domínio. Isto não se limita à auditoria das decisões do próprio sistema, mas também dos procedimentos e ferramentas utilizadas durante o processo de desenvolvimento. Sempre que for relevante e prático, o sistema deve gerar registos acessíveis a humanos dos processos internos do sistema de IA.



FIGURA 2
Etapas sobre como aplicar a estrutura de Ética desde o Design proposta por Brey e Dainow (2023).

Os requisitos de projeto baseados nos seis valores são então traduzidos em procedimentos e ações adicionais em diferentes estágios do processo de projeto.

Brey e Dainow (2023) oferecem um esboço de como a EdD-IA é usada no desenvolvimento de sistemas de IA – não se trata de uma descrição de como aplicar EdD-IA, mas sim de um resumo das principais etapas do processo de aplicação da EdD-IA a qualquer desenvolvimento. Estas etapas estão presentes na Figura 2:

Finalmente, para a implementação da EdD-IA, os autores oferecem um modelo genérico de desenvolvimento de sistemas (Figura 3) que foi intencionalmente concebido para ser mapeado em qualquer metodologia de design e desenvolvimento:

O modelo enquadra o desenvolvimento do sistema em termos de seis “fases” genéricas. Embora sejam, inevitavelmente, descritos numa sequência linear de passos discretos, devem ser entendidos como classes de operação que podem ser misturadas, reorganizadas e analisadas conforme apropriado para se adaptarem aos processos de desenvolvimento que serão realmente utilizados na construção de um sistema de IA (Brey; Dainow, 2023, tradução nossa).

O modelo de seis passos está presente na Figura 3, a seguir:

As contribuições de Brey e Dainow (2023) por meio do seu modelo de Ética desde o Design para IA colaboram para tornar acionáveis as questões de ordem ética que surgem – ou podem surgir – como consequência da interação humano-IA. Prever como inserir valores morais tornando-os requisitos de projeto é o grande ganho dessa abordagem que pode colaborar de modo significativo para produções mais éticas.

