REVISTA DESIGN & TECNOLOGIA
ISSN: 2178-1974
2024, Vol. 14, No. 28
DOI 10.23972/det2024iss28pp100-115

PGDESIGN
www.pgdesign.ufrgs.br

Design e ética em IA: reflexões a partir do impacto social do Lensa enquanto produto digital baseado em Inteligência Artificial

Carolina V Pizarro[1]

1 *Departamento de Design, Universidade Estadual Paulista (UNESP), Faculdade de Arquitetura, Artes, Comunicação e Design, Bauru, Brasil*

RESUMO

As técnicas de Inteligência Artificial – IA – vêm sendo aplicadas em diferentes projetos de produtos, sistemas e serviços. Atrelados ao avanço do uso da IA, surgem desafios de ordem ética que permeiam a interação dos seres humanos com tais tecnologias, o que motivou a UNESCO a publicar no ano de 2022 documento oficial com recomendações sobre a ética da IA como uma referência para boas práticas em projetos desta natureza. A partir desse cenário, o presente artigo tem como objetivo refletir e discutir sobre ética, IA e seus impactos sociais, tendo como base o documento de recomendações da UNESCO. Para tal fim, foi analisada a documentação de um produto digital voltado a edições fotográficas denominado Lensa. Esse produto, quando lançado, obteve ampla aceitação por produzir ilustrações originais generativas, ao mesmo tempo em que suscitou questionamentos de ordem ética sobre o tratamento e o uso de dados sensíveis dos usuários, bem como impactou diretamente a comunidade de ilustradores e designers digitais. Esta pesquisa tem perfil qualitativo, do tipo estudo de caso, conduzida por meio de análise documental, na qual foi realizado, à luz dos dez princípios para a ética da IA propostos nas recomendações da UNESCO, um exame da documentação de software do produto Lensa que está disponível aos usuários. Como resultado, verificou-se que o aplicativo, até o momento da publicação deste estudo, apresenta uma série de pendências do ponto de vista ético, demandando um maior cuidado especialmente relacionado às questões de transparência e explicabilidade da IA.

PALAVRAS-CHAVE

Design;
Ética;
Inteligência Artificial;
Projeto.

Design and ethics in AI: reflections from the social impact of Lensa as an AI-based digital product

ABSTRACT

Artificial Intelligence – AI – techniques have been applied to different product designs, systems, and services. Alongside the increasing the use of AI, ethical challenges arise that permeate the interaction of human beings with such technologies, prompting UNESCO to publish an official document in 2022 with recommendations on AI ethics, as a reference for good practices in such projects. From this scenario, this article aims to reflect and discuss ethics, AI, and their social impacts, drawing upon UNESCO recommendations.

To this end, the documentation of a digital product focused on photo editing, named Lensa, was analyzed. This product, upon its release, gained wide acceptance for producing original generative illustrations, but it also raised ethical concerns about the treatment and use of users' sensitive data, as well as directly impacting the community of illustrators and digital designers. This research adopts a qualitative approach, a case study, conducted through document analysis. Considering the ten principles for AI ethics proposed in the UNESCO recommendations, an examination of the Lensa software documentation available to users was undertaken. As a result, as of the publication of this study, it was found that the Lensa application presents a series of ethical shortcomings, requiring greater attention, especially concerning issues of transparency and explainability of AI.

KEYWORDS

Design;
Ethics;
Artificial Intelligence;
Project.

CONTATO: Carolina V Pizarro – *caroldipp@gmail.com*

© 2024 – Revista Design & Tecnologia

1. INTRODUÇÃO

Atualmente na área de design, as técnicas de Inteligência artificial (IA) têm sido aplicadas em diferentes projetos e para soluções diversas. Assim, é possível encontrar a IA atuando em produtos físicos com suporte digital, tais como os assistentes virtuais conversacionais; soluções digitais embarcadas em produtos tradicionalmente mecânicos, como os automóveis, por exemplo – nos quais tanto manobra quanto estacionamento podem ser assistidos por IA. Também é possível encontrar a aplicação da IA em instituições como bancos e hospitais, auxiliando no gerenciamento de riscos; em jogos de vídeo game e computador, nos quais a IA oferece desafios variados e variáveis aos jogadores; também nos governos, auxiliando, por exemplo, na tomada de decisão em diferentes áreas, além das simulações em ambientes como museus, apenas para citar alguns casos de uso.

Entre as diferentes possibilidades que a IA traz enquanto tecnologia, algumas delas são transformadas em modelos de negócios que até mesmo terminam por tangenciar áreas de atuação do design e, uma vez que se tornam de fácil acesso ao público em geral, são capazes de permitir aos usuários a produção, por conta própria, de conteúdos que demandariam conhecimento técnico específico ou mesmo um projeto de design. O alcance desses resultados – mesmo que de qualidade questionável – é possibilitado pela IA e pelo aprendizado de máquina, os quais se valem da grande quantidade de dados disponíveis em rede para operar.

Recentemente, um sistema em especial baseado em IA ganhou repercussão e popularidade no âmbito da internet e das redes sociais por produzir ilustrações originais geradas a partir de fotos pessoais entregues pelos próprios usuários. Denominado Lensa, esse sistema funciona com base em aprendizado de máquina e utiliza como amostra de treinamento da IA um conjunto de fotos pessoais dos usuários entregues com seu consentimento – e, inclusive, mediante pagamento – retornando a eles os denominados *Magic Avatars*: ilustrações com o rosto da própria pessoa – resultado do aprendizado de máquina a partir de suas fotos –, representados em diferentes estilos de ilustração – também fruto do aprendizado de máquina –, cujos estilos são captados a partir das muitas imagens e lustrações disponíveis na web (Prisma [...], 2022d).

A transparência em relação à operação do Lensa na produção das imagens traz importantes questões do ponto de vista ético referentes àquelas que não são resultado de um processo de design, mas entregam soluções como se fossem. Alguns questionamentos relevantes para que se possa fazer uma leitura crítica desse contexto seriam: como a empresa esclarece o uso que faz dos dados dos usuários que aceitam os termos de serviço? O que é feito com esses dados? Como o Lensa é capaz de retornar tantos estilos de ilustração diversos se não existem designers ou artistas digitais dedicados à essa produção? Que impacto as produções estão causando na comunidade de designers e ilustradores?

Sem a intenção de realizar uma análise jurídica do cenário posto – mas entendendo que implicações desta natureza existem ou podem vir a existir no contexto tratado –, o presente artigo pretende refletir sobre tais questões, tendo como base o documento intitulado *Recomendação sobre a ética da Inteligência Artificial*, elaborado pela UNESCO, e tecer, a partir dele, considerações sobre os desdobramentos da tecnologia empregada no referido aplicativo – doravante denominado *app* – a qual pode interferir na produção da área do design, sinalizando as preocupações do ponto de vista ético que estas interações provocam. Entendendo que atualmente a área de design desenvolve variadas soluções em produtos, sistemas e serviços que integram a IA, também cabe a ela estar atenta às produções que não necessariamente são originadas de um processo de design, mas têm poder de influenciar na produção da área positiva ou negativamente.

Avaliar as implicações éticas das produções que envolvem em algum nível a interação ou integração dos campos da IA e do Design é o que vai permitir o desenvolvimento de novos pensamentos, estratégias e processos capazes de ponderar riscos e benefícios dessas aplicações, não somente para produções oriundas do campo do design, mas também para produções que, mesmo não realizadas pelo design, têm como consequência um impacto social que reverbera não apenas na própria área, mas na sociedade como um todo.

2. REFERENCIAL TEÓRICO

O campo da Inteligência Artificial vem se expandindo e se desenvolvendo desde a década de 1950. As diferentes técnicas associadas ao desenvolvimento da Inteligência Artificial – especialmente o aprendizado de máquina – têm evoluído de maneira veloz, impactando com a mesma velocidade o cotidiano das pessoas, as quais buscam os benefícios da IA, mas nem sempre têm consciência ou condições de identificar os riscos que seus diferentes usos podem trazer.

Para Suleyman e Bhaskar, a evolução da IA institui uma nova onda tecnológica entre tantas outras que a humanidade já testemunhou historicamente, a qual trará muitas oportunidades e também riscos. Para os autores, "está claro que, se formos bem-sucedidos em replicar a inteligência humana, esse será não somente um negócio lucrativo, mas também uma mudança sísmica para a humanidade, inaugurando uma era na qual oportunidades sem precedentes serão acompanhadas de riscos sem precedentes" (Suleyman; Bhaskar, 2023, local. 7597).

Entendendo a complexidade do panorama que se forma, fruto da interação entre humanidade e IA, o campo do Direito vem discutindo as implicações éticas desta digitalização da sociedade, bem como os impactos que a IA tem nos direitos fundamentais humanos e na própria prática do Direito (Peixoto; Silva, 2019; Alves; Guimarães, 2023).

Para Santos (2023, p. 87), na era da IA, um dos pontos de atenção em relação à garantia de direitos é a preservação da autonomia humana. Nesse sentido, o autor destaca que: "o que se encontra em questão diante da digitalização da sociedade é a centralidade da pessoa humana na ordem jurídica e social, enquanto sujeito dotado de autonomia para realizar escolhas e tomar decisões. A denominada "era digital" surge como desdobramento lógico da aceleração da sociedade da informação e, apropriando-se gratuitamente dos dados sobre a experiência humana, logo se converte em centro de controle e de tomada de decisões que desconsidera a autonomia humana". O prejuízo da autonomia envolve outros fatores que se tornam também pontos de atenção, tais como: a privacidade dos dados, a transparência no seu uso, a explicabilidade de como eles são utilizados, e com qual finalidade.

Nesse cenário, a ética constitui um fator orientador essencial em dois âmbitos: no de projeto de produtos, sistemas e serviços baseados em IA; e no de uso dessas produções. O primeiro diz respeito a projetar produtos, sistemas e serviços que integrem a IA de maneira ética, tendo a própria ética como ponto de partida para o projeto e o que ele vai oferecer. O segundo, relativo ao uso, implica disponibilizar o produto com transparência perante os usuários, explicando nas políticas – ou mesmo dentro do produto – os benefícios e riscos atrelados, para que as pessoas possam fazer um uso consciente, crítico e informado daquela solução.

Considerando ambos os espaços de ação – projeto e uso –, cabe também ao design, enquanto área interdisciplinar responsável por formular as diferentes interfaces entre ser

humano e tecnologia, zelar por um futuro ético para a aplicação da IA tendo em conta seus impactos sociais, problematizando, refletindo sobre a temática e visando formular novos aportes teóricos e práticos para contribuir e atuar nesse processo. E nesse sentido, a ética da profissão figura como parte essencial da discussão.

A reflexão sobre a ética da profissão adentra ao âmbito da deontologia das práticas profissionais, a qual de acordo com Passos (1993, p. 60) caracteriza-se como "[...] um conjunto de normas e princípios que tem por fim orientar as relações dos profissionais com os seus pares, destes com os seus clientes, com a sua equipe de trabalho, com as instituições a que servem entre outros. Como a sua margem de aplicação é limitada ao círculo profissional, faz com que essas normas sejam mais específicas e objetivas".

No Brasil as Associações dos Designers Gráficos – ADG, (200-) – e dos Designers de Produtos – ADP, (2004) – contam com instruções de ordem ética estabelecidas em códigos próprios, contudo, ainda não há um código de ética unificado para toda a categoria de profissionais de design. Em um cenário no qual a profissão não é regulamentada, cabe destacar a significativa contribuição de ambas as iniciativas em fortalecer a prática profissional com normas de conduta humanísticas e voltadas a fazer do design um campo dedicado a contribuir de diversas e melhores maneiras com a sociedade. Entretanto, entende-se que seria benéfica a criação de um código de ética unificado para toda a categoria, com flexibilidade de atualização suficiente para se adaptar às novas demandas que surgem para a área, principalmente diante das constantes mudanças de ordem social, cultural, econômica e tecnológica que se abatem sobre o campo.

Partindo da perspectiva do design, apoiando-se em Löbach (2001), os autores Grilo, Van der Linden e Ribeiro (2022, p. 10) destacam que o designer "[...] não está descolado dos impactos do artefato na sociedade e, por isso, espera-se uma postura de responsabilidade do designer diante das contradições entre os processos econômicos e seus impactos sociais e ambientais. O designer pode, então, conformar-se ao que está estabelecido pela organização ou produzir iniciativas que redirecionem a produção tecnológica". Nesse sentido, tem-se que a produção do design é fruto indissociável da responsabilidade de quem a projeta e do quanto este profissional assume tal papel no seu cotidiano e nas relações estabelecidas na sua prática profissional. Um projeto ético nasce de uma prática ética.

