

## ÉTICA E RESPONSABILIDADE SOCIAL NA ERA DA INTELIGÊNCIA ARTIFICIAL<sup>23</sup>

Gustavo da Rocha Machado<sup>24</sup>

### Resumo

Este artigo discute os desafios éticos e as implicações da responsabilidade social na era da IA, contextualizando-os nas esferas política, social, econômica, tecnológica e ambiental. Parte-se de fundamentos filosóficos clássicos, como Sócrates, Platão e Aristóteles, para refletir sobre dilemas contemporâneos intensificados pela IA generativa, *deepfakes*, automação e economia da atenção. Busca-se analisar tanto os riscos — desinformação, manipulação algorítmica, vieses, perda de privacidade, impactos ambientais, precarização do trabalho e fadiga cognitiva — quanto as oportunidades, como avanços na saúde, educação, inclusão produtiva, mobilidade urbana e sustentabilidade. O papel da IA na mediação de decisões morais é problematizado, questionando seus limites na substituição ou apoio à reflexão ética humana. São abordadas questões como regulação (LGPD, AI Act), *compliance*, governança corporativa, ESG, modelos de negócios éticos, design centrado na vida (*life-centered design*) e metodologias de análise de impacto social e ambiental. Destaca-se a necessidade de aprendizagem contínua (*upskilling* e *reskilling*), comitês de ética, transparência algorítmica e construção colaborativa de futuros desejáveis por meio de *foresight* estratégico. Conclui-se que a tecnologia é neutra em essência, podendo potencializar tanto benefícios quanto danos. Assim, a ética na era da IA deve ser compreendida como prática cotidiana e responsabilidade compartilhada entre indivíduos, empresas e Estado, exigindo pensamento crítico, regulação equilibrada, inclusão social e compromisso com o bem-estar coletivo e a sustentabilidade planetária.

**Palavras-chave:** ética digital; inteligência artificial; responsabilidade social.

Mais do que nunca, a temática da ética e da responsabilidade social corporativa é superimportante, e estamos vivenciando um momento, agora, no século XXI, de grande complexidade em diversas esferas. Na política, temos uma série de elementos, como a geopolítica. Na esfera social, estamos enfrentando grandes mudanças por conta da inteligência artificial (IA). Na tecnológica, a cada semana temos alguma novidade em IA. Na esfera econômica, também há grandes impactos por conta dessa nova tecnologia. E na ambiental, existe tanto a possibilidade da IA poder nos munir com um

---

<sup>23</sup> Texto adaptado da palestra homônima, integrante da Semana de Responsabilidade Social de 2025, do Centro Universitário Padre Anchieta (UniAnchieta), cuja gravação em vídeo está disponível em: <https://www.youtube.com/watch?v=LmXSXvOGQNC>. A palestra foi mediada pela prof. Antonio Valini e pela aluna Jéssica Rodrigues Rosa.

<sup>24</sup> Mestre em Design pela Universidade de Miami. Graduado em Comunicação Social pela PUC-Campinas.

entendimento melhor das questões climáticas, mas, por outro lado, os computadores também demandam grandes *data centers*, utilizando muitos recursos hídricos e energia elétrica. É muito importante podermos trazer um pouco dessas temáticas.

Na Grécia Antiga, Platão, Aristóteles e Sócrates abordavam dilemas que são também contemporâneos. Percebemos, então, que uma forma de prevermos o futuro é olharmos para as obras de ficção científica do passado. Tudo está se tornando cada vez mais presente, e a ficção científica está se tornando o nosso dia a dia. Temos como exemplo o filme *Her* (2013), com o Joaquim Phoenix. Assim como na obra, hoje, já existe uma tendência de as pessoas terem relacionamentos com a IA. Isso ocorre mesmo com a IA desmaterializada; imagine quando ela estiver um corpo robótico!

O objetivo desta apresentação é explorar esses dilemas e desafios, nos quais não existem respostas fáceis. Além disso, abordaremos muitos questionamentos e provocações, mas não de fato respostas, porque precisamos realmente lidar com essas questões em grupos e debater mais profundamente esses temas, examinar soluções, propor os *frameworks* e as novas metodologias. Um exemplo é a análise SWOT (*strengths, weaknesses, opportunities e threats*), por meio da qual olhamos internamente para forças e fraquezas, sendo aplicada tanto nas empresas quanto pessoalmente; em operacionais e riscos, bem como nos externos do ambiente do mercado. Hoje em dia, não é que essa ferramenta (FOFA, em português) esteja obsoleta, mas ela requer o dinamismo de ser realizada semanalmente, com uma curiosidade muito grande. Isso porque tudo muda a cada semana, levando a inspirações e ações práticas.

Partimos desse conceito da ética como algo que faz parte de uma filosofia e é responsável pela investigação dos princípios que acabam motivando e distorcendo. Veremos também essa questão dos *deepfakes*, com vídeos e imagens cada vez mais realistas, que acabam disciplinando e orientando o nosso comportamento humano. Então, muitas vezes, se fala que a ética está vinculada quando as pessoas não estão olhando os nossos comportamentos, quando não tem ninguém olhando se estamos fazendo algo errado, ou se, por princípios, independentemente de alguém estar olhando ou não, fazemos a coisa certa.