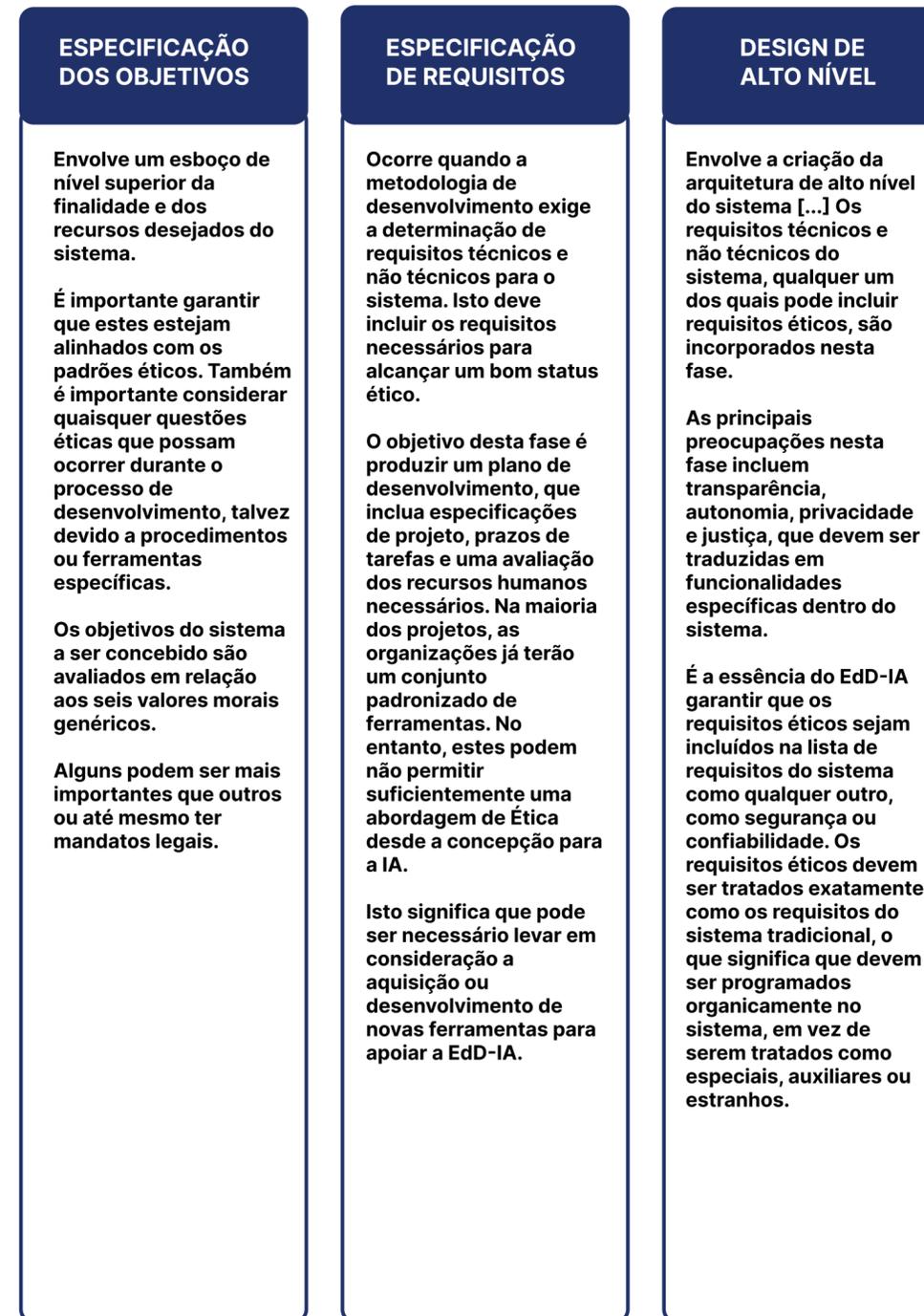


FIGURA 3

Síntese do modelo genérico de desenvolvimento de sistemas para implementação da Ética desde o Design para IA, proposto por Brey e Dainow (2023). Fonte: Brey e Dainow, 2023, tradução nossa, imagem elaborada pela autora.

FIGURA 3

Síntese do modelo genérico de desenvolvimento de sistemas para implementação da Ética desde o Design para IA, proposto por Brey e Dainow (2023). Fonte: Brey e Dainow, 2023, tradução nossa, imagem elaborada pela autora.



PARA ALÉM DO PROJETO: O PAPEL DO DESIGNER NOS DESENVOLVIMENTOS ÉTICOS QUE INTEGREM A IA

Para qualquer desenvolvimento que envolva o design, deve-se considerar que há um fator primordial que pauta o desenrolar do projeto: a compreensão do contexto.

Do ponto de vista do design, o contexto – não somente o de uso, mas também o contexto geral – é um dos aspectos orientadores dos rumos dos projetos na área, muitas vezes definindo as tomadas de decisão nos desenvolvimentos (Bonsiepe, 2011; Moraes, 2005; Denis, 2008; Forty, 2013; Schneider, 2010). Também é possível afirmar que analisar o contexto é uma ação essencial na gestão do design no seu nível mais avançado, o

estratégico, no qual ocorre a gestão do conhecimento e de redes tornando o design um elemento transformador da realidade (Best, 2012; Mozota; Klöpsch; Costa, 2011; Martins; Merino, 2011).

Isso posto, cabe destacar que o contexto complexo atual tem sido marcado pelos desenvolvimentos em IA e também pelos impactos sociais e éticos destas soluções, as quais originam novas demandas de projeto, inclusive – e principalmente – de ordem ética.