2.1 Inteligência Artificial e aprendizado de máquina

A Inteligência Artificial é um ramo da computação que foi criado por Alan Turing com a publicação do estudo de sua autoria denominado *Computing machinery and intelligence* no ano de 1950. Contudo, o termo Inteligência Artificial em si foi cunhado pelo pesquisador John McCarthy, no ano de 1956, durante a conferência *The Dartmouth Summer Research Project on Artificial Intelligence*, um evento dedicado a pensar e discutir as relações sobre a inteligência e as possibilidades de imitação desta por parte das máquinas. Até esse evento o termo utilizado era "simulação computacional" (Buchanan, 2005, p. 57; Boden, 2020, p. 34). Buscando sintetizar os objetivos da IA, Boden (2020) identifica que ela procura preparar os computadores para executar ações que a mente humana é capaz de fazer. A autora destaca ser esse um campo amplo, pois a inteligência não é uma dimensão única, mas sim um "[...] espaço ricamente estruturado com diferentes habilidades de processar informação" (Boden, 2020, p. 13). Dessa maneira, segundo a autora, a IA emprega técnicas diversas, se dedicando, portanto, a muitas tarefas distintas. Em termos de estrutura, para alcançar soluções para os problemas em diferentes âmbitos, a IA se utiliza de técnicas de Aprendizado de Máquina – amplamente conhecido pelo termo em inglês *Machine Learning*.

O aprendizado de máquina ocorre quando, um computador – a partir de conhecimentos estatísticos sofisticados – é capaz de aprender ou melhorar o processamento de dados sem precisar ser programado (Taulli, 2020, p. 35). No que se refere às técnicas de Aprendizado de Máquina, segundo Cortiz (2019, p. 2-3), citando Russel e Norvig (2009), essas podem ser divididas em duas principais abordagens: o aprendizado supervisionado e o não-supervisionado. Na primeira abordagem, os sistemas atuam observando alguns exemplos de pares de entrada e saída e aprendem uma função capaz de mapear uma nova entrada para uma saída específica. Isto é possível mediante um grande volume de dados rotulados, ou seja, dados que tenham sido previamente "explicados" ou identificados, por meio de informações específicas dadas ao sistema sobre eles. Na segunda abordagem – denominada não supervisionada – não é previsto ou demandado que os dados sejam obrigatoriamente rotulados ou "explicados" para o sistema, uma vez que essas técnicas têm a capacidade de criar por si só agrupamentos e segmentações, fazendo isso por meio da identificação de semelhanças nos perfis, o que lhes permite interpretar melhor os próprios dados.

De acordo com o autor, como as técnicas de aprendizado de máquina dependem de dados para aprendizado a partir de exemplos, cabe o cuidado, por parte de projetistas e desenvolvedores, em relação à fonte e à qualidade desses dados. É preciso buscar que os algoritmos – número finito de regras e ações executáveis com a finalidade de resolução de um problema – sejam treinados da melhor maneira e de forma responsável, uma vez que o modo como os dados são escolhidos influenciará no aprendizado e comportamento do sistema, o qual oferecerá, portanto, respostas que serão fruto desse processo e poderão estar enviesadas (Cortiz, 2019). Estabelece-se assim, a necessidade de desenvolver um trabalho cuja ética deve ser um aspecto norteador para lidar corretamente com tais dados, uma vez que que as decisões tomadas pelos algoritmos e máquinas podem gerar impactos em larga escala.

Considerando os impactos de tais tecnologias, deve-se atentar aos riscos os quais os seres humanos podem estar submetidos quando da interação com elas. Segundo Boechat (2019, p. 54), uma vez que a I.A. vem sendo utilizada para gerar respostas quase humanas, a ética passa a ser uma parte central das discussões nesse campo do conhecimento. O autor utiliza como exemplo, citando PTI (2020), que se um desenvolvedor dedicado a um projeto de IA tiver um viés negativo em relação a determinada demanda, esse viés pode se infiltrar na solução que ele projeta e, assim, afetar o resultado das respostas. O viés nos resultados é também um problema destacado por Kaufman (2022, p. 16-17), uma vez que pode trazer resultados discriminatórios como, por exemplo, de gênero, idade, etnia, entre outros. Embora, em geral, se atribua o viés às bases de dados tendenciosas, ele pode emergir muito antes de se coletar os dados para treinar a IA, relacionando-se diretamente às decisões tomadas pelos desenvolvedores, posto que as variáveis que são inseridas no modelo de IA a ser treinado, inclusive, irão determinar quais dados serão selecionados. Tal aspecto da tecnologia se torna particularmente desafiador e preocupante do ponto de vista ético, considerando a existência da chamada "caixa preta" da IA, especialmente no Aprendizado Profundo – ou *Deep Learning*.

O *Deep Learning* é uma subárea do *Machine Learning* que permite o processamento de enormes quantidades de dados com capacidade para identificar relações e padrões que os

seres humanos poderiam não conseguir identificar se fossem postos a desempenhar a mesma tarefa (Taulli, 2020). Segundo Taulli (2020, p. 98) a palavra "*deep*" se refere ao número de camadas ocultas na rede neural – um modelo de Inteligência artificial projetado para imitar o cérebro que contém diversas camadas de análise as quais fornecem grande parte do poder de aprendizagem à máquina.

A denominada "caixa preta" da IA constitui a área do sistema onde ele é alimentado com dados e que, a partir deles, entrega respostas. Mas, como nesses contextos o sistema se "auto ensina" "[...] não há como saber o que ele efetivamente sabe nem quais parâmetros usou para chegar aos seus resultados" (Boechat, 2019, p. 54). É nesse intervalo que se estabelecem os riscos. Kaufman (2022) destaca que, como toda tecnologia, a IA é também social e humana e que seus efeitos percebidos e vivenciados na sociedade irão depender do que os seres humanos fazem com a tecnologia, como a percebem, como experimentam suas possibilidades e a usam, como a inserem nos ambientes técnico-sociais e enfatiza que cabe à sociedade debater e decidir "[...] sobre se a IA deve ser aplicada em todos os domínios e para executar todas as tarefas, e se o uso da IA em aplicações de alto risco se justifica" (Kaufman, 2022, p. 24). Para a autora, o desafio reside em buscar um equilíbrio entre mitigar (ou eliminar) os riscos e preservar o ambiente de inovação, defendendo que isso deve ser feito de modo que não supervalorize ou demonize a IA enquanto tecnologia. Atingir esse equilíbrio é um desafio que se coloca não somente para a Ciência da Computação, área-base da produção de conhecimento da IA, mas também das demais áreas que a tangenciam ou atuam com ela de maneira interdisciplinar, como é o caso do Design.

O design é a área responsável por traduzir a tecnologia em uma linguagem que o usuário compreenda por meio das interfaces (Johnson, 2001) e também por criar o corpo de conhecimento teórico e crítico encarregado de buscar caminhos pelos quais serão construídas as conexões entre as potencialidades das diversas tecnologias e os seres humanos, posicionando-se, portanto, na intersecção entre a cultura da vida cotidiana, da tecnologia e da economia, humanizando essa relação (Bonsiepe, 2012). Essa visão crítica da profissão permite que se busque constantemente uma prática humanística, colocando o ser humano no centro do processo, como defende Norman (2010, p. 147) ao identificar que, à medida que se avança tecnologicamente, é necessário considerar o impacto das várias mudanças sobre os indivíduos e as sociedades e que "os designers estão na linha de frente destas preocupações, pois é o designer que traduz ideias em realidade". Cabe então, também a este campo, a discussão dos impactos sociais que as soluções baseadas em IA podem apresentar quando acessadas pelas pessoas usuárias. Acompanhar os impactos destes projetos quando em uso na coletividade permite a construção de novas inferências e análises, buscando novas perspectivas para pesquisas e políticas públicas, de modo que os desenvolvimentos se tornem cada vez mais adequados do ponto de vista ético.

2.2 Ética para a IA: recomendações da UNESCO

A IA está presente e de maneira crescente em várias atividades e aspectos da vida humana. Santaella (2021) destaca que conjuntamente à escalada da tecnologia ampliam-se também suas instâncias e os dilemas inerentes à ela. Por isso, o assunto faz parte da agenda de governos e organismos internacionais de porte. De acordo com a autora, ao mesmo tempo em que há uma corrida do mundo empresarial rumo à transformação digital "[...] 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).

As preocupações com as implicações éticas da IA estiveram no centro das discussões da 41ª sessão da Conferência Geral da Organização das Nações Unidas para a Educação, a Ciência e a Cultura reunida em Paris de 9 a 24 de novembro de 2021, na qual a UNESCO apresentou um documento intitulado "Recomendações sobre a ética da Inteligência Artificial". O documento é resultado de um processo que teve início em novembro de 2019, na 40ª sessão da mesma conferência, a qual aprovou a 40 C/Resolução 37 a qual determinou que a diretora-geral elaborasse um instrumento internacional normativo sobre a ética da inteligência artificial (IA) na forma de uma recomendação, a qual deveria ser apresentada à Conferência Geral em sua 41ª sessão, em 2021. O documento de 44 páginas, publicado em 2022, considera as diferentes dimensões nas quais o uso da IA pode levantar questões éticas, destacando valores e princípios orientadores a serem seguidos pelos Estados no desenvolvimento de estruturas ou mecanismos políticos, para que estes sejam seguidos por outras partes interessadas, como, por exemplo, empresas do setor privado, instituições acadêmicas e de pesquisa e sociedade civil. O objetivo principal das discussões que fundamentaram e estão contidas no documento é salvaguardar os seres humanos e seus direitos ao interagirem com produções que envolvam IA.

Como parte do escopo do documento apresentado pela UNESCO é destacada a abordagem ética da IA como uma reflexão normativa sistemática, pensada de maneira holística, de modo que ela seja abrangente e também multicultural, estando em constante evolução de valores, princípios e ações interdependentes, podendo orientar as sociedades para lidar de forma responsável com os impactos, sejam eles conhecidos ou desconhecidos, das tecnologias de IA, os quais podem se abater sobre os seres humanos, sociedades, meio ambiente e ecossistemas "[...] oferecendo-lhes uma base para aceitar ou rejeitar essas tecnologias" (UNESCO, 2022, p. 10). A recomendação da instituição presente no documento entende a ética como uma base dinâmica para que se possa avaliar e orientar de maneira normativa as tecnologias de IA fazendo referência "[...] à dignidade humana, ao bem-estar e à prevenção de danos – como uma bússola e tendo como fundamento a ética da ciência e da tecnologia" (UNESCO, 2022, p. 10).

A publicação traz valores e princípios que devem ser respeitados por todos os atores envolvidos em algum grau no ciclo de vida dos sistemas de IA e, quando necessário, devem ser promovidos por meio de alterações à legislação existente e da elaboração de novas leis, regulamentos e diretrizes empresariais, sendo que "[...] isso deve estar de acordo com o direito internacional, incluindo a Carta das Nações Unidas e as obrigações de direitos humanos dos Estados-membros" (UNESCO, 2022, p. 18). Como valores a publicação apresenta quatro pontos (UNESCO, 2022, p. 18-20), sendo o primeiro valor *Respeito, proteção e promoção dos direitos humanos, das liberdades fundamentais e da dignidade humana*, o qual estabelece que os sistemas de IA, quando desenvolvidos, devem garantir o respeito, proteção e promoção da dignidade, dos direitos humanos e liberdades fundamentais, conforme estabelecido pelo direito internacional, incluindo as leis internacionais relativas aos direitos humanos, priorizando que a qualidade de vida dos seres humanos seja reforçada, não permitindo que ocorra nenhuma violação ou abuso de direitos humanos ou liberdades fundamentais, ou da dignidade humana (UNESCO, 2022, p. 18).