Nesse ponto, há uma provocação: o *ChatGPT* poderia ser o pai da ética moderna? Isso porque Sócrates foi considerado fundador da ética racional, e há essa distinção

entre a IA antiga, que existe por décadas e décadas e é muito voltada à estatística, à matemática, e agora temos a IA generativa, que surgiu com o *ChatGPT* e outras ferramentas como uma visão mais criativa. Mas será que a IA poderia substituir a reflexão moral humana ou contribuir de alguma forma?

Isaac Asimov, um escritor prolífico de ficção científica, escreveu centenas de obras. Ele criou algumas leis da robótica na obra *Eu, Robô* (1950), que existe tanto em livro quanto no cinema, com o Will Smith. É provável que a IA pode realmente colaborar em muitas frentes e muitos dilemas, mas não como uma substitutiva da ética. Hoje, temos o dado surpreendente de que dois terços dos usuários confiam em conselhos éticos dados pela IA, até mesmo utilizando-a como terapeuta e psicóloga. Essa é uma evolução do “Dr. Google”, da tendência comportamental, no passado, do faça você mesmo. Atualmente, temos essa nova evolução da saúde física. Anteriormente, usávamos o Google para pesquisar sobre uma determinada doença; agora, usamos a inteligência generativa, criativa. E é por isso que ela alucina, porque ela é criativa. Essa inteligência não dá a resposta da mesma forma cada vez que você interage com ela. O Dr. Google era aplicado à saúde física; a IA é aplicada à saúde mental.

Entretanto, não podemos ficar só com visões negativas em relação ao futuro. Isso faz muito sucesso entre as pessoas: filmes de guerra, dramas, noticiários, Datena etc. *Black Mirror* (2011), por exemplo, é uma série da Netflix que acaba abordando vários dilemas éticos e muito negativos, como a pessoa ser cancelada no mundo real e ninguém a enxergar mais, da mesma forma que algumas pessoas foram canceladas nas redes sociais, por conta da polarização, e até mesmo parentes acabaram brigando por questões políticas, e assim por diante.

Quando foi lançado o *Veo 3*, do Google, uma ferramenta de vídeo, começaram a surgir diversos vídeos super-realistas, como pessoas conversando na rua, temas dos mais diversos, que parecem reais, mas são pessoas em IA. Inclusive, na semana passada, fiz um *deepfake* do bem. Como vou dar outro curso on-line em outra instituição, também remoto, sobre IA, ao invés de gravar um vídeo de um minuto e meio sobre o curso, eu “me hackeei”, “me clonei”, criei o meu avatar, a minha entidade IA a partir de uma foto estática minha e fiz essa gravação. Então, não precisei gaguejar e simplesmente gravar várias vezes até o vídeo ficar bom. No final das contas, a instituição

achou que ficou criativo esse vídeo. E realmente, tanto o movimento labial quanto o gesticular e a expressão facial estão cada vez mais fidedignos, mais reais.

Por outro lado, se minha mãe umas duas vezes já entrou em contato comigo por receber mensagens de WhatsApp e ficar na dúvida se eram ou não verdadeiras, imaginem o que os golpistas podem fazer com imagens de vídeo. Então, precisamos ficar muito atentos. Mais ou menos no mês passado, o Google também lançou uma ferramenta revolucionária de imagem onde é possível “fazer *selfie*” com pessoas famosas, com quem você quiser, como em um iate com Bill Gates, por exemplo. Ela se chama *Nano Banana*, que tem esse nome engraçado porque decidiram manter o nome usado durante a fase de testes. Com essa ferramenta, conseguimos criar imagens realistas e, a partir delas, cenas de vídeo.

Assim, tendo já uma imagem realista, é meio caminho andado para se gerar um vídeo realista. Isso pode gerar aquela confiança, seja na família, nas instituições, nas empresas ou nas próximas eleições no Brasil e no mundo. Portanto, precisamos, efetivamente, de mecanismos e tecnologias com pensamento crítico, para desconfiar se aquilo é, de fato, real ou não. Por exemplo, o *Sora 2*, ferramenta de vídeo do *ChatGPT*, foi lançado semana passada e tem uma marca d'água que vai se movimentando ao longo do vídeo, tornando um pouco mais difícil de retirá-la. Se nas eleições passadas, nos Estados Unidos e no Brasil, já havia desinformação e *fake news*, infelizmente, daqui a três ou quatro anos, isso só tende a piorar, pois estamos ainda na infância da IA generativa.