No design, pensar projetos de maneira ética implica, portanto, uma etapa anterior que envolve tratar eticamente o contexto para o qual se projeta. Isto não significa que o contexto modificará os valores morais a serem incorporados, mas sim que ele traz particularidades e subjetividades que precisarão ser consideradas para inclusão de maneira ética no projeto.

Como exemplo, pode-se considerar a realidade do Norte Global que difere substancialmente da realidade do Sul Global – região que contempla

nações marcadas pelo colonialismo, neocolonialismo e que apresentam grandes desigualdades. Nesse sentido, problemas específicos que atingem o sul, com capacidade de originar graves questões de ordem ética, podem nem existir no norte do globo e vice-versa.

Tem-se então, que em projetos de design, além dos requisitos de desenvolvimento e processos pautados na ética como proposto pela abordagem EdD-IA, um fator precisará ser constantemente acompanhado, monitorando, analisado e revisto para informar o projeto: o contexto.

Na era da IA os contextos para os quais se projeta são múltiplos e se apresentam variáveis e fluidos. Isto implica que os projetos de design envolvendo a IA, devem ser pensados com flexibilidade o suficiente para suportar iterações cada vez mais velozes. Desse modo, os designers podem constantemente visitar o contexto altamente variável, detectar os impactos sociais dessas produções, analisar criticamente o cenário e iterar rapidamente respondendo às adaptações necessárias.

Essa iteração é parte do processo de design, ou seja, a solução projetada está constantemente sendo revisitada e redesenhada visando a melhoria contínua. Também é verdade que a dinâmica das relações estabelecidas entre as pessoas, os artefatos e suas interfaces estão em constante mudança e isto é um fato que fundamenta a própria iteração no projeto (Flusser, 2007; Jordan, 2005; Norman, 2008). Isso posto, cabe enfatizar que na era da IA essa dinâmica se dá em um contexto no qual a tendência é que as mudanças sejam cada vez mais aceleradas também pelos impactos sociais da própria IA sendo acrescida de efeitos nas dimensões psicológicas, físicas e emocionais das pessoas – o que forçará novas mudanças no contexto de referência dos projetos.

Nessa perspectiva, entende-se ser essencial para uma prática e uma produção éticas, que o desenvolvimento de análises aprofundadas e constantes dos contextos nos quais se dão a interação entre pessoas e IA – considerando os impactos sociais destas produções – faça parte do processo de design, não somente com o olhar

focado no uso, mas antes, com um olhar preditivo dos diferentes contextos gerais para os quais se produz, incluindo aspectos políticos, econômicos e sociais que podem – e deverão – afetar diretamente também o contexto específico das pessoas usuárias.