O segundo valor versa sobre a *Prosperidade ambiental e ecossistêmica* e diz respeito ao meio ambiente e os

ecossistemas, entendendo-os como uma necessidade existencial aos seres humanos e demais seres vivos e que, por isso, devem ser reconhecidos, protegidos e promovidos ao longo do ciclo de vida dos sistemas de IA. Assim sendo, a documentação orienta que se deve zelar pela proteção e restauração do meio ambiente e do ecossistema, bem como para o desenvolvimento sustentável, reduzindo o impacto dos sistemas da IA no meio ambiente (UNESCO, 2022, p. 19).

Como terceiro valor, a recomendação da UNESCO destaca *Garantir diversidade e inclusão* salientando que no ciclo de vida de uma solução em IA devem ser garantidos o respeito, a proteção e a promoção da diversidade e da inclusão em consonância com o direito internacional, assim como o incentivo à participação ativa de todos os indivíduos ou grupos, independentemente de raça, cor, ascendência, gênero, idade, língua, religião, opinião política, nacionalidade, origem étnica ou social, condição econômica, social ou de nascimento, deficiência e ou quaisquer outros motivos (UNESCO, 2022, p. 19-20).

O quarto valor denominado *Viver em sociedades pacíficas, justas e interconectadas* se refere à necessidade de os agentes de IA atuarem de maneira favorável para garantir sociedades pacíficas e justas para que, ao longo do ciclo de vida dos sistemas de IA, sejam promovidas paz, inclusão, justiça, equidade e interconectividade, sendo que também devem ser realizados esforços para superar quaisquer obstáculos nesse sentido e não obter vantagens de possíveis ausências de infraestruturas tecnológicas, educação ou habilidades, bem como estruturas legais essenciais, de modo a causar algum prejuízo (UNESCO, 2022, p. 20). Segundo o documento (p. 18), tais valores desempenham o papel de ideais motivadores para a elaboração de políticas e normas jurídicas e também representam os fundamentos dos princípios que são em número de dez.

O primeiro princípio é o de *Proporcionalidade e não causar dano*, o qual diz respeito ao cuidado para que nenhum dos processos relacionados ao ciclo de vida dos sistemas de IA exceda o necessário para atingir objetivos ou metas legítimas, sendo que eles devem ser adequados a cada contexto. Também devem ser asseguradas 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).

Como segundo princípio, a publicação traz *Segurança e proteção*, o qual destaca que em se tratando de sistemas de IA deve-se garantir que os danos potenciais (riscos de proteção) e as vulnerabilidades a ataques (riscos de segurança) sejam evitados, impedidos e eliminados durante o ciclo de vida desses sistemas. Também, a criação de marcos sustentáveis de proteção à privacidade dos dados de acesso permitirá que melhores modelos de IA sejam desenvolvidos, favorecendo uma IA segura e protegida (UNESCO, 2022, p. 20).

O terceiro princípio refere-se a *Justiça e não discriminação* e diz respeito ao cuidado que se deve ter na promoção de justiça social, equidade e não discriminação de qualquer tipo, buscando a conformidade com o direito internacional e a partir de uma abordagem inclusiva, de modo a garantir que os benefícios das tecnologias de IA estejam disponíveis e acessíveis a todos. Obviamente, essas ações devem levar em consideração as especificidades dos diferentes grupos que possam ser desfavorecidos ou marginalizados, além de atuar contra a discriminação e a determinação algorítmica tendenciosa (UNESCO, 2022, p. 20-21).

No quarto princípio, o documento aborda a *Sustentabilidade*, destacando que o desenvolvimento de sociedades mais sustentáveis depende também de ações no âmbito tecnológico. Assim, a avaliação contínua dos impactos das tecnologias de IA nos aspectos humanos, sociais, culturais, econômicos e ambientais deve ser realizada com pleno conhecimento das implicações que tais tecnologias trazem para a sustentabilidade, criando um conjunto de objetivos a serem monitorados e entendendo-os como aspectos em constante evolução (UNESCO, 2022, p. 21).

Como quinto princípio, o documento traz o *Direito à privacidade e proteção de dados,* destacando que a privacidade é um direito essencial para proteger a dignidade, a autonomia e a capacidade de ação humanas, necessitando ser respeitada, protegida e promovida ao longo do ciclo de vida dos sistemas de IA. Nesse contexto, os dados devem ser coletados, utilizados, compartilhados, arquivados e apagados em consonância com o direito internacional e em alinhamento com os valores e princípios estabelecidos na recomendação da UNESCO aqui discutida. Não obstante, devem também ser respeitados os demais marcos jurídicos nacionais, regionais e internacionais relevantes (UNESCO, 2022, p. 22).

O sexto princípio, *Supervisão humana e determinação*, visa garantir a possibilidade de atribuição de responsabilidade ética e legal em qualquer estágio do ciclo de vida dos sistemas de IA, compreendendo a supervisão como não somente humana individual, mas também como pública desses sistemas e seus impactos. Neste ponto entende-se que as pessoas podem recorrer à IA por motivos de eficácia para alguma necessidade ou tarefa, mas que a decisão de ceder o controle para tomadas de decisão em contextos limitados continua sendo de responsabilidade humana sendo, portanto, também sua a prestação de contas. É destacado neste princípio que decisões envolvendo vida e morte não devem ser transferidas a sistemas de IA (UNESCO, 2022, p. 22).

O sétimo princípio, denominado *Transparência e explicabilidade*, destaca a importância das soluções em IA serem transparentes para os envolvidos e também explicáveis, características essenciais para garantir o respeito, a proteção e a promoção dos direitos humanos, das liberdades fundamentais e dos princípios éticos. Com transparência, permite-se que as pessoas entendam o impacto do uso da tecnologia e como cada estágio dela é colocado em prática, considerando o contexto e a sensibilidade do sistema de IA. Já explicabilidade significa tornar inteligível e fornecer informações sobre o resultado dos sistemas de IA, a qual também está relacionada com a transparência, pois esses resultados devem ser compreensíveis e rastreáveis, conforme o contexto (UNESCO, 2022, p. 22-23).

O princípio de número oito se refere a *Responsabilidade e prestação de contas*, segundo o qual os atores de IA e os Estados-membros devem assumir responsabilidades éticas e jurídicas de acordo com o direito nacional e internacional. Segundo este princípio, devem ser desenvolvidos mecanismos para supervisão, avaliação de impacto, auditoria e diligência, incluindo proteção de denunciantes, para se garantir que exista a responsabilização pelos sistemas de IA, bem como seu impacto na sociedade (UNESCO, 2022, p. 23).

Como princípio de número nove, *Conscientização e alfabetização*, destaca-se a importância e a necessidade de se promover a compreensão pública das tecnologias de IA e do valor dos dados por meio de educação aberta e acessível, engajamento cívico, habilidades digitais e treinamento em ética da IA, bem como de alfabetização midiática e informacional, além de treinamento conduzido em conjunto por diferentes atores como governos, organizações intergovernamentais, sociedade civil, universidades, meios de comunicação, líderes comunitários e setor privado, sendo que a compreensão dos sistemas de IA deve ser baseada em seu impacto nos direitos humanos, no meio ambiente e nos ecossistemas (UNESCO, 2022, p. 23).

O último princípio diz respeito a *Governança e colaboração adaptáveis e com múltiplas partes interessadas*, identificando que o direito internacional e a soberania nacional devem ser respeitados no uso de dados e que os Estados podem regulamentar tanto os dados gerados dentro de seus territórios, quanto os que os atravessam. Com isso, podem também tomar medidas para uma regulamentação de dados eficaz baseada 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 operacionalizar os valores e princípios apresentados anteriormente, a UNESCO ainda propõe 11 áreas de ação política que, em linhas gerais, envolvem os pontos sintetizados na Figura 1.



Figura 1 Visão geral das áreas políticas estabelecidas para operacionalizar as recomendações da UNESCO sobre a ética da Inteligência Artificial. Fonte: UNESCO, 2022, p. 26-39. Adaptado pela autora.

Ao criar essa documentação a UNESCO institui um marco significativo fornecendo orientações sobre áreas da sociedade nas quais é urgente que Estados, instituições e coletividade em geral ampliem os debates sobre como lidar com a aplicação e uso da Inteligência Artificial tanto no presente, quanto no futuro.

Entre essas soluções, muitas passarão pelo design e serão entregues na forma de produtos, sistemas e serviços. Dessa maneira, é necessário que o debate seja feito dentro da área para que se busque não somente caminhos melhores para os projetos que integrem a IA, mas também uma ampliação da produção científica, teórica e crítica do design no acompanhamento, análise e crítica destas produções, bem como de seus resultados e impactos na sociedade.

Para Flusser (2007, p. 200), há a necessidade de os designers se interessarem também pela ética envolvida em sua atuação e produção. Segundo o autor, há tempos atrás o designer tinha como meta principal a produção de objetos úteis, realizados com exatidão científica, apresentando aspecto bonito e sendo também funcionais. Naquele contexto, considerações morais ou políticas raramente estavam em jogo, diferentemente dos dias atuais, nos quais, a questão da responsabilidade moral e política do designer adquiriu nova importância e, diante das quais, o autor enfatiza que "[...] o eventual desinteresse dos designers por essas questões poderá levar à total ausência de responsabilidade" (Flusser, 2007, p. 202).

Assim, as mudanças tecnológicas que atingem o design e, ao mesmo tempo, influenciam sua produção, trazem aos profissionais da área tanto oportunidades quanto desafios que precisam ser abordados. Em relação aos desafios, os de natureza ética serão os balizadores dessa produção para que ela seja humanizadora e emancipatória do ponto de vista do seu impacto, mesmo que artificial, e do ponto de vista tecnológico.

Também Dignum (2019) enfatiza que desenvolver uma IA responsável não é sobre as características dos sistemas de IA em si, mas sim sobre o papel dos seres humanos que criam e operam tais tecnologias. Nesse sentido, somos então responsáveis por "[...] como construímos sistemas, como

usamos sistemas e quanto permitimos que esses sistemas decidam e ajam por conta própria" (Dignum, 2019, p. 7, tradução nossa). Há portanto um papel a ser desempenhado pelo design enquanto área não somente produtora, mas também problematizadora dessa realidade, uma missão que deve ser desempenhada pelos pesquisadores e docentes da área, estudantes de design, designers atuantes em campo, sendo também uma contribuição pretendida por este artigo.

3. PROCEDIMENTOS METODOLÓGICOS

O presente artigo configura uma investigação qualitativa do tipo exploratória a qual, de acordo com Gil (2008), tem o objetivo de oferecer uma visão geral aproximada de determinado fato sendo um tipo de pesquisa realizado especialmente quando o tema escolhido "[...] é pouco explorado e torna-se difícil sobre ele formular hipóteses precisas e operacionalizáveis" (Gil, 2008, p. 27). O método científico adotado é o estudo de caso, que Segundo Yin (2001, p. 32) "[...] investiga um fenômeno contemporâneo dentro de seu contexto da vida real, especialmente quando os limites entre o fenômeno e o contexto não estão claramente definidos". Para o autor, os estudos de caso têm como poder diferenciador a capacidade de lidar com uma ampla variedade de evidências, tais como documentos, artefatos, entrevistas e observações, sendo "[...] generalizáveis a proposições teóricas, e não a populações ou universos. Nesse sentido, o estudo de caso, como o experimento, não representa uma 'amostragem', e o objetivo do pesquisador é expandir e generalizar teorias (generalização analítica) e não enumerar frequências (generalização estatística)" (Yin, 2001, p. 29).

Para a realização deste estudo, procedeu-se análise documental realizada a partir da documentação do referido *app*, disponível para acesso por parte dos usuários, operando sua averiguação à luz dos dez princípios para a ética da IA propostos pela UNESCO e citados anteriormente. Optou-se por esse tipo de análise pelo fato de a documentação ser um registro público e inequívoco de como o produto tem sido apresentado e conduzido. A análise documental, conforme Cellard (2012, p. 304-305) tem como vantagem a possibilidade de favorecer que o material seja desconstruído e reconstruído visando responder ao questionamento que originou a pesquisa a partir das ligações que o pesquisador faz entre os fatos acumulados e os elementos de informação: "A fim de estabelecer essas ligações e de constituir configurações significativas, é importante extrair os elementos pertinentes do texto, compará-los com outros elementos contidos no *corpus* documental onde [...] é a leitura repetida que permite, finalmente, tomar consciência das similitudes, relações e diferenças capazes de levar a uma reconstrução admissível e confiável. As combinações possíveis entre os diferentes elementos contidos nas fontes estabelecem-se em relação ao contexto, à problemática, ou ao quadro teórico [...]". A essa altura ocorre, segundo o autor, a sincronicidade na qual uma associação de ideias forma uma explicação. Ou seja, "[...] em que um certo raciocínio se constrói repentinamente, e em que uma ligação se estabelece entre vários fatos e faz-se à luz" (*Ibid.*, p. 305).