Atualmente, há cerca de três quartos da população acreditando, em algum momento, em notícia falsa, seja no WhatsApp ou em outros meios. Infelizmente, as redes sociais lucram com essa questão. Vivenciamos, hoje, a economia da atenção, com o TikTok e a rolagem infinita do Instagram, o que propicia com que as pessoas passem mais tempo naquela plataforma, gerando lucros maiores. Também, não há muito discernimento. Um exemplo é 100ª Corrida São Silvestre, em São Paulo, que esgotou muito rápido, e logo um cambista começou a vendê-la de 320 reais por 500. Houve propaganda no Instagram, mas sem um discernimento por parte da Meta, do Facebook, a respeito desse anunciante, que, na verdade, era um falsário. Pior que lidar com cambista, é lidar com cambista falso.

A percepção da realidade vai sendo distorcida cada vez mais. Hoje, eu posso pegar uma foto minha com a minha esposa correndo num determinado local e pedir para trocar o fundo com o *Nando Banana*, por exemplo. Se estou correndo em Campinas, posso trocar o fundo por outra parte do mundo, seja a Muralha da China ou onde quer que eu queira. Isso é criado com um certo realismo. O problema é que, com o passar do tempo e, eventualmente, com algum declínio cognitivo com o passar da idade, podemos nos confundir em termos de “será mesmo que nós vivenciamos isso nessa foto que está aqui ou não?”. Isso porque, hoje em dia, praticamente ninguém tem mais câmera analógica, é tudo digital.

Temos alguns número maliciosos de *fakes*, como de pornografia (95%), o que é péssimo. Felizmente, surgiu a Lei n. 14.132/21, a partir de 2021, criminalizando isso. Entretanto, a taxa de sucesso de algumas ferramentas de identificação ainda deixa um pouco a desejar; deveríamos chegar pelo menos em uns 80%. Há *deepfake* até mesmo do Tom Cruise com sobrepeso. É possível haver também riscos corporativos, como o CEO de uma empresa, um presidente ou um líder de Estado falando algo que não disse. Então, como tudo na vida, a tecnologia é algo neutro, que pode ser utilizada tanto de forma positiva quanto negativa.

No entretenimento, já existem atores e dublagem em IA. Por outro lado, recentemente, saiu o *AirPod 3*, o fone de ouvido da Apple, que faz tradução simultânea. Se viajamos para outro país, incorporamos ferramentas como o *Zoom*, o *Meet* e o *Teams*, com ação simultânea, abrindo um campo de trabalho para quem é “nômade digital”. É fantástico você poder trabalhar em qualquer parte do mundo. Outros pontos positivos são a preservação histórica, a educação e as simulações. O lado negativo tem as informações falsas da política, a falta financeira, a pornografia, a manipulação de mercado, e assim por diante.

A série na Netflix tem toda a questão da gamificação, mas o que é gamificação? Os elementos presentes nos jogos, na verdade, tratam de psicologia comportamental, em como reconhecer, como recompensar o participante. Inclusive, temos, hoje, um grande problema com as *bets* também. Então, toda essa tecnologia e esse design das redes privilegia exatamente a economia da atenção, o engajamento maximizado e a polarização social. Todo tipo de “treta” também dá audiência. Todo mundo gosta de ver

um “barraco”, seja no Big Brother, seja nas redes sociais. E isso, obviamente, gera impacto na nossa saúde mental, na questão da atenção.

Como essas ferramentas são gratuitas, nossos dados acabam sendo o produto. Conforme uma busca em um determinado produto, aparecem propagandas de anunciantes, seja no Google, seja no Facebook. O documentário *O Dilema das Redes* (2020) é bem interessante sobre isso. O psicólogo Skinner trabalha exatamente com os efeitos e a reação a partir de estímulos. Existem as *bets*, os cassinos digitais, com recompensas variáveis, que ativam ciclos de neurotransmissores, como dopamina. Temos essa economia dos *likes*, com a preocupação de quantas curtidas recebemos no nosso vídeo e na foto postada. Além disso, há todo aquele design persuasivo de rolagem infinita no Instagram e vídeos curtos no TikTok, que possuem uma acessibilidade fácil de até uma criança conseguir, com o dedinho, rolar para o próximo vídeo.

Existem os desafios em relação ao *deepfake*, que envolvem a questão da segurança digital, da *cyber* segurança. Também, houve a interrupção de serviços aéreos, alguns meses atrás, em algumas partes do mundo. Então, a segurança, infelizmente, não é semelhante a uma nova pandemia, não sendo uma questão de quando vai acontecer, pois precisamos nos preparar para isso.

Outro tema importante é sobre os vieses algorítmicos, que envolvem o comprometimento da privacidade. Afinal, se acessamos alguma plataforma gratuitamente, abrimos mão dos nossos dados. É um “toma lá, dá cá”. Se você não compartilha seus dados com o *ChatGPT*, por exemplo, ele não terá o seu histórico e, conseqüentemente, também não será tão assertivo. Então, à medida que ele sabe tudo de você, é possível responder de uma maneira muito mais personalizada e assertiva do que se o histórico não fosse gravado. Há, portanto, uma nova tendência de *brain rot*, da fadiga cognitiva.