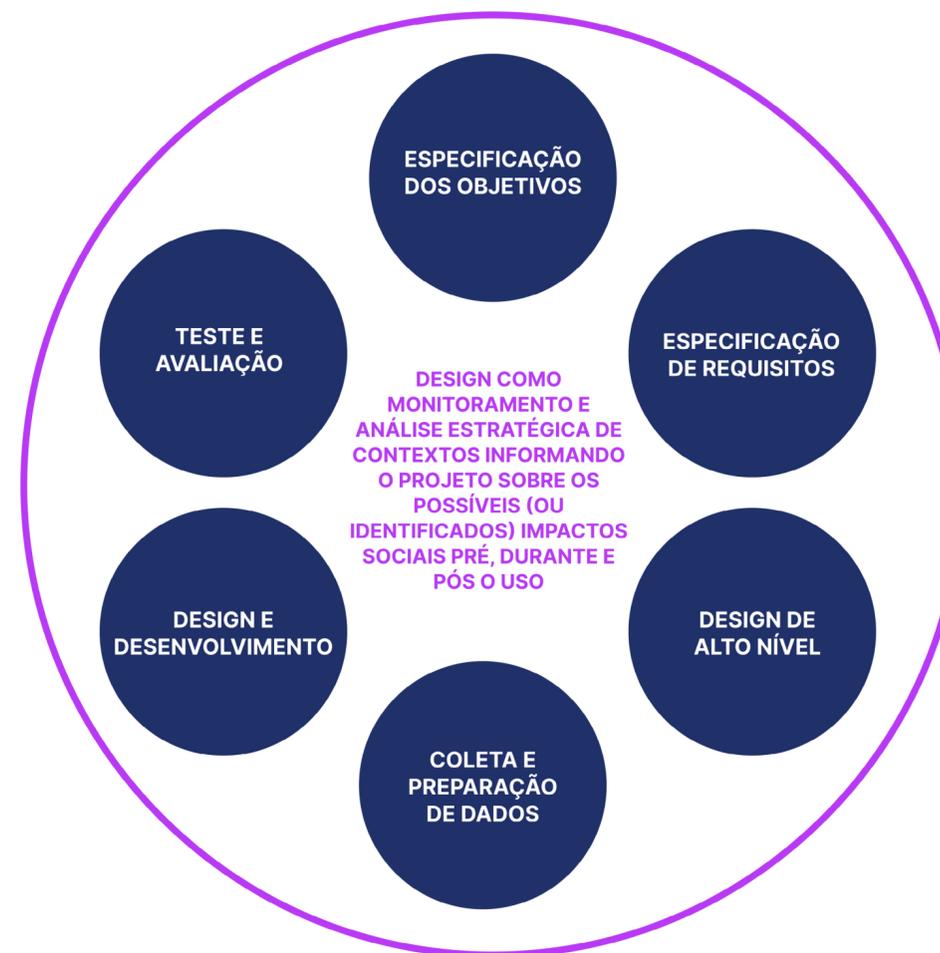
Dessa maneira, entendendo a complexidade do cenário que se apresenta – bem como a interdependência acelerada entre design e contexto – e considerando os impactos sociais da IA, acredita-se que o campo do design pode, para além do projeto, contribuir em um escopo mais amplo: o de monitoramento, análise e compreensão aprofundada dos diferentes contextos e suas particularidades para, de maneira estratégica e preditiva, informar o projeto.

O caráter interdisciplinar da área faz com que ela apresente recursos para atuar não somente no projeto em si incorporando aspectos éticos, mas também como uma força de caráter vigilante, analítico e preditivo, podendo identificar os contextos complexos para os quais se projeta e os

impactos sociais destas produções, atuando de modo transdisciplinar com diferentes áreas como: Computação, Sociologia, Antropologia, Psicologia, Educação, Direito entre outras.

Essa atuação do design como uma força estratégica preditiva dos impactos sociais possíveis ou identificados, informando o projeto em cada etapa, é defendida neste ensaio e pode ser sintetizada como uma contribuição holística que perpassa o projeto em si e o próprio processo de projeto em EdD-IA proposto por Brey e Dainow (2023), como apresentado na Figura 4:

A Figura 4 registra a visão proposta por esta autora englobando a EdD-IA e defendendo que o papel do design nos projetos éticos para IA será – para além do projeto em si – também muito voltado à análise e interpretação dos diferentes contextos para os quais se produz considerando os impactos sociais pré-uso, durante o uso e após o uso.



- Etapas do modelo genérico de desenvolvimento de sistemas EdD - IA de Brey e Dainow (2023).
- Design contribuindo para além do projeto: como monitoramento e análise estratégica de contextos para informar o projeto.

FIGURA 4

Proposta de atuação do design enquanto análise estratégica de contextos para informar o projeto considerando as etapas propostas por Brey e Dainow, 2023. Fonte: A autora.

Acredita-se, portanto, que embora conhecer o contexto seja parte integrante do processo de design, na era da IA ela se dará de modo mais complexo, em cenários velozmente mutantes e provavelmente demandará designers dedicados somente a esta função estratégica de análise e interpretação para que se possa prever os impactos sociais possíveis, ou mesmo compreender os impactos já identificados do projeto com IA na sociedade. Reside aí uma tarefa desafiadora e que, ao mesmo tempo, exige significativo comprometimento para o campo do design.