Assim, com base na análise realizada em documentação disponível aos usuários, busca-se discutir as implicações éticas envolvidas e os impactos sociais verificados a partir da ampla aceitação e uso do aplicativo Lensa, uma tecnologia para produção de ilustrações generativas pessoais. Cabe destacar que, durante a análise documental, foram buscados pontos de atenção que envolvessem algum risco de ordem ética para as pessoas no uso do produto.

Dessa maneira, os tópicos da documentação que se mostraram adequados e corriqueiros do ponto de vista ético no uso de produtos digitais não foram discutidos, uma vez que compõem o mínimo esperado pela sociedade em relação à prestação de serviços via digital.

Por fim, importa reiterar que a análise aqui apresentada não é jurídica, mas sim realizada a partir da lente do design, pautada na consciência dos direitos fundamentais humanos e com foco na ética da experiência de uso por parte das pessoas que acessam a tecnologia por meio do produto.

4. ANÁLISE DO LENSA: UM EXEMPLO DE DESAFIO ÉTICO PARA O DESIGN ENVOLVENDO INTELIGÊNCIA ARTIFICIAL

A partir das discussões anteriormente apresentadas, é possível afirmar que a Inteligência Artificial, atualmente, proporciona o desenvolvimento de diversos projetos com diferentes funcionalidades podendo atender variadas necessidades, mas junto com estes avanços surgem também desafios de múltiplas ordens. Algumas dessas produções, no entanto, podem tangenciar ou mesmo interferir na produção de alguma área em específico, como, por exemplo a do design. Um caso que recentemente apresentou impacto desta natureza envolve o aplicativo Lensa da empresa Prisma Labs Inc. com sua popularidade inesperada e a significativa interferência gerada na área de atuação de ilustradores/designers.

A Prisma Labs é uma empresa norte americana que atua com o objetivo de oferecer soluções para edição de vídeos e fotos, afirmando que suas tecnologias "[...] permitem que qualquer pessoa sem experiência profissional ou habilidades de edição obtenha ótimos resultados em questão de segundos" (Prisma [...], 2023a, tradução nossa).

De acordo com o site da empresa, ela foi fundada no ano de 2016 e ganhou popularidade com o aplicatvo Prisma, o qual permitia, por meio da aplicação de filtros à escolha do usuário, transformar fotografias em peças visuais com efeitos variados como, por exemplo, fotos que pareciam ter sido desenhadas à mão ou pintadas à tinta (Prisma [...], 2023b), como mostra a Figura 2.



Figura 2 Exemplo de atuação do aplicativo Prisma. À esquerda do leitor a foto original e à direita a mesma foto com filtro aplicado simulando pinceladas de pintura. Fonte: Prisma [...], 2023b.

O sucesso do aplicativo Prisma permitiu à Prisma Labs continuar seus desenvolvimentos em edição de fotos e vídeos, passando a aplicar Inteligência Artificial para o aprimoramento da edição e manipulação de imagens, o que resultou em uma prática mais refinada, sintetizada em um produto mais complexo e poderoso: o Lensa. Lançado em 2018, o Lensa (Prisma [...], 2023c) é o segundo aplicativo da empresa e, no momento da finalização deste artigo, computa mais de 10 milhões de downloads na loja de aplicativos do Google – a loja da Apple não disponibiliza o número de downloads dos produtos disponíveis em seu catálogo e, por este motivo, os dados não figuram aqui.

Entre as funcionalidades de edição, o Lensa traz o conceito do que a empresa chama de *Magic Avatars* o qual é divulgado

da seguinte maneira: “Não é edição de fotos, nem filtro ou efeito [...] Gere avatares alucinantes em vários estilos artísticos a partir de seus retratos com uma das IAs mais avançadas já criadas” (Prisma [...], 2023d, tradução nossa). A produção dos *Magic Avatars* funciona com os usuários aceitando termos e condições e cedendo no mínimo 10 fotos suas, do tipo autoretrato, com o rosto em diferentes posições, para que o *app* por meio da Inteligência Artificial analise as imagens e, portanto, os rostos contidos nelas. A partir dessa amostra de fotos, a IA é então retreinada e devolve ao usuário uma série de imagens de si mesmo ricamente ilustradas, adotando diferentes estilos de ilustração como mostra a Figura 3, a qual contém exemplos captados do próprio site do *app*. Existem testes gratuitos, mas a monetização do produto é feita por meio da compra, por parte dos usuários, de planos de pacotes de imagens geradas.



Figura 3 Exemplos dos denominados *Magic Avatars* produzidos pelo *app* Lensa. Diferente do *app* Prisma, que apenas alterava a imagem original aplicando filtros, no Lensa existe a criação de estilos de ilustração, roupas e cenários completamente novos para a mesma pessoa. Fonte: Prisma [...], 2023c.

O que chama a atenção nesta produção é que diferentemente do primeiro *app* da empresa, o previamente citado Prisma – que entregava como serviço a aplicação de filtros artíticos – , no *app* Lensa, os chamados *Magic Avatars* não são resultado da simples aplicação de filtros nas fotos para causar efeitos, eles efetivamente são imagens completamente novas, baseadas no treinamento da IA a partir das fotos dos usuários, a qual aplica estilos de ilustração altamente elaborados e diversos. A partir dessa dinâmica, duas preocupações de ordem ética começaram a surgir em relação ao produto e seu impacto social: o primeiro diz respeito a como a empresa trata os dados faciais utilizados para treinar a IA e retornar as imagens; o segundo refere-se ao questionamento de como a IA é capaz de produzir ilustrações sem que existam artistas contratados para criar a base de dados que treina a IA, ou seja, de onde saem as referências de estilo de ilustrações tão variados que treinam a IA para elaborar essas imagens.

No caso da primeira preocupação de ordem ética, os dados faciais, ao se visitar a política de privacidade – datada de 25 de outubro de 2023 – disponível no site e acessada para análise neste artigo em 09 de dezembro de 2023, a empresa identificava no item *Introdução e escopo* que poderia coletar vários tipos de informações, incluindo, entre outros, os denominados “dados pessoais” e faz o alerta “se você não deseja que processemos seus Dados Pessoais conforme descrito nesta Política de Privacidade, não use o Lensa” (Prisma [...], 2023e, tradução nossa). Apenas esse aviso, do modo que se apresenta redigido, praticamente transfere para as mãos dos usuários toda a responsabilidade sobre o que possa vir a acontecer com seus próprios dados quando sob a tutela da empresa para operacionalizar o produto.

Sabe-se que ao utilizar qualquer serviço baseado em software, a concordância com a política de privacidade e os termos de uso, implica a ciência e em certa medida o “aceite dos riscos” que tal tipo de produção pode oferecer. Entretanto, cabe vigilância redobrada quando a política de privacidade estabelece que se o potencial usuário não aceita correr riscos, então não deve utilizar o produto. Quando o adequado seria a empresa se comprometer de maneira transparente para que a experiência de uso fosse o mais segura possível, inclusive destacando direitos, ensinando estratégias de segurança e indicando caminhos, tanto administrativos quanto jurídicos, possíveis para o usuário acionar a empresa caso algo fuja do previsto com o uso do produto ou, eventualmente, ocorra alguma situação de exceção.

Mais à frente, no item de número três – *Como processamos, compartilhamos e mantemos suas fotos, vídeos e dados de imagem* –, a mesma política informa que os dados faciais seriam tratados juntamente a outros elementos da foto sob a denominação “Dados de Imagem” e admite que “[...] quando você envia fotos ou vídeos para o Lensa, processamos suas fotos e vídeos usando tecnologia que identifica os dados que estão nas fotos e vídeos e nos fornece informações, como informações sobre sua posição facial, orientação e topologia (‘Dados Faciais’) e outros elementos da sua foto (coletivamente, ‘Dados de Imagem’)" (Prisma [...], 2023e, tradução nossa). Quando se acessa o item quatro *Dados pessoais que coletamos automaticamente*, é possível encontrar, entre outros dados – tais como compra, assinatura e etc. – novamente o termo “Dados de Imagem”. Mas, ao tentar identificar em específico o que é coletado automaticamente, a política sugere voltar ao item anterior. Fica entendido, desse modo, que o que existe de informação sobre os dados de imagem está no item três da política e que, portanto, os “Dados Faciais” lá descritos são tratados como dados pessoais, sobre os quais a empresa alega ter amplos direitos.

Neste ponto, vale a ressalva de que não se encaixam – até o momento da escrita deste texto – na categoria de dados pessoais os *Magic Avatars*, ou seja, aquelas ilustrações geradas por Inteligência Artificial a partir das fotos originais e que contêm uma cópia do mapa facial oss usuários. Assim sendo, como estes últimos não são considerados “Dados Pessoais” pela empresa, não possuem qualquer diretriz de privacidade esclarecida na política vigente.

Sobre os *Magic Avatars*, o item três da referida política se mostra ambíguo ao identificar no subtópico *Magic Avatars* que “Imediatamente após a geração bem-sucedida de Avatares, suas fotos originais, a cópia do modelo de difusão estável, bem como quaisquer dados de imagem associados, são excluídos permanentemente de nossos servidores AWS. Não usamos seus dados pessoais para treinar e/ou criar produtos/inteligência artificial separados” (Prisma [...], 2023e, tradução nossa). Apesar dessa afirmação, no subitem seguinte – *Seus Avatares* – há uma contradição ao afirmar que “seus Avatares são armazenados em nossos servidores (fornecidos pela Amazon Web Services (EUA)) e estão disponíveis para você no aplicativo Lensa a qualquer momento, por meio de qualquer dispositivo a partir do qual você faça login no aplicativo Lensa, até que você decida (i) excluir os Avatares ou (ii) excluir sua conta Lensa” (Prisma [...], 2023e, tradução nossa).

Se o subitem *Magic Avatars* identifica que imediatamente após a geração dos avatares são excluídos permanentemente dos servidores as fotos, o modelo de treinamento da IA e quaisquer dados de imagem associados, como os avatares ainda ficam disponíveis para os usuários nos servidores conforme exposto no subitem seguinte *Seus Avatares*?

Informações ambíguas dessa natureza identificam o quão urgente se torna exigir que as políticas de produtos envolvendo IA tragam maior transparência e explicabilidade não só sobre o modelo de IA utilizado e como ele funciona, mas também sobre os processos envolvendo o produto e o uso de dados. Pois não é identificado com clareza o que acontece com as ilustrações do tipo *Magic Avatars* originadas a partir do mapeamento facial das amostras de fotos entregues pelas pessoas, uma vez que

esses avatares contêm uma cópia dos dados faciais de cada uma delas.

A referida política no item cinco, que trata *As finalidades e as nossas bases legais para o tratamento dos seus dados pessoais*, identifica que "podemos usar seus Dados Pessoais para diversos fins, dependendo da categoria de Dados Pessoais e da maneira como você usa e interage com o aplicativo Lensa [...]" (*Ibid.*, tradução nossa). Ali aponta-se que os dados são usados para, entre outros, apresentar o produto, cumprir obrigações e exigir direitos, personalizar ofertas, fornecer suporte técnico, favorecer a comunicação, realizar pesquisas etc. Mas no item seguinte, de número seis, *A quem divulgamos dados pessoais* é afirmado que: "poderemos divulgar as informações que processamos sobre você, incluindo quaisquer Dados Pessoais [...]" (*Ibid.,* tradução nossa), sendo que a empresa identifica que o faz para: empresas que façam parte do mesmo grupo; prestadores de serviços, parceiros de negócios e fornecedores; fornecedores terceirizados de análise e parceiros de publicidade; no caso de serem fundidos, adquiridos, alienados, financiados, vendidos, alienados ou dissolvidos; e em resposta a solicitações legais e para fins de prevenção de danos. Deve-se atentar aqui que, também neste item, em nenhum momento é esclarecido se esses dados pessoais incluem os mapas faciais – pelo que foi visto anteriormente, esta possbilidade poderia existir – o que seria um risco importante ao compartilhar estes dados com terceiros.