Anteriormente, já lidávamos com um “tsunami” de dados, de informações, de estímulos visuais, auditivos e audiovisuais. Agora, com a IA, em uma era ainda mais acelerada do que a própria internet, isso acaba levando a quase que um *burnout*, ou seja, gerando dificuldade de concentração, ansiedade e redução da capacidade crítica. Então, é preciso também fazer um “detox” digital de tempos em tempos, ficar off-line, curtir a natureza e os amigos. Esse equilíbrio é importante na relação homem–máquina.

Quem acaba vivenciando muito isso é a nova geração nativa digital, a geração Alpha. Assim, quem é pai ou tio realmente precisa, de alguma forma, ficar atento e, eventualmente, intervir em algumas medidas. Os jovens ficam presos em telas, celulares e tablets, utilizando-os por muitas horas, sem realmente a supervisão dos pais e responsáveis.

Algumas formas de mitigação são o tempo de uso, o modo noturno com filtro azul, para não atrapalhar no sono, e pausas de notificação. Há também o bem-estar digital, que são ferramentas que monitoram e limitam os aplicativos. Existe a técnica Pomodoro, na qual, a cada hora, você pode tirar 10 minutos de *break*, podendo fazer *sprints* de trabalho, por exemplo. Assim, dedica-se intensamente ao trabalho por 50 minutos, com 10 minutos para “esticar as pernas”, se levantar, alongar, ir ao banheiro e tudo mais. Isso acaba sendo bom para a mente e para a coluna, para o físico.

São bilhões e bilhões de dispositivos conectados, muita gente on-line. Obviamente, temos um *gap*, com países em desenvolvimento que não têm acesso, mas, em geral, esse não é mais o problema. Atualmente, há um problema maior com pessoas que se alimentam mal, principalmente de ultraprocessados, do que com pessoas que passam fome no mundo. O Brasil é muito diverso. Mesmo a cidade de São Paulo, por exemplo, tem regiões completamente diferentes. Imagina, então, as barreiras econômicas, os desafios educacionais e a infraestrutura das regiões do Nordeste, Sul, Sudeste, Centro-Oeste. Não são só Norte e Nordeste que convivem com essas questões.

A privacidade é o novo petróleo. Essas ferramentas entram no modelo de serem gratuitas exatamente porque nós somos os produtos, como na obra *1984*, de George Orwell (1949). As marcas acabam nos vigiando e, às vezes, até controlando, persuadindo e manipulando. O que nos protege disso são as Leis Gerais de Proteção de Dados (no Brasil, LGPD, na Europa, GDPR), que proporcionam alguns avanços, mas, ainda assim, insuficientes. Então, precisamos efetivamente evoluir, seja na proteção de dados, seja na transparência.

Muitas empresas acabam não sendo muito transparentes nas suas práticas, apesar de termos a sigla, recentemente, do ESG (social, governança e meio ambiente). Há a possibilidade, hoje mais do que nunca, de desenvolvimento de modelos de negócios éticos. Então, existe a economia circular, que não olha linearmente (extração, produção e descarte), mas de uma forma que seja possível reutilizar materiais e

recursos, com tecnologias inclusivas e acessíveis. A IA, cada vez mais, vem evoluindo. O *Gemini*, do Google, ainda peca um pouco, pois, às vezes, ele corta a fala da pessoa, mas o *ChatGPT* está muito bom para interface por voz, e qualquer pessoa, de qualquer idade, consegue acessar e interagir com ele.

Com o empreendedorismo ético, há empresas que já consideram o meio ambiente e o social, seja no comércio justo, em cafés, redes ou produtos locais. Inclusive, há um sistema de busca com IA muito bom, o *Perplexity*, que, gratuitamente, tem certos limites diários, mas possui quase que um modelo socrático de interação, porque ele apresenta três ou quatro perguntas de exemplo após a busca do tema, para que seja possível dar continuidade e se aprofundar, se tornando praticamente uma rolagem quase infinita, semelhante ao Instagram.

Em relação aos produtos físicos, não digitais, existe, hoje, a possibilidade de materiais éticos, desde a extração responsável de matéria-prima até produtos biodegradáveis e de manufatura justa. No caso de construções humanas, um exemplo é a Índia e outros países, como a China, que faz a reciclagem de celulares, materiais que podem ser tóxicos. A gestão de resíduos também está vinculada a isso, com o lixo eletrônico.

Assim, existem dois tipos de gatilhos: um é psicológico, o outro é físico. A obsolescência programada faz com que as geladeiras atuais não durem tanto quanto as das nossas avós, que duravam décadas, por exemplo. Já a obsolescência apercebida é a psicológica, como, por exemplo, a diferença física entre um monitor de cristal líquido, de LCD, em relação àquele antigo de CRT, que era enorme, fazendo com que a pessoa fique tentada, por uma questão de design e estética, a adquirir um produto novo.