Versando sobre as incertezas da prática do design para a sustentabilidade, Rafael Cardoso Denis traz uma reflexão que cabe também para o presente momento pautado por incertezas referentes ao pensamento do design frente aos desafios da IA:

Como prever um contexto de uso que não existe ainda? Como pensar nas necessidades de um usuário que não se conhece? Como planejar aquilo que, por definição, não pode ser antevisto? Parece impossível! Dá trabalho, com certeza. A necessidade de projetar para o

pós-uso abre um imenso desafio que promete revolucionar o pensamento em design (Denis, 2012, p. 161).

O “pós-uso” no cenário da IA significará lidar especialmente com os impactos decorrentes da interação com agentes inteligentes, mas também as fases de “pré-uso” – que implica a tomada de decisão e opção pelo uso daquela tecnologia – e “durante o uso” trarão seus próprios desafios e dilemas que não serão só de projeto, mas de pensamento sobre o próprio design enquanto campo e sobre os designers frente ao seu papel nessa nova realidade.

CONSIDERAÇÕES FINAIS

De produtos a serviços a IA traz, e trará ainda mais no futuro, avanços em diferentes setores tais como saúde, educação, comunicação, segurança, entre outros.

Na esteira desse avanço, surgem também diferentes implicações de ordem ética sobre como essas tecnologias têm sido aplicadas e com quais finalidades, especialmente no cenário atual no qual poucas empresas privadas detêm grandes poderes de ordem econômica, política, midiática, de desenvolvimento e, como consequência, de ações que impactam a sociedade envolvendo as IAs.

Atentos a tais questões e aos impactos sociais da IA, diferentes órgãos, governos e instituições têm buscado caminhos para regular produções visando salvaguardar os direitos fundamentais humanos sem, contudo, inibir o desenvolvimento da tecnologia. E ao tratar do projeto dessas soluções entra-se no campo de ação do design.

As discussões apresentadas neste capítulo versam sobre como inserir valores morais dentro dos projetos que envolvem a IA de modo a conseguir acioná-los enquanto requisitos de projeto em um processo estruturado, visando alcançar desenvolvimentos mais éticos em IA.

Nesse sentido, a abordagem “Ética desde o Design para IA” proposta por Brey e Dainow (2023) traz importantes contribuições para a materialização de um processo de desenvolvimento orientado pela ética.

Deve-se, contudo, ter em conta que na aplicação do modelo proposto, a atenção aos diferentes contextos, às características e particularidades das diferentes sociedades é imperativa para que se possa traçar de maneira adequada os rumos do projeto mantendo o foco na ética.

A análise do contexto há tempos faz parte do processo de design, entretanto, na era da IA o contexto muda rapidamente e os impactos sociais de projetos que integrem a IA podem se tornar imprevisíveis, o que faz com que a análise constante dos contextos seja primordial e estratégica para informar as iterações nos projetos.

A partir da abordagem EdD-IA proposta por Brey e Dainow (2023), este capítulo defende que o papel

do designer na era da IA será ampliado para além do de projetista, sendo o designer também um analista estratégico de contextos com capacidade preditiva e um elemento integrador de diversas disciplinas no projeto.

Essa atuação permitirá informar de maneira contínua a iteração no processo de design, de modo que a equipe envolvida possa se antecipar aos impactos sociais das soluções projetadas, antevendo riscos, mitigando-os ou eliminando-os de maneira mais célere, o que, no limite, potencializa os benefícios almejados de projetos para a sociedade.

Acredita-se, por fim, que na era da IA, fazer design implicará conhecer, analisar e monitorar cada vez mais e melhor a alteração do contexto e isso será tão essencial quanto conhecer as motivações, as limitações, os desejos e as necessidades das pessoas para quem se projeta, objetivo primeiro do design.