O item seguinte da política, o sete, versa sobre *Cookies, kits de desenvolvimento de software e outras tecnologias de rastreamento* comuns no uso de sistemas on-line, mas é o item oito *Os seus direitos em relação aos seus dados pessoais*, que identifica que as pessoas, com base na lei de proteção de dados, têm total direito sobre seus dados, podendo solicitar a qualquer momento o acesso, a modificação, a correção, a restrição de processamento, a portabilidade, o apagamento, a retirada do consentimento e a reclamação.

No item de número nove são tratadas questões de comunicação, contudo é no tópico dez S*egurança de dados* que se estabelece a visão da empresa em relação aos dados dos usuários. Pelo menos até o fechamento do presente artigo, o que se lê é: "[...] embora tomemos medidas para fazer o nosso melhor para proteger os seus Dados Pessoais, não podemos garantir a segurança das informações coletadas, transmitidas para ou através do Lensa, ou uma garantia absoluta de que tais informações não podem ser acedidas, divulgadas, alteradas ou destruídas. Qualquer transmissão dos seus Dados Pessoais é por sua conta e risco. Não nos responsabilizamos pela evasão das medidas de segurança contidas no Lensa. Por favor, entenda que não existe tecnologia ou medida ideal para manter 100% de segurança".

Desse ponto em diante evidencia-se que a postura adotada é de transferir para o usuário toda e qualquer responsabilidade por danos a ele mesmo no uso do produto oferecido pela empresa. Entende-se que, de fato, existem brechas no ambiente on-line que podem ser acessadas por pessoas com diferentes interesses, entretanto, isto não isenta a empresa de suas responsabilidades ao lidar com dados extremamente sensíveis. A empresa se coloca identificando que faz seu melhor, mas o seu melhor de fato seria, para além de oferecer um bom produto, amparar o usuário em situações adversas e de risco, que são justamente aquelas nas quais ele está mais vulnerável e cujas consequências e danos – em se tratando de dados faciais – podem ser imprevisíveis.

Os demais itens da polítca identificam que a empresa pode reter os dados por um período de tempo maior caso necessário e também dão conta da transferência de dados entre países, da proteção de crianças no acesso e uso, dos sites de terceiros, de como contatar um oficial de proteção de dados e sobre a atualização constante da política de privacidade.

Na mesma data de atualização da *Política de privacidade* – 25 de outubro de 2023 – o documento *Termos de uso do aplicativo Lensa* também foi atualizado. Conforme o site da empresa os termos "[...] constituem um contrato vinculativo entre você e nós [...]" (Prisma [...], 2023f). Desse termo constam alguns pontos de atenção referentes à responsabilidade da empresa em relação ao produto que vende.

Logo no começo dos termos, um alerta é destacado em negrito: "se você não concordar com estes Termos, incluindo a disposição de arbitragem obrigatória e renúncia a ações coletivas na Seção 15 (Resolução de disputas; Arbitragem vinculativa), não use o Lensa. Seu uso contínuo do Lensa confirmará sua aceitação destes Termos" (Prisma [...], 2023f).

Tal destaque evidencia que a pessoa usuária do produto deve renunciar aos seus direitos, inclusive de mover ações coletivas. Esse ponto é preocupante, uma vez que o produto oferecido pela empresa foi amplamente aceito, impulsionado por um engajamento orgânico em redes sociais que provocou alta capilaridade e adesão justamente de ordem coletiva. Assim, do modo como está apresentado, a empresa parece se isentar de assumir responsabilidades sobre as externalidades negativas que o produto possa favorecer. Tal ponto fica ainda mais evidente ao longo do texto dos *Termos de uso*.

Os termos são compostos por 23 temas, alguns muito comuns nos termos de uso de produtos digitais, tais como licença, compras no *app*, feedback, comunicação etc. Mas merecem atenção, contudo, três desses itens.

O item 9 estabelece informações sobre violação de direitos autorais, identificando que a empresa respeita os direitos e responde a notificações de supostas violações de direitos autorais que cumpram a lei aplicável, incluindo os critérios para alguma reclamação de violação de direitos autorais. Cabe destacar que este item em especial, só passou a figurar dos termos de uso do Lensa a partir do termo datado de 6 de junho de 2022, já que no termo anterior, de 5 de agosto de 2021, ele não era tratado. Tem-se aí a evidência de que a questão de violação de direitos autorais – e, portanto, a necessidade de estabelecer um termo para ela – somente surgiu no ano de 2022, após os impactos sociais que o uso e a popularidade do produto originaram e a partir dos questionamentos da sociedade, que tiveram ampla repercussão na imprensa (Hunter, 2022; Javaid, 2022).

O item 15, *Resolução de disputas; processo arbitral*, destaca, em negrito e caixa alta, que "[...] qualquer arbitragem sob estes termos, por você ou por nós, será somente individual; arbitragens de classe, ações de classe, ações representativas e consolidação com outras arbitragens não são permitidas. Você renuncia a qualquer direito de ter seu caso decidido por um júri e renuncia a qualquer direito de participar de uma ação coletiva contra nós" (Prisma [...], 2023f). Este item justamente inibe a movimentação coletiva por pessoas usuárias que se sintam prejudicadas de alguma maneira e que busquem mover ações conjuntas contra a empresa. Esse ponto estar expresso nos termos em 2023 é também um reflexo de como a empresa lidou com a ampla repercussão dos protestos e discussões sobre a segurança dos dados faciais captados e, especialmente, os direitos autorais perante o Lensa em relação aos *Magic Avatars* no ano de 2022. Importa aqui esclarecer como isso se deu.

No caso da aplicação de estilos de arte dentro do Lensa, a pessoa usuária envia uma foto que fica armazenada nos servidores por 24 horas, sendo destruída após esse prazo. No caso da criação dos *Magic Avatars*, ocorre processo similar ao supracitado, agora com o envio de 10 a 20 fotos de rosto para o Lensa, as quais são armazenadas nos servidores e utilizadas

para retreinar uma cópia do modelo *Stable Diffusion*. Essa cópia é gerada para personalizar o modelo e então criar os Avatares. Segundo a empresa, imediatamente após a geração bem-sucedida dos avatares, as fotos originais são excluídas, mas, como visto anteriormente em um ponto ambíguo da *Política de privacidade*, os Avatares são armazenados nos servidores para que os usuários possam acessá-los posteriormente.

Ainda assim, como analisado nessa mais recente política de privacidade, não é explicado como são desenvolvidos os estilos artísticos aplicados nas ilustrações dos *Magic Avatars*, os quais configuram a segunda preocupação de ordem ética que o produto suscitou e que foi anteriormente destacada no início deste tópico.

De acordo com Javaid (2022), jornalista do The Washington Post, o modelo *Stable Diffusion* "[...] é alimentado por um banco de dados chamado LAION-5B, construído por pesquisadores de IA" (Javaid, 2022, tradução nossa), algo semelhante a lançar uma rede na Internet para captura. Na mesma reportagem, o designer Jon Lam disse à jornalista que tal banco de dados capta imagens e obras de arte de sites, incluindo milhões de imagens pertencentes a artistas – alcançando o número de 5,85 bilhões de imagens –, as quais são combinadas com texto e "esses pares imagem-texto 'treinam' o modelo *Stable Diffusion* sobre como criar conteúdo como os *Magic Avatars"* (Lam *apud* Javaid, 2022, tradução nossa). A jornalista destaca que a Stability AI, empresa proprietária do modelo *Stable Diffusion* não havia retornado qualquer informação ao jornal quando questionada sobre o assunto.

Muitos artistas reconheceram características de suas produções nos *Magic Avatars* produzidos pelo Lensa, entre eles a designer Lauryn Ipsum, que em entrevista para Javaid, destaca que embora a IA devolva avatares originais, os elementos menores que alimentam esta criação como "[...] paletas de cores, pinceladas, texturas, estilos individuais – foram retirados de artistas como ela sem consentimento, crédito ou compensação" (Ipsum *apud* Javaid, 2022, tradução nossa). Foi inclusive a busca por um reconhecimento pautado na ética que fez com que designers e artistas digitais protestassem contra a arte generativa produzida pela IA.

No mês de dezembro de 2022 um dos principais portais que hospedam o portfólio de diversos designers e artistas digitais, o *ArtStation,* foi tomado por protestos desses profissionais. Muitos deles retiraram todas as suas produções da plataforma para que elas não pudessem mais ser captadas como dados de livre acesso para o treinamento de IA e adicionaram ao seu portfólio uma única imagem com o símbolo de proibido sobrepondo a sigla AI. A intenção foi deixar claro o descontentamento com as produções generativas realizadas pela tecnologia e que utilizam como base artes originais, sem sem consentimento, crédito ou compensação aos artistas, como mostra a Figura 4.

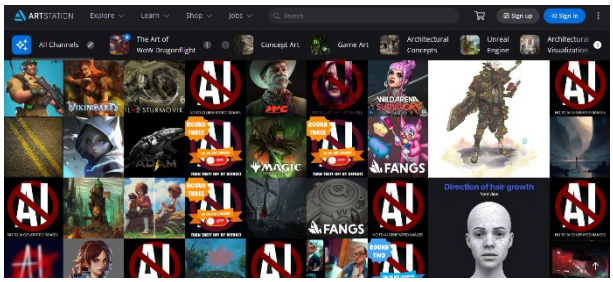


Figura 4 Página inicial do site ArtStation captada em 17/12/2022 na qual é possível verificar a substituição de vários portfólios pela imagem de protesto contra a dinâmica em que a Inteligência Artificial é treinada. Fonte: Artstation, 2022.

À época, adentrando a um dos portfólios presentes na página inicial, o qual continha, ao invés do símbolo de protesto, uma nota escrita, foi possível verificar o conteúdo da nota, a qual explicava sua posição, como mostrado na Figura 5. A íntegra da nota pode ser verificada na Figura 6.

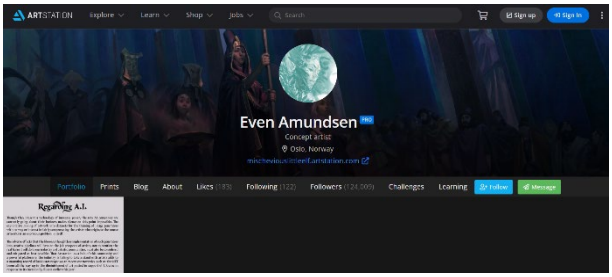


Figura 5 Página do artista digital Even Mehl Amundsen com uma única imagem: sua nota explicativa. Fonte: Amundsen, 2022.



Regarding A.I.

Though they present a technology of immense power, the way AI companies are currently going about their business makes silence on this point impossible. The exploitative mining of artwork into datasets for the training of image generators with no way or interest in fairly compensating the artists who originate the source artworks is an enormous problem in itself.

The adverse effects that the likewise thoughtless implementation of such generators into creative pipelines will have on the job prospects of artists, not to mention the real harm it will do to our industry and artistic communities, must also be considered and mitigated as best possible. That Artstation, as a hub of this community and a powerful platform in the industry, is failing to take a stand with artists adds to a mounting record of lackluster responses to recent controversies such as the NFT boom all the way up to the diminishment of art posted in support of Ukraine in response to its invasion by Russia earlier this year.

Clearly, the calculus of action at Artstation HQ is faulty. Hopefully, they will, when they remember that their platform is sustained by those who make and share art, change their approach and lend their power to the cause of stopping exploitative AI practices and in the future act in such a way as to be able to stand as an integral pillar of this community, not one that supports its exploitation. That choice is theirs. This is mine. Until Artstation shows solidarity, my gallery there will show nothing but this image. Solidarity and fair practices are key to an artistic community, so take care of each other and pay it forward.

Sincerely, Even Mehl Amundsen

Figura 6 Nota explicativa na página do artista digital Even Mehl Amundsen. Fonte: Amundsen, 2022.

Em seu posicionamento, o artista destaca ser impossível ficar em silêncio diante da maneira como as empresas de IA estão conduzindo seu negócio atualmente. Para ele, esta prática atual configura uma mineração exploratória de obras de arte transformando-as em conjuntos de dados para o treinamento de geradores de imagens.E também destaca que conduzir processos neste perfil "[...] sem nenhuma maneira ou interesse em compensar de forma justa os artistas que originaram as obras de arte base é um enorme problema em si" (Amundsen, 2022, tradução nossa).