Recentemente, descobriram que ao dizer “obrigado” e “por favor”, ser gentil, gera, por parte da IA, curiosamente, respostas melhores. Entretanto, se começamos a acrescentar mais palavras na nossa interação com a IA, dizendo “por favor” e “obrigado”, gastaremos mais *tokens*, mais água, mais energia elétrica. Então, existe esse dilema: vamos gastar mais água e energia elétrica em benefício próprio para receber uma resposta melhor do *GPT* ou do *Gemini*, ou não e talvez receber uma resposta não tão boa, mas economizar recursos em relação ao meio ambiente? Estamos vivenciando essa singularidade, uma velocidade de mudança tecnológica sem precedentes.

A IA é uma das tecnologias exponenciais, mas se conectamos IA com *big data*, com impressão 3D, com sensores de internet das coisas nas cidades e com computação quântica, vai ser um outro grande salto de uma tecnologia que, nos próximos anos, vai se tornar mais popular. Serão computadores muito mais potentes, poderosos e velozes, se comparados aos atuais. Tudo isso só tende a acelerar e, obviamente, teremos mais paradigmas éticos. Então, as empresas e os profissionais precisam se adaptar a essas mudanças.

Uma das formas é por meio do *upskilling*, que é a atualização profissional. Então, como fica a atividade de um advogado, agora, na era da IA? Ou de um médico veterinário, um educador, um psicólogo etc.? Existe atualização profissional. Para ex-alunos, é comum haver esse tipo de serviço, de forma rápida, especificamente sobre IA. Outra maneira é pelo *skilling*, que é a requalificação profissional. Assim, num caso mais radical e disruptivo, há a necessidade das pessoas se requalificarem, eventualmente, até mudarem de carreira. Realmente, é imperativo trazer esses questionamentos, tanto no meio acadêmico, nas instituições, como os próprios alunos, formandos e profissionais, em relação aos fundamentos morais, pois essa responsabilidade é compartilhada. Dessa forma, para recém-graduados, seria interessante que houvesse uma política de, por exemplo, desconto para fazer uma atualização profissional junto à IA, dentro da área de formação.

Algumas estratégias para as empresas responsáveis seriam: formação de comitês de ética; *compliance*; análise de impactos sociais e ambientais de determinadas tecnologias e estratégias; capacitação contínua. Essa nova era da aprendizagem é para a vida toda. Não é o diploma que finalmente define se você encerrou a carreira, mas, mais do que nunca, precisamos de capacitação contínua. O futurista Alvin Toffler (1980), no século XX, falou que, no século XXI, o maior desafio para todos os profissionais seria aprender, desaprender e reaprender o tempo todo. Estamos vivenciando esses desafios.

Percebemos, em questões éticas, alguns *cases* controversos no mercado, no Brasil e no mundo, como, por exemplo, todas as campanhas de “se beber, não dirija”, mas, por outro lado, haver bebida alcoólica sendo vendida em postos de gasolina. Também, há alimentos ultraprocessados, salgadinhos, *snacks*, biscoitos e bolachas sendo vendidos em hospitais. Outro exemplo é o patrocínio esportivo das *bets*. Essas

questões geram debates, podendo trazer mobilização e, eventualmente, até mesmo chegando ao boicote, como em outros países.

À medida que a IA está trazendo impactos positivos fantásticos, como na mobilidade urbana, com veículos autônomos, tudo isso deve, em algum momento, expandir de forma mais massiva para diversas cidades no mundo. Por outro lado, existem os diagnósticos médicos feitos cada vez mais com IA, bem como dados sendo compartilhados, que podem ou não afetar apólices de seguros. Isso leva a discussões sobre a privacidade desses dados médicos, decisões judiciais, leis da robótica e regulação no mundo.

Há casos, como na União Europeia, em que há uma política bastante restritiva, que acaba engessando inovação. Entretanto, também existe o outro extremo, dos Estados Unidos, que dão total liberdade às *big techs*, permitindo que elas façam o que quiser, seja o Meta, o Google, a Apple ou qualquer outra. A China tem uma política que protege bastante o Estado. Já o Brasil precisa se posicionar.

Uma forma interessante de escrever o futuro é olhar para a ficção científica. Em *Blade Runner* há a discussão sobre o direito das inteligências sociais, dos robôs. Em *Eu, robô* (2004), há a questão da responsabilidade dos robôs. Em *Her* (2013), filme no qual o Joaquim Phoenix se apaixona por um sistema operacional, há a questão da propriedade intelectual. *Westworld* (2016–2022), série de faroeste da Netflix, traz a questão dos precedentes e pessoas jurídicas com direitos distintos.

Outro exemplo de *case*, antes da IA, é o traje de natação da marca Speedo, que era uma bermuda com uma textura semelhante às escamas de barracuda, que deslizava maravilhosamente bem na água, gerando o questionamento se isso é *doping* tecnológico, representando uma competição desleal. Uma vez, numa piscina pública, percebi algumas pessoas nadando com pé-de-pato, o que é uma vantagem, fazendo com que passassem por mim numa velocidade maior, ou seja, não tem como competir, pois são mais rápidas e podem até mesmo gerar lesões. Por isso, no Brasil, não é permitido o uso de pé-de-pato ou de pranchas de mão na piscina do Sesc, porque podem machucar alguém. Essas técnicas de *doping* acabam trazendo as questões do aprimoramento humano.