REFERÊNCIAS

ACEMOGLU, Daron; JOHNSON, Simon. Big tech is bad. Big A.I. will be worse. **The New York Times**. 9 jun. 2023. Disponível em: <https://www.nytimes.com/2023/06/09/opinion/ai-big-tech-microsoft-google-duopoly.html>. Acesso em: 03 out. 2023.

BEST, Kathryn. **Fundamentos de gestão do design**. Porto Alegre: Bookman, 2012.

BRASIL. **Projeto de Lei nº 2338, de 2023**. Dispõe sobre o uso da inteligência artificial. Brasília, DF: Plenário do Senado Federal (Secretaria Legislativa do Senado Federal), [2023] Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233#tramitacao_10494842. Acesso em: 10 jul. 2023.

BOECHAT, C. Inteligência Artificial, empatia e inclusão: um problema de design. **Ergodesign & HCI**, n especial, v. 7, ano 7 (2019). Disponível em: <http://periodicos.puc-rio.br/index.php/revistaergodesign-hci/article/view/1306/735>. Acesso em: 8 abr.2023.

BONSIEPE, Gui. **Design, cultura e sociedade**. São Paulo: Blucher, 2011.

BREY, Philip; DAINOW, Brandt. Ethics by design for artificial intelligence. **AI Ethics**. Springer Link, 2023. DOI: <https://doi.org/10.1007/s43681-023-00330-4>. Disponível em: <https://link.springer.com/article/10.1007/s43681-023-00330-4>. Acesso em: 10 out. 2023.

CASSINO, João Francisco. 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, 2021.

CEREJO, Joana; CARVALHAIS, Miguel. **Design after the rise of AI-driven services: learning from literature review**. DIGICOM 3rd International Conference on Digital Design & Communication. 2019. Disponível em: https://www.researchgate.net/publication/337631988_Design_After_the_Rise_of_AI-Driven_Services_Learning_from_Literature_Review. Acesso em: 8 dez. 2023.

COULDRY, Nick; MEJIAS, Ulises A. **The costs of connection: how data is colonizing human life and appropriating it for capitalism**. Califórnia: Stanford University Press, 2019.

COZMAN, Fábio G.; NERI, Hugo. 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, 2021. Disponível em: <https://www.livrosabertos.abcd.usp.br/portaldelivrosUSP/catalog/book/650>. Acesso em: 10 set.2023.

COZMAN, Fábio G.; KAUFMAN, Dora. Viés no aprendizado de máquina em sistemas de inteligência artificial: a diversidade de origens e os caminhos de mitigação. **Revista USP**, n. 135, p. 195-210, out./nov./dez. 2022. DOI: <https://doi.org/10.11606/issn.2316-9036.i135p195-210>. Acesso em: 10 set.2023.

MORAES, Dijon de. **Análise do design brasileiro**: entre mimese e mestiçagem. São Paulo: Blucher, 2005.

DENIS, Rafael Cardoso. **Uma introdução à história do design**. 3. ed. São Paulo: Blucher, 2008.

_____. **Design para um mundo complexo**. São Paulo: Cosac Naify, 2012.

FLUSSER, Vilém. **O mundo codificado**: por uma filosofia do design e da comunicação. São Paulo: Cosac Naify, 2007.

FORTY, Adrian. **Objetos de desejo**: design e sociedade desde 1750. São Paulo: Cosac Naify, 2013.

JORDAN, Patrick W. **Designing pleasurable products**: an introduction to the new human factors. Filadélfia: Taylor & Francis, 2005.

KAUFMAN, Dora. **A inteligência artificial irá suplantar a inteligência humana?** Barueri: Estação das Letras e Cores, 2018. E-Book.

_____. Resenha do livro "Ethics of artificial intelligence", de Matthew Liao. **Teccogs**. n. 23, jan./jun. 2021a. DOI: <https://doi.org/10.23925/1984-3585.2021i23p157-163>. Disponível em: <https://revistas.pucsp.br/index.php/teccogs/article/view/55974>. Acesso em: 10 out. 2023.