O artista segue discorrendo, entre outros fatores, que é necessário ponderar sobre o impacto e o dano real que essa prática poderá causar à indústria e comunidades artísticas se estas não se organizarem e se posicionarem para mitigar os riscos exploratórios, sendo então impescindíveis a solidariedade e as práticas justas, uma postura que era esperada de atores como a plataforma ArtStation.

O posicionamento do artista une-se ao da designer Lauryn Ipsum que em sua entrevista à jornalista Mahan Javaid afirmou não temer a tecnologia, mas ansiar por uma atuação mais justa: "[...] nenhum de nós acha que a IA pode substituir os artistas. Tudo o que queremos é a opção de participar, o crédito que ganhamos e o pagamento que merecemos" (Ipsum *apud* Javaid, 2022, tradução nossa).

De fato, embora as ilustrações geradas pela IA como *Magic Avatars* sejam competentes, relatos de usuários dão conta de imagens que vieram com características faciais distorcidas ou membros fora de lugar que causaram entranhamento (Hunter, 2022). O que revela existir uma distância óbvia de qualidade entre uma peça criativa realizada de modo exclusivo, com método, dentro de um projeto de design e uma ilustração generativa realizada por IA com base em captação de imagens aleatórias. Esse fato, contudo, não pareceu ser um problema para o público usuário menos atento, uma vez que o *app* teve grande aceitação e é neste ponto que ele se torna tanto um desafio para a área quanto um sintoma da falta de criticidade no uso do produto.

O exemplo do Lensa é apenas um dos tantos outros impactos que poderão ser verificados quando da aplicação da IA para fins mercadológicos. Nesse caso em especial, a produção de ilustrações originais com base em dados demonstrou ser um desafio para a área de ilustração no design, especificamente no que tange a ilustração como projeto e o reconhecimento dos criadores das artes que treinaram a máquina.

É imperativo ter em mente que tanto o design quanto outras áreas poderão ser afetadas por projetos que possam surgir baseados em IA como resposta a necessidades variadas das pessoas usuárias – além, também, das necessidades futuras que ainda não estão evidentes no âmbito social e, portanto, não foram mapeadas e demandam um projeto. Importa salientar que não se trata de demonizar a tecnologia, como destaca Kaufman (2022), mas saber como extrair dela uma parceria que promova o melhor para os seres humanos. Versando sobre este aspecto, a autora faz uma analogia com os automóveis, afirmando que embora os motoristas em geral não saibam fabricar ou consertar um carro, eles têm consciência de que diante de um obstáculo devem frear e não acelerar, assim também deve ser com a IA. “Temos de saber quando ’frear’ e quando ’acelerar’ a inteligência artificial” (Kaufman, 2022, p. 20).

Dentro do Lensa, em ambos os casos, tanto no uso de dados faciais dos usuários, quanto no uso de dados criativos captados em rede para a geração dos avatares, tem-se implicações de natureza ética. Nesse sentido, importa destacar que a existência do “dado inovador resultante” que são os *Magic Avatars*, responsáveis pela alta monetização do produto Lensa para além dos típicos filtros, está condicionada à existência, ao tratamento e à manipulação de outros dois conjuntos de dados: os faciais dos usuários e os criativos disponíveis na internet. Somente a partir destas duas entradas – dados faciais e criativos – é possível que o app retorne a saída na forma de *Magic Avatars*.

A partir de tais considerações e tendo por base os dez princípios promovidos pela UNESCO para uma IA ética, pode-se inferir que o Lensa apresenta pendências. Relizando uma análise comparativa do produto a partir das informações disponíveis para os usuários e tendo como referência cada um dos dez princípios da UNESCO, tais pendências são apresentadas a seguir.

1º Princípio – Proporcionalidade e não causar dano: a política de privacidade do app precisa apresentar com maior clareza se o produto de fato atua coletando dados de maneira proporcional ao que precisa/se propõe e a quais danos as pessoas usuárias e designers/artistas digitais estariam submetidas, disponibilizando inclusve suporte e orietações para que estes busquem seus direitos e/ou reparação/compensação.

2º Princípio – Segurança e proteção: embora a política de privacidade deixe claro que zela pela segurança e proteção dos dados, a mesma política afirma que não tem conduções de garantir 100% da segurança e se isenta em ser implicada nas diferentes consequências que podem advir com o uso do produto. Além disso, o modo como o produto opera e as repercussões na sociedade a partir do uso de dados variados e de maneira questionável do ponto de vista ético abalam a percepção coletiva do real comprometimento com este princípio, transformando-o em uma pendência a ser resolvida.

3º Princípio – Justiça e não discriminação: em sua política de privacidade, o Lensa não aborda este tópico, contudo, existem relatos públicos (Hunter, 2022) que identificam a existência de falhas e distorções nos Avatares produzidos pelo Lensa para as pessoas que não são brancas, o que reforça as preocupações sobre a equidade étnica na geração de conteúdos pela IA. Também a geração de Avatares femininos frequentemente sexualizados suscita preocupações dado o caráter machista e de objetificação da mulher que estes resultados trazem (Hunter, 2022).

4º Princípio – Sustentabilidade: não foram encontrados dados em relação a esse princípio nas informações públicas do Lensa, um ponto que também merece atenção, uma vez que o poder computacional para geração destas imagens em alto volume demanda um grande uso de recursos energéticos nos *data centers* responsáveis, os quais ocupam muito espaço e consumem muita energia durante o treinamento das máquinas

5º Princípio – Direito à privacidade e proteção de dados: a maneira inconsistente como o Lensa lida com este princípio foi a principal discussão ao longo desta pesquisa juntamente com os questionamentos relativos a direitos autorais. Verificou-se que a política de privacidade foi atualizada diversas vezes desde que o *app* foi lançado, entretanto, para fundamentar a análise aqui realizada, foi tomada como referência sua última atualização. A diferença entre ela e as anteriores reside no fato de que foi ao longo do tempo que as novas versões passaram a abordar a produção dos *Magic Avatars,* os direitos das pessoas sobre eles e incluíram a violação de direitos autorais. Mas tais esclarecimentos só surgiram após a repercussão na mídia quanto a insegurança referente ao uso do produto. Ainda assim, ela não é totalmente transparente e em muitos pontos isenta a empresa de responsabilização ou prestação de contas.

Verifica-se, dessa maneira, que muitas das preocupações de ordem ética que deveriam ter sido norteadoras desde o princípio do projeto do produto – e, portanto, constar de sua política de privacidade já a primeira versão –, só foram abordadas, atendidas ou esclarecidas em parte e depois da ampla repercussão na mídia, tanto sobre a insegurança em relação ao tratamento dos dados faciais dos usuários, quanto em relação ao uso indevido de obras de artistas digitais para monetização do produto. Cabe destacar que esclarecimentos sobre a questão da captação de artes em rede para treinamento da IA sem consentimento ou compensação dos artistas digitais ainda não estavam presentes na documentação pública do Lensa até o fechamento deste artigo, tornando-se assim outra pendência a ser destacada.

6º Princípio – Supervisão humana e determinação: este princípio visa garantir que seja sempre possível atribuir responsabilidade ética e legal em qualquer estágio do ciclo de vida dos sistemas de IA sendo que a referida supervisão humana pode ser pública também, originada na sociedade. Embora a política de privacidade informe meios de contato com a empresa para diferentes solicitações que possam partir do usuário referente aos seus dados ou privacidade, até o momento do fechamento deste artigo não constava na documentação um nome ou razão social como responsável ético e legal para questões envolvendo o Lensa – há apenas o contato do oficial de proteção de dados designado. Também não existe um item que esclareça em quais condições o usuário teria direito de acionar juridicamente a empresa ou seu

representante legal se assim necessitasse. Fica subentendido, então, que um processo dessa natureza envolveria, como tradicionalmente ocorre, o binômio cidadão versus empresa. No caso, usuário versus Prisma Labs, a proprietária e responsável pelo Lensa.

Importa, contudo, ressaltar que ao mesmo tempo em que houve ampla aceitação do produto, ocorreu também um amplo movimento de supervisão humana que a própria sociedade engendrou diante dos questionamentos públicos que surgiram em relação ao produto e o uso de dados. Esse foi um posicionamento importante que deflagrou a pressão necessária para que a empresa se manifestasse, muito embora ainda existam pendências referentes aos direitos e ao uso das obras dos artistas digitais.

O caso do Lensa demonstrou que o sexto princípio da UNESCO será essencial para garantir a autonomia e autopreservação humanas quando em contato com a IA, conforme as produções que envolvem essa tecnologia sejam ampliadas. Este princípio permitirá que a partir de uma consciência crítica de uso, as próprias pessoas possam questionar e, se necessário, acionar juridicamente os responsáveis na busca por salvaguardar os seus direitos.

7º Princípio – Transparência e explicabilidade: este princípio destaca a importância das soluções em IA serem transparentes para os envolvidos e também explicáveis. O presente artigo demonstrou que ambos os pontos, tanto transparência quanto explicabilidade, estiveram em certa medida negligenciados na trajetória do Lensa enquanto produto digital. Embora existam a política de privacidade e os termos de uso, nos quais ficam evidentes os esforços da empresa em esclarecer os diferentes âmbitos que suportam o produto e o que ele oferece para os usuários – também apresentando direitos e deveres das partes envolvidas –, muitos pontos essenciais não foram tratados de maneira transparente e tampouco foram adequadamente explicados. Essa postura acaba refletindo nos questionamentos e polêmicas envolvendo o uso de dados, o qual é a base de produção da principal inovação que o produto traz – para além de aplicar filtros em fotos – que são os *Magic Avatars*. O uso de dados criativos produzidos por designers e artistas digitais sem consentimento ou compensação persistiu – até o fechamento deste artigo – como um ponto nevrálgico que ainda carece ser não só explicado, mas de maneira transparente.

8º Princípio – Responsabilidade e prestação de contas: segundo este princípio os atores de IA e os Estados-membro devem assumir responsabilidades éticas e jurídicas de acordo com o direito nacional e internacional. No caso do Lensa, a pressão pública é o que tem gerado as discussões de ordem ética presentes na mídia e nas redes sociais. Entretanto, até o momento do fechamento deste artigo, não foram encontradas informações públicas sobre processos questionando juridicamente a empresa responsável pelo produto. De todo modo, não se pode descartar completamente a possibilidade de isso estar ocorrendo ou vir a ocorrer.

Nesse sentido, destaca-se que a política de privacidade e os termos de uso do *app* redigidos como estão, são muito marcados pela isenção da empresa frente às externalidades negativas do seu produto e pela transferência de responsabilidade total dos riscos para as mãos das pessoas usuárias, além de coibir movimentações judiciais de ordem coletiva por parte delas.

Independentemente da existência de processos jurídicos em curso, o caso do Lensa promove reflexões sobre como as pessoas e governos poderão ou deverão se posicionar juridicamente diante de produtos envolvendo a IA e de semelhante tratamento de dados. Até o momento, o caminho pelo qual a Prisma Labs presta contas em relação ao Lensa é a atualização da política de privacidade do produto e dos termos de uso, contudo, esse é um procedimento tradicional para todo e qualquer produto digital. Ainda não há, portanto, uma prestação de contas oficial com notas à imprensa ou mesmo em âmbito jurídico. Isso poderá ser alterado conforme leis regulatórias forem sendo implementadas nos países.

9º Princípio – Conscientização e alfabetização: tal princípio refere-se à importância e à necessidade de se promover a compreensão pública das tecnologias de IA e do valor dos dados. O que se dá por meio da educação aberta e acessível, engajamento cívico, habilidades digitais e treinamento em ética. O sucesso do Lensa junto ao público e, ao mesmo tempo, o questionamento por parte da sociedade dos pontos críticos do produto, demonstrou o quão importante é para, além de promover o acesso ao mundo digital, garantir que esse acesso seja pautado por um pensamento crítico. Tal garantia viabiliza que se possa promover um uso consciente desses artefatos ponderando riscos, benefícios e promovendo a autonomia necessária para que, quando preciso, as pessoas afetadas possam se organizar e exigir seus direitos.

A criticidade para um uso consciente e emancipador dos projetos que envolvem as técnicas de IA somente poderão ser promovidas por meio da educação. Sobre o Lensa, não foi encontrada iniciativa por parte da empresa responsável pelo produto em educar o público usuário sobre ele, a qual apresente os benefícios e também os riscos inerentes ao uso do *app*. Com isso, o nono princípio torna-se, também, um fator pendente.