Já a IA, com neurosensores, *neurochips* e próteses biônicas, faz com que essa simbiose na integração humano–máquina esteja cada vez mais presente. Temos cada

vez mais automação no mercado de trabalho, necessitando de transparência e equilíbrio, como tudo na vida. No filme *2001: uma odisseia no espaço* (1968), o robô se rebela e não responde à solicitação humana. Com a evolução da IA para a chamada IA geral, ela atingirá um nível superior a profissionais de qualquer área. Inclusive, é provável que, no futuro, os prêmios Nobel serão todos para a IA. Em 2024, o Prêmio Nobel em Biologia já foi para um dos fundadores do *Google Deep Minds*, que é uma tecnologia do Google que proveu uma identificação de proteínas através de IA. Mas quando chegarmos à IA geral, provavelmente outros prêmios Nobel devem ser voltados a ela.

Provavelmente na física, na química, na medicina e em todas as áreas, teremos IA. O teste de Turing, que identifica se é uma IA ou se é um humano, assim como o teste de retina e a entrevista em *Blade Runner* para identificar se o indivíduo é um robô, são recursos que já estão superatuais. Isso porque, à medida que os alunos começaram a utilizar a IA nas faculdades, em tudo quanto é canto, os professores começaram a se questionar sobre que ferramentas podem usar para saber se aquele texto foi feito com IA ou não.

Assim como ocorre no filme *Her* (2013), será que a IA vai delegar nossas emoções e algoritmos? No filme *Ex\_Machina* (2014), há uma robô sexy que manipula o protagonista. Em *2001: uma odisseia no espaço* (1968), o HAL (supercomputador) toma decisões e se revolta contra a solicitação humana. Inclusive, quando pedimos para a IA reescrever um código completo, provavelmente por programação, para gastar menos *tokens*, ela acaba não retornando esse código completo, dizendo “insira essa parte, nessa seção aqui”, e às vezes, eu quero o código completo, mas a IA já se recusa. Assim, já estamos vivenciando a ponta do iceberg em relação a essa insurgência.

Algo incrível é que Zac Azenov, na obra *Eu, robô* (1950), décadas e décadas atrás, já criou algumas leis que são superatuais. Na obra, as duas primeiras leis são: 1. não ferir os humanos; 2. obedecer aos humanos, exceto se houver conflito com a primeira lei. Assim, se um humano pedir para aquele robô virar um caçador e matar um outro humano, ele vai se recusar. A lei 3 é a da autopreservação, então, o robô não pode deixar ser destruído, exceto que, novamente, acabe entrando em conflito com a primeira ou a segunda lei. Já a lei zero é proteger a humanidade da autodestruição. Essas leis no mundo real gerariam uma esperança também, já que acabaríamos não fazendo muita

mudança em relação ao clima. Nossas ações, infelizmente, são mínimas, se observarmos a Eco92, o Acordo de Paris e o Protocolo de Kyoto, por exemplo. Quem sabe, com a evolução da IA, ela não deixe mais que continuemos “destruindo o planeta”.

Trazendo um pouco dessa nova tendência e conversando com algumas pessoas, percebemos que as novas gerações se sentem mais à vontade de se abrir, de se confidenciar e de receber feedback nas empresas de máquinas do que de humanos. Então, isso acaba impactando relações de trabalho, emocionais e de ordem psicológica. Alguns estudos demonstram exatamente essas questões vinculadas à geração mais nova. Dessa forma, obviamente, existem esses dilemas da IA com a terapêutica e a privacidade. Se a pessoa começa a se confidenciar com o *ChatGPT* ou com o *Gemini*, mostrando que está com tendências suicidas, é o momento dessa IA entrar em contato com um parente próximo ou com o Serviço de Apoio à Vida, ou não? Existem, então, as discussões dos limites de intervenções, de quem programa esses juízos de valor, de empatia.

Obviamente, há benefícios, como uma maior acessibilidade a serviços de saúde mental, pois, às vezes, as pessoas não têm dinheiro para pagar um psicólogo. Além disso, desestigmatiza essa busca por ajuda e fornece suporte a qualquer momento. Se a pessoa está em desespero, ela pode contar com a IA. Entretanto, existe a dependência e a falta de empatia, porque é uma máquina simulando o comportamento humano. Outro risco é a questão da privacidade, que é percebida claramente na música. Artistas como Elton John e Paul McCartney, por exemplo, têm defendido que a IA realmente muitas vezes infringe direitos autorais. O *Suno* e o *Udio* são ferramentas usadas para criar alguns hits, em qualquer estilo, com uma qualidade incrível, mas que, por muitas vezes usa dados dos melhores músicos do mundo sem consentimento.