MARTINS, Roseane Fonseca de Freitas; MERINO, Eugênio Andrés Díaz. **A gestão do design como estratégia organizacional**. Londrina: Eduel; Rio de Janeiro: Rio Books, 2011.

MOZOTA, Brigitte Borja de; KLÖPSCH, Cássia; COSTA, Filipe Campelo Xavier da. **Gestão do design**. Porto Alegre: Bookman, 2011.

NORMAN, Donald A. **Design emocional**: por que adoramos (ou detestamos) os objetos do dia-a-dia. Rio de Janeiro: Rocco, 2008.

ONU. **Press remarks launching high-level advisory body on artificial intelligence**. The Secretary-General. 2023. Disponível em: https://www.un.org/sites/un2.un.org/files/sgs_remarks_announcing_high-level_advisory_body_artificial_intelligence_26_october_2023.pdf. Acesso em: 27 out. 2023.

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

SANTAELLA, Lucia. **Inteligência artificial e cultura**: oportunidades e desafios para o sul global. [S.l.], UNESCO. CETIC.BR/NIC.BR. 2021. Disponível em: <https://cetic.br/pt/publicacao/inteligencia-artificial-e-cultura-oportunidades-e-desafios-para-o-sul-global/>. Acesso em: 03 out. 2023.

_____. **Há como deter a invasão do ChatGPT?** São Paulo: Estação das Letras e Cores, 2023. E-book.

SCHNEIDER, Beat. **Design – uma introdução**: o design no contexto social, cultural e econômico. São Paulo: Editora Blucher, 2010.

SULEYMAN, Mustafa; BHASKAR, Michael. **A próxima onda**: inteligência artificial, poder e o maior dilema do século XXI. Rio de Janeiro: Record, 2023. E-book.

SCHMIDT, Sarah. 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**. ed. 331, set. 2023. Disponível em: <https://revistapesquisa.fapesp.br/os-desafios-para-regulamentar-o-uso-da-inteligencia-artificial/>. Acesso em: 10 set. 2023.

SILVA, Tarcízio. **Racismo algorítmico: inteligência artificial e discriminação nas redes digitais**. São Paulo: Edições Sesc São Paulo, 2022.

TAULLI, Tom. **Introdução à Inteligência Artificial: uma abordagem não técnica**. São Paulo: Novatec, 2020.

UNESCO. Organização das Nações Unidas para a Educação, a Ciência e a Cultura. **Recomendações sobre a ética da inteligência artificial**. 2022. Disponível em https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. Acesso em: 08 jun. 2022.

_____. Organização das Nações Unidas para a Educação, a Ciência e a Cultura. **Orientação para IA generativa em educação e pesquisa. Guidance for generative AI in education and research**. 2023. Disponível em <https://unesdoc.unesco.org/ark:/48223/pf0000386693>. Acesso em: 10 out. 2023.

XAVIER, Antônio Carlos. Letramento digital: impactos das tecnologias na aprendizagem da Geração Y. **Calidoscópio**. v. 9, n. 1, p. 3-14, jan./abr. 2011. DOI: 10.4013/cld.2011.91.01. Disponível em: <https://revistas.unisinos.br/index.php/calidoscopio/article/view/748>. Acesso em: 11 out. 2023.



CAROLINA VAITIEKUNAS PIZARRO

Professora Assistente Doutora da Universidade Estadual Paulista (UNESP) atuando no Departamento de Design do campus de Bauru. Seus interesses de pesquisa situam-se nas áreas de Prática de projeto e o papel do designer na sociedade; Teoria e crítica do design; Design centrado no ser humano; Design e Tecnologias inteligentes baseadas Inteligência artificial e seus impactos sociais.

EMAIL: carolina.pizarro@unesp.br