10º Princípio – Governança e colaboração adaptáveis e com múltiplas partes interessadas: este princípio alega que o direito internacional e a soberania nacional devem ser respeitados no uso de dados e que os Estados podem regulamentá-los. No caso do Lensa, sua política de privacidade prevê a posibilidade dos usuários invocarem as leis que oferecem direitos legais sobre seus dados pessoais, as quais têm origem nas leis de proteção de dados vigentes na Europa e nos EUA. Mas não menciona especificidades de outras regiões, deixando este aspecto também pendente.

Esta breve análise documental de produto – sem qualquer pretensão de domínio dos termos da lei ou valor jurídico, mas dedicada ao design da experiência do usuário e com foco na ética – permitiu verificar que ainda existem lacunas a serem preenchidas pela Prisma Labs no que se refere principalmente à transparência e explicabilidade do produto Lensa para o público.

Concluída a análise, pode-se inferir que é possível utilizar os dez princípios éticos elaborados pela UNESCO como uma lista de verificação para produtos que envolvam a IA, inclusive podendo gerar estruturas de protocolos de análise para outros produtos, com potencial para se tornar um instrumento a ser utilizado tanto pelos usuários, quanto pelos designers.

Para os usuários, os dez princípios podem servir como uma referência mínima de transparência e qualidade para se aceitar ou rejeitar um produto que contenha a IA. Para os designers, podem servir como uma lista de requisitos de projeto que atenda aos princípios éticos mais elementares desde sua fase inicial.

Portanto, há a necessidade de promover constantemente uma reflexão crítica sobre a ética das soluções de design que contenham IA e que são – e serão – acessadas pelos usuários na medida em que produções dessa natureza só são possíveis mediante o amplo uso de dados. Para Zuboff (2019) os dados sobre os usuários passaram a ser essenciais e a compor o cenário para o que a autora denomina Era do Capitalismo de Vigilância. Nesse conceito, o capitalismo como se conhece se transformou em “uma nova ordem econômica que reivindica a experiência humana como matéria-prima gratuita para práticas

comerciais dissimuladas de extração, previsão e vendas" (Zuboff, 2019).

Dentro desse cenário, muitos são os modelos de negócios que pautam seus desenvolvimentos com base na captação de dados dos usuários, utilizando-os como matéria-prima para a aplicação em seus produtos sob a justificativa de prever comportamentos e aprimorar experiências. Tais dados captados, quando empregados de modo adequado e ético podem gerar valor percebido para os usuários. Contudo, na dinâmica atual dos negócios, dados são, geralmente, captados em maior volume do que o necessário e de de maneira pervasiva – com os meios de informática distribuídos perceptível e imperceptível no ambiente – e ubíqua, tornando a interação do ser humano com a computação invisível ao integrar a informática aos comportamentos naturais humanos (Zuboff, 2019).

Esses grandes volumes de dados captados em sociedade, quando não utilizados em sua totalidade, gera um excedente, o qual Zuboff (2019) denomina de "excedentes comportamentais". Esses dados se tornaram um tipo de matéria-prima que pode ser negociado entre diferentes agentes no atual modelo de capitalismo.

A preocupação com os rumos que a captação em massa de dados sensíveis, seus desdobramentos e possíveis impactos sociais foi, inclusive, a justificativa da IBM para se retirar do cenário de negócios e de pesquisas voltados para a tecnologia de reconhecimento facial no ano de 2020 (Reuters, 2020).

Na ocasião, casos de violência policial claramente motivados por racismo foram algumas das razões que fizeram o então CEO da empresa Arvind Krishna afirmar, em carta enviada ao Congresso Norte Americano, que não ofereceria mais software de análise ou reconhecimento facial de uso geral e que, além disso, a empresa se oporia a qualquer uso de tal tecnologia para fins de vigilância em massa e criação de perfis raciais.

Na carta, o CEO afirmou que a IBM gostaria de trabalhar com o Congresso Norte Amreicano na busca por justiça e equidade racial e que propunha essa parceria com foco incialmente em três áreas políticas consideradas como principais: a reforma da polícia, o uso responsável da tecnologia e a ampliação de habilidades e oportunidades educacionais (Krishna, 2020). No mesmo documento, o CEO afirmou que a empresa não tolerará o uso de qualquer tecnologia, incluindo a de reconhecimento facial, oferecida por outros fornecedores que envovlam "[...] violações de direitos e liberdades humanos básicos ou qualquer finalidade que não seja consistente com nossos valores e princípios de Confiança e Transparência" (Krishna, 2020, tradução nossa).

Krishna destacou ainda a importância do debate coletivo diante dos desafios postos na adoção de tecnologias deste tipo e que este é o tempo de iniciar um diálogo nacional "[...] sobre se e como a tecnologia de reconhecimento facial deve ser empregada pelas agências policiais domésticas" (Krishna, 2020, tradução nossa).

Debates e estudos para regulação da IA vêm ocorrendo no mundo todo, mas até que se estabeleçam marcos regulatórios coerentes, que legislem sobre o tema sem inibir o desenvolvimento, nos quais a IA possa ser utilizada como recurso para cocriação sem infringir direitos autorais e que permitam a responsabilização no uso indevido de dados e conteúdos, algumas alternativas para proteção do uso de imagens surgiram recentemente, como a *Nightshade* e a *PhotoGuard*.

A primeira iniciativa, *Nightshade* "envenena" as IA generativas de Arte como o *Dall-E, Midjourney* e *Stable Diffusion* – que operam de maneira semelhante ao Lensa. De acordo com o *MIT Technology Review*, o professor e pesquisador de ciência da computação da Universidade de Chicago, Ben Zhao, esclarece que há uma total assimetria de poder entre as empresas de tecnologia e os artistas, por isso elaborou o *Nightshade* com a intenção de mudar essa dinâmica de poder. Segundo a reportagem, "algumas empresas, como a OpenAI e a Stability AI, se ofereceram para permitir que os artistas optem por não participar dos conjuntos de treinamento ou disseram que respeitarão as solicitações para que seu trabalho não seja extraído. No entanto, no momento, não há nenhum mecanismo para forçar as empresas a cumprirem sua palavra. Zhao diz que o Nightshade poderia ser esse mecanismo. É extremamente caro criar e treinar modelos de IA generativos e pode ser muito arriscado para as empresas de tecnologia extraírem dados que poderiam quebrar suas joias da coroa" (MIT [...], 2023). O *Nightshade* funciona fazendo alterações sutis nos pixels da imagem, que são invisíveis ao olho humano, mas que: "[...] enganam os modelos de aprendizado de máquina, fazendo-os pensar que a imagem representa algo diferente do que realmente representa. Quando os artistas aplicam isso em seus trabalhos e essas imagens são coletadas como dados de treinamento, esses 'pixels envenenados' entram no conjunto de dados do modelo de IA e causam o mau funcionamento do modelo. Imagens de cachorros se tornam gatos, chapéus se tornam torradeiras, carros se tornam vacas. Os resultados são realmente impressionantes, e atualmente não há defesa conhecida" (*Id.*, 2023).

A segunda iniciativa, o *PhotoGuard*, foi desenvolvido por pesquisadores do MIT CSAIL – Computer Science and Artificial Intelligence Laboratory – os quais reconhecem a capacidade e o potencial das IA generativas e seus riscos: "à medida que entramos numa nova era em que as tecnologias alimentadas pela inteligência artificial podem criar e manipular imagens com uma precisão que confunde a linha entre a realidade e a fabricação, o espectro do uso indevido torna-se cada vez maior" (Gordon, 2023, tradução nossa). De acordo com a autora, o *PhotoGuard* é "uma técnica que usa perturbações – alterações minúsculas em valores de pixels invisíveis ao olho humano, mas detectáveis por modelos de computador – que efetivamente atrapalha a capacidade do modelo de manipular a imagem" (*Ibid*., tradução nossa). Essa proteção torna quase impossível que as imagens de pessoas captadas na internet possam ser manipuladas para criar as informações falsas, por exemplo.

Ambos os desenvolvimentos mostram que as Universidades e Institutos de pesquisa independentes contribuem de maneira importante para o debate, pesquisando e trazendo soluções alternativas para outras necessidades que surgem como consequência do uso da IA.

Atualmente, pode-se verificar que, conforme Cerejo e Carvalhais (2019, p. 367) a tecnologia da IA está desempenhando um papel significativo em moldar o futuro do Design, compondo recursos e pontos de contato novos e complexos a serem implementados em projetos. É preciso salientar que nesse processo, e como consequência dele, as questões de ordem ética serão primordiais para que projetos de design estejam adequados ao acesso e uso humano. Assim, entendendo que a produção de projetos que envolvem IA pode demandar a captação de dados para treinar os modelos de aprendizado de máquina, é possível afirmar que preocupações de natureza ética e seus impactos sociais devem ser parâmetros mandatórios para qualquer projeto envolvendo essa tecnologia, além de compor o cenário das pesquisas em design – área responsável por humanizar as tecnologias.

Os desdobramentos da captação, do armazenamento e da manipulação de dados das pessoas usuárias como ferramentas para a IA, bem como os impactos sociais dos produtos baseados nessa tecnologia, são temas sensíveis e que vêm sendo tratados

em diversas áreas do conhecimento para além da ciência da computação, matemática, engenharia e ciências cognitivas (Cortiz, 2019; Taulli, 2020). Todavia, ainda carecem de maior atenção no campo do design, que é justamente responsável pelo projeto de produtos, incluindo os digitais, como o Lensa. Cabe então a reflexão sobre como o design pode contribuir para ponderar os riscos e benefícios envolvendo a aplicação da IA e de que maneira garantir que a produção seja o mais ética possível. Nesse sentido, as diretrizes apontadas pela UNESCO no relatório aqui apresentado podem indicar um caminho e auxiliar a balizar políticas públicas que salvaguardem os direitos das pessoas.

5. CONSIDERAÇÕES FINAIS

A Inteligência Artificial já faz parte de muitas das produções do design voltadas aos seres humanos. Ao passo que essas tecnologias avançam, avançam também as preocupações derivadas do acesso e do uso das produções resultantes delas, especialmente quando trazem impactos do ponto de vista ético. Em uma sociedade na qual os dados das pessoas usuárias são captados e tratados em grande volume, cresce a preocupação sobre como esses estão sendo utilizados em conjunto com a IA.

Diante desse cenário, as preocupações do ponto de vista ético atreladas ao uso da IA fazem parte da agenda atual de diferentes órgãos, instituições e governos. Toda e qualquer produção tecnológica que preveja o acesso e o uso por parte de seres humanos deve ser pautada por aspectos éticos, assegurando que os direitos fundamentais e a dignidade humana sejam respeitados. Foi nesse sentido que a UNESCO elaborou um primeiro documento oficial com recomendações éticas para o emprego da IA por parte dos Estados-membro.

Tais recomendações prevêem ações em diferentes âmbitos da vida social nas quais se pode prever o contato – atual e futuro – de seres humanos com a Inteligência Artificial, bem como a interação entre ambos, ponderando-se riscos e benefícios.

Longe de ser um documento acabado, o texto da UNESCO é o marco inicial de um acompanhamento que deve ser constante e que certamente provocará atualizações em suas recomendações a cada novo salto tecnologico da IA com impacto na sociedade. E, também, em cada nova instância e inter-relação formada entre a IA e a sociedade por meio de projetos de produtos, sistemas ou serviços, uma tarefa que cabe também ao design.

O presente artigo tomou o produto Lensa como exemplo de impacto social do emprego da IA. As produções do aplicativo criado pela Prisma Labs em 2018 foram aceitas e utilizadas em massa no ano de 2022, sendo o *app* disponibilizado até os dias atuais e, neste processo, não só foram captados milhões de dados, como certamente produziu-se outros milhões de novos dados a partir deles.

Embora a política de privaciadade do produto receba atualizações constantes para adequação às leis, é preocupante que princípios básicos e elementares como segurança e proteção de dados, bem como a transparência na captação e explicabilidade do uso destes dados, somente tenham figurado na política de privacidade em data posterior à notícias que discutiam o uso dos dados faciais dos usuários, bem como sobre as polêmicas envolvendo artistas digitais e a captação de suas artes em rede.