Hollywood também tem discussões sobre essa questão da propriedade intelectual. O filme *O Congresso Futurista* (2013) é exatamente sobre uma atriz já em processo de decadência, de fim de carreira, que, então, decide fazer um contrato cedendo os direitos de imagem para filmes futuros através de hologramas e IA. Outro exemplo é o ator Robert Downey Jr., que já incluiu em seu testamento que não quer ser digitalizado, que quer manter os direitos autorais, e ninguém tem direito a utilizar a imagem dele em futuros filmes após sua morte. Além disso, semana passada, uma atriz de IA gerou toda uma revolta.

É importante acreditar, não na substituição, mas no aumento das nossas capacidades humanas como um *doping* tecnológico do bem, algo legal. Mas o problema todo é que teremos a possibilidade, no futuro, de utilizar tecnologias como *neuralink*, neurosensor, tiara, ou o próprio óculos Ray-Ban da Meta, que permitirão olhar para uma pessoa e filmar sem o consentimento dela, ou mesmo olhar para uma pessoa e já ser conectado ao seu LinkedIn ou às redes sociais.

Normalmente, tudo isso traz uma série de dilemas, como a edição genética, pela qual as pessoas podem decidir a cor dos olhos dos filhos e outras características. É muito interessante evitar uma doença de nascença, como a síndrome de Down, por exemplo, mas agir em questões de ordem estética já é um outro cenário e um outro espaço para debate e conversa. Então, há problemas na música, plágio, publicidade digital, publicidade infantil etc. Esse debate já existe muito antes da IA: jogos digitais, mecânicas viciantes, as próprias redes sociais, questões de representatividade, violência e valores.

Em relação às inteligências inteligentes, na China, algumas cidades têm centenas de câmeras espalhadas pelas ruas. Se a pessoa atravessa fora da faixa, ela já recebe uma notificação no celular informando a infração, havendo um sistema como um consórcio, ou seja, um Serasa, um *score* social. Então, se alguém comete muitas infrações, não consegue nem comprar uma passagem para eventualmente fugir do país. Dessa forma, por um lado, podemos utilizar esses sensores e dispositivos tanto em prol de cidadania como, no outro extremo, poder “virar um *Black Mirror*”. É um controle total.

Outra discussão é a ética de trabalho remoto. Houve um caso recente de um desligamento, por baixa produtividade do Itaú, de centenas de colaboradores, porque o banco alegou que eles não estavam trabalhando de fato. Temos também o Nubank, com os modelos auditados, a Magazine Luiza, com a transferência em recomendações, e organizações com políticas formais de ética digital, que são mais de 50 hoje.

Além do lado corporativo, existem também os marcos regulatórios, os incentivos positivos e a proteção social. A Diella, que significa “sol” em albanês, é uma ministra em IA na Albânia, com a proposta de auditar os contratos públicos, ou seja, por tabela, para coibir a corrupção no país. É o primeiro experimento no mundo nesse sentido. Então, será que podemos combater a corrupção no Brasil e em outros países com IA? Ou, em alguma parte do sistema, isso vai se corromper?

Alguns marcos regulatórios são: ISO 26000 (diretrizes internacionais sobre a responsabilidade social corporativa), LGPD (artigo 20 exige explicabilidade de decisões algorítmicas), AI Act (da Europa, com classificação de risco para sistemas de IA), PL 21/2020 (projeto de lei brasileiro para regulamentação da IA).

Podemos cultivar a ética digital por meio de *hackathons* éticos, eventos, prêmios de inovação, capacitação, *frameworks*, metodologias, workshops de co-criação de cenários éticos, com projetos de futuros participatórios, nos quais cenários futuros são explorados, criando visões futuras e implementando de forma colaborativa e inclusiva, trazendo uma diversidade de vozes. Uma possibilidade é utilizar fóruns, nos quais agimos, questionamos, incluímos e orientamos, mas, principalmente, no qual agimos.

Todos nós temos nossos vieses cognitivos. O mapa de vieses cognitivos mostra que há o efeito manada, de você seguir o seu grupo social, e o viés da confirmação, se você já tem uma predefinição sobre aquela pessoa, o que acaba gerando a polarização política. É a mesma coisa. Assim, se você acha que a esquerda ou a direita é de determinada forma, qualquer notícia no WhatsApp, sobre o que quer que seja, acaba confirmando isso.

Na máquina moral do Massachusetts Institute of Technology (MIT), em Boston, há o cenário, por exemplo, de um carro autônomo com problema, gerando o questionamento de qual cenário você escolhe: matar uma criança, um idoso ou um pet? É a escolha de Sofia, de quem vai viver e quem vai morrer. Esse foi um cenário atual na época da pandemia. Infelizmente, não havia equipamentos suficientes nos hospitais, então os médicos tiveram que fazer essas escolhas de Sofia naquela época, da mesma forma que escolher Sofia na Segunda Guerra. Percebemos que essas decisões têm um caráter cultural. Existem países que privilegiam mais o idoso, outros, mais as crianças. No próprio Titanic, primeiro evacuaram crianças e mulheres.