Neste último caso em específico, dois pontos importantes devem ser destacados e trazer reflexões. O primeiro deles, diz respeito a como dados pessoais sensíveis – inclusive no nível extremamente sensível de mapeamento facial – têm sido tratados e como as pessoas usuárias aparentemente ainda não têm uma compreensão clara dos riscos aos quais podem estar sendo submetidos, inclusive pagando para entregar os próprios dados, em razão do poder de engajamento dos *Magic Avatars*.

O segundo deles, refere-se a quais dinâmicas podem se estabelecer quando a IA passa a ser treinada a partir de produções da criatividade humana. Nesse âmbito, em que ainda são obscuros tanto os riscos quanto as possibilidades, um fato se destaca: a necessidade de desenvolver diretrizes que pautem as produções por IA de modo transparente, bem como o reconhecimento e a compensação financeira dos envolvidos.

Chama a atenção que o juízo da própria sociedade deflagrou o debate público sobre os riscos envolvidos no uso do produto, sendo que questionamentos partiram de diferentes atores sociais, seja na figura de pessoas usuárias, organizações sociais ou mesmo profissionais que foram afetados pelos impactos da dinâmica que se estabeleceu na popularização de uso do Lensa. O debate público sobre os aspectos éticos do produto forçou a empresa a se posicionar e revisar sua política de privacidade, que ainda necessita esclarecer os pontos pendentes já mencionados.

Esse efeito, causado pela reação da sociedade mediante o impacto do Lensa, trouxe à luz a importância de uma "alfabetização digital" das pessoas usuárias, para que possam adquirir cada vez mais conhecimento sobre os direitos dos seres humanos diante do mundo o digital e da IA aplicada em projeto de produtos e dos riscos a que podem estar submetidos. Essa alfabetização será essencial para fomentar a consciência crítica da coletividade na adoção e no uso de produtos, sistemas e serviços baseados em IA.

As iniciativas de regulação da IA que vêm ocorrendo em diversos países serão também importantes pontos de apoio e referência para o desenvolvimento de políticas púlbicas, tanto para as empresas produzirem soluções em IA atentando a aspectos éticos, quanto para as pessoas poderem utilizar estas produções de maneira segura. Contudo, sabe-se que quando uma solução alcança a sociedade e passa a ser utilizada em massa, externalidades negativas podem ocorrer. Nesse sentido, acompanhar o que ocorre como consequência do contato e do uso de tais soluções será crucial e é nesse ponto que a transparência, a explicabilidade e a responsabilização precisarão ser claros.

Também as tecnologias alternativas, como *Nightshade* e *PhotoGuard*, vêm surgindo e movimentando o cenário dos desenvolvimentos e das possibilidades em IA, trazendo para integrar o debate o fator de "enfrentamento" quando alguma tecnologia parece, em algum sentido, exploratória. Esses desenvolvimentos alternativos demonstram que um dos papéis das Universidades e centros de pesquisa científica independentes reside em "equilibrar a balança" de poder frente à assimetria das grandes empresas de tecnologia – também conhecidas como *big techs* – e as pessoas usuárias.

Em casos como o do Lensa – e em outros tantos que podem surgir –, o campo do design, enquanto área criadora de soluções para a humanidade na forma de projetos, deve estar atento, acompanhando e desenvolvendo tanto produções nos âmbitos teórico e crítico, problematizando a realidade social impactada por essas produções, quanto no âmbito da prática de projeto, momento no qual os projetos são criados e desenvolvidos.

Manter a sinergia entre os dois domínios, o teórico e o prático, é o que vai permitir que o design consiga identificar as dinâmicas que podem se estabelecer na sociedade a partir das interações entre seres humanos e projetos que envolvem a IA. Para que assim seja possível produzir propostas teóricas emancipatórias, zelando pela ética com poder de influenciar a prática, bem como produzir práticas mais sensíveis que possam retroalimentar a teoria – seja a partir de novas perspectivas,

captadas em sociedade, ou mesmo com caráter de denúncia dos aspectos resultantes do uso.

Mais do que docentes, pesquisadores e projetistas, os profissionais do campo do design na era da IA deverão ser vigilantes e zelar pela ética no projeto, entendendo que somente assim será possível garantir que o propósito da profissão seja cumprido: o de melhorar continuamente a vida do ser humano.

REFERÊNCIAS

1. ADG. Associação dos Designers Gráficos. **Código de ética profissional do designer gráfico**. [200-]. Disponível em: https://adg.org.br/novosite/wp-content/uploads/2020/07/ADGBrasil_CodigoEtica.pdf. Acesso em: 29 jan. 2024.
2. ADP. Associação dos Designers de Produto. **Código de ética profissional do design de produto**. [2004]. Disponível em: https://turmadod.com/alunos/downloads/4s2010_2/etica_legislacao/Codigo_de_Etica_Design_Produto.pdf. Acesso em: 29 jan. 2024.
3. ALVES, J. A. S.; GUIMARÃES, R. V. S. A. (org.). **Direitos humanos e a ética na era da inteligência artificial.** Indaiatuba, SP: Editora Foco, 2023. 232 p.; *E-book*.
4. AMUNDSEN, E. M. Home page in ArtStation. Disponível em: https://www.artstation.com/mischeviouslittleelf. Acesso em: 17 dez. 2022.
5. ARTSTATION. **Community**. Disponível em: https://www.artstation.com/?sort_by=community. Acesso em: 17 dez. 2022.
6. BODEN, M. **Inteligência artificial**: uma brevíssima introdução. São Paulo: Editora UNESP, 2020.
7. BONSIEPE, G. **Design como prática de projeto**. São Paulo: Blucher, 2012.
8. 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. 2021.
9. BUCHANAN, B. G. A (very) brief history of artificial intelligence. **AI Magazine**, v. 26, n. 4, p. 53–60, 15 dez. 2005. Disponível em: https://ojs.aaai.org//index.php/aimagazine/article/view/1848. Acesso em: 4 abr. 2021.
10. CEREJO, J.; CARVALHAIS, M. **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 abr. 2021.
11. CELLARD, A. A análise documental. *In*: POUPART, Jean *et al*. **A pesquisa qualitativa**: enfoques epistemológicos e metodológicos. Petrópolis: Vozes, 2012.
12. CORTIZ, D. **O design pode ajudar na construção de inteligência artificial humanística?** Anais do 17° Ergodesign – Congresso Internacional de Ergonomia e Usabilidade de Interfaces Humano Tecnológica: Produto, Informações Ambientes Construídos e Transporte 17° USIHC – Congresso Internacional de Ergonomia e Usabilidade de Interfaces Humano Computador. 2019. Disponível em: http://pdf.blucher.com.br.s3-sa-east-1.amazonaws.com/designproceedings/ergodesign2019/1.02.pdf. Acesso em: 8 abr. 2021.
13. DIGNUM, V. **Responsible artificial intelligence**: how to develop and use AI in a responsible way. Artificial Intelligence: Foundations, Theory, and Algorithms. Springer, 2019.
14. FLUSSER, V. **O mundo codificado**: por uma filosofia do design e da comunicação. São Paulo: Cosac Naify, 2007.
15. GIL, A. C. **Métodos e técnicas de pesquisa social**. São Paulo: Atlas, 2008.
16. GORDON, R. Using AI to protect against AI image manipulation: "PhotoGuard," developed by MIT CSAIL researchers, prevents unauthorized image manipulation, safeguarding authenticity in the era of advanced generative models. **MIT News**. 31 jul. 2023. Disponível em: https://news.mit.edu/2023/using-ai-protect-against-ai-image-manipulation-0731. Acesso em: 10 dez. 2023.
17. GRILO, A.; VAN DER LINDEN, J. C. S.; RIBEIRO, V. G. Competências e ética do designer nas tecnologias da informação: a deontologia de design implícita em textos de profissionais de UX. **Design em Questão**, v. 2, n. 3, p. 6-20, jul. 2022. Disponível em: https://design.revistas.ufcg.edu.br/index.php/designemquestao/article/view/66. Acesso em: 29 jan. 2024.
18. HUNTER, T. AI selfies – and their critics – are taking the internet by storm. Ethical Issues. **The Washington Post**. Washington D.C., 8 dez. 2022. Disponível em: https://www.washingtonpost.com/technology/2022/12/08/lensa-ai-portraits/. Acesso em: 28 dez. 2022.
19. JAVAID, M. The Magic Avatar you paid $3.99 for is probably stolen, artists say. **The Washington Post**. Washington D.C., 9 dez. 2022. Disponível em: https://www.washingtonpost.com/technology/2022/12/09/lensa-apps-magic-avatars-ai-stolen-data-compromised-ethics/. Acesso em 28 dez. 2022.
20. JOHNSON, S. **Cultura da interface**: como o computador transforma nossa maneira de criar e comunicar. Rio de Janeiro: Zahar, 2001.
21. KAUFMAN, D. **Desmistificando a inteligência artificial**. Belo Horizonte: Autêntica, 2022.
22. KRISHNA, A. **IBM CEO's letter to congress on racial justice reform**. International Business Machines – IBM. Disponível em: https://www.ibm.com/policy/facial-recognition-sunset-racial-justice-reforms/. Acesso em: 10 dez. 2022.
23. MIT TECHNOLOGY REVIEW. Nova ferramenta pode dar aos artistas vantagem sobre a IA. **MIT Technology review**. 17 nov. 2023. Disponível em: https://mittechreview.com.br/nova-ferramenta-pode-dar-aos-artistas-vantagem-sobre-a-ia/. Acesso em: 10 dez. 2023.
24. NORMAN, D. **O design do futuro**. Rio de Janeiro: Rocco, 2010.
25. PASSOS, E. S. Tendências da ética profissional na modernidade. **Revista Brasileira de Enfermagem**, Brasília, v. 46, n. 1, p. 56-62, jan./mar. 1993. Disponível em: https://www.scielo.br/j/reben/a/cS5nKnNMfJR8fjB58JwmHKg/?lang=pt. Acesso em: 29 jan. 2024.
26. PEIXOTO, F. H.; SILVA, R. Z. M. **Inteligência artificial e direito.** Curitiba: Alteridade Editora, 2019.
27. PRISMA LABS. **Company**. Disponível em: https://prisma-ai.com/about. Acesso em: 09 dez. 2023a.

28. _____. **Prisma**. Disponível em: https://prisma-ai.com/prisma. Acesso em: 09 dez.2023b.

29. _____. **Lensa**. Disponível em: https://prisma-ai.com/lensa. Acesso em: 09 dez. 2023c.

30. _____. **Magic Avatars**. Disponível em: https://land.prisma-ai.com/magic-avatars/. Acesso em: 09 dez. 2023d.

31. _____. **Privacy policy**. Disponível em: https://lensa.app/privacy. Acesso em: 09 dez. 2023e.

32. _____. **Lensa app terms of use**. Disponível em: https://lensa.app/terms. Acesso em: 10 dez. 2023f.

33. REUTERS. **IBM gets out of facial recognition business**. 2020. Disponível em: https://www.reuters.com/article/ibm-facial-recognition-idUSL4N2DM00L. Acesso em: 4 abr. 2021.

34. SANTAELLA, L. **Inteligência artificial e cultura**: oportunidades e desafios para o sul global. UNESCO. 2021. Disponível em: https://cetic.br/pt/publicacao/inteligencia-artificial-e-cultura-oportunidades-e-desafios-para-o-sul-global/. Acesso em: 02 out. 2022.

35. SANTOS, R. B. Impactos das tecnologias digitais sobre a fruição dos direitos sociais à saúde, à educação e aos benefícios. *In*: ALVES, J. A. S.; GUIMARÃES, R. V. S. A. (org.). **Direitos humanos e a ética na era da inteligência artificial.** Indaiatuba, SP: Editora Foco, 2023. *E-book* (232 p.) p. 76-113.

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

37. TAULLI, T. **Introdução à inteligência artificial**: uma abordagem não técnica. São Paulo: Novatec, 2020.

38. 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.

39. ZUBOFF, S. **A era do capitalismo de vigilância**: a luta por um futuro humano na nova fronteira do poder. Rio de Janeiro: Intrínseca, 2019.

40. YIN, R. K. **Estudo de caso**: planejamento e métodos. Tradução: Daniel Grassi. 2. ed. Porto Alegre: Bookman, 2001.