Renato Jannuzzi propôs o *Canvas de impacto*, que, diferentemente daquele canvas de modelo de negócios, é feito para analisar os impactos do negócio, seja no meio ambiente, seja em recursos, seja em pessoas ou nas comunidades locais, de onde se está, por exemplo, extraindo recursos minerais. Isso é um marco, pois já tivemos alguns problemas no passado com mineradoras. Então, deve-se não só pensar em um modelo de negócios que faça sentido para as pessoas, que faça esse sentido em termos de lucratividade para a empresa, mas que também possa avaliar sob essa ótica.

A Ideo é uma empresa que tem, em inglês, *cards* de ética, que podem ser levados para uma dinâmica colaborativa, trazendo junto com a equipe o desenvolvimento de um produto físico ou um produto digital utilizando questões éticas. O *design thinking* era uma metodologia de inovação centrada no ser humano, mas que, com essa evolução do planeta, acabou ficando um pouco obsoleto. Então, precisamos ir além do *design thinking* de inovação centrada no ser humano.

O *life-center design* é uma perspectiva centrada na vida e no planeta, levando em conta o olhar local e dando voz aos animais. Há, portanto, um campo de persona, de arquétipo, no qual você coloca o que o animal pensa, fala, escuta, da mesma forma que fazíamos na época do *design thinking*, que fazia esse tipo de atividade com pessoas. Pode-se dar voz até aos animais, às plantas, às árvores. Isso pode parecer maluquice, mas faz sentido. Design sustentável também com IA pode ajudar a produzir produtos sustentáveis, com uso especulativo, quando se considera exatamente esses cenários futuros e especulação.

A Roda de Futuros é como um mapa mental onde coloca-se uma temática no centro e verifica-se os impactos de primeira, segunda e terceira ordem. No caso do exemplo do carro autônomo (temática central), o primeiro impacto seria a redução de acidentes de trânsito; de segunda ordem, a redução de óbitos de fatalidade; o impacto de terceira ordem poderia ser a redução da doação de órgãos, que seria um impacto negativo, porque, no final das contas, muitas doações de órgãos no Brasil acontecem por conta de acidentes de trânsito. Por outro lado, existem tecnologias exponenciais, com impressão 3D, que, em breve, permitirão a produção hiperpersonalizada de um órgão para as pessoas não precisarem mais ficar na fila de doação.

A proposta é não nos tornarmos reféns de futuros, de futuros dos outros, futuro das *big techs*, mas antecipá-los, projetarmos futuros de forma colaborativa, não só porque a minha utopia pode ser a distopia de outros e vice-versa. Hábitos saudáveis, pra quem gosta de ultraprocessados, é uma distopia, por exemplo, e vice-versa. Então, precisamos juntar outras vozes. Existe também o *foresight* estratégico, que é olhar para o futuro de forma estratégica e cocriar. Assim, o futuro não existe, mas você cocria sua visão de futuro preferível. Tudo isso para gerar o bem-estar coletivo, que gere valor compartilhado e lucro de forma responsável. Realmente, as empresas precisam lucrar, mas não às custas de destruir o planeta ou comunidades locais.

“O futuro não é inevitável. O futuro é agora” é uma frase atribuída a Arthur Clarke nos anos 1970, associada ao livro *Profiles of Future*. Isso mais do que nunca, porque temos ferramentas e informação através da IA. Hoje em dia, podemos perguntar qualquer coisa. Há, portanto, um chamado à ação: educar-se continuamente; integrar ética no trabalho e nos estudos; advogar por mudanças; colaborar amplamente. Então, o futuro está em nossas mãos; não está para ser previsto, mas para ser construído e cocriado. Temos a responsabilidade coletiva de adquirir proteção, cultivo e legado.

Estudos em neurociência têm demonstrado uma menor ativação de algumas regiões do cérebro se começamos a delegar demais à IA, além de, se usarmos o princípio de Pareto com 20% de trabalho humano e 80% delegar à IA, podemos sofrer um declínio cognitivo das nossas faculdades mentais. Com a IA cada vez mais presente, tende a crescer o uso de assistentes em formato de “pingue-pongue”, nos quais o usuário pergunta, recebe uma resposta, corrige, adiciona contexto e segue refinando. Porém, há um nível acima disso: a automação por agentes, que recebem uma missão completa e são capazes de pesquisar, consolidar informações, diagramar um documento com imagens e gráficos e entregar um material final pronto, o que pode liberar tempo das pessoas para outras atividades.

Em cenários mais radicais, será necessário ressignificar o que é ser humano, porque não há como competir com máquinas em produtividade e velocidade. Por outro lado, valores e experiências humanas não são reproduzíveis pela máquina, o que tende a aumentar a valorização de competências humanas e do fazer artesanal. A ideia de “feito por humanos”, inclusive, será um diferencial no futuro. Existem tendências e contratendências, então é importante olhar para diferentes arquétipos de cenários, para não ficar preso a uma visão única e fatalista. É preciso sair do “atualismo”, daquela visão de que “o Brasil só está piorando”. Também há coisas positivas acontecendo.

É possível haver fadiga digital. Portanto, é preciso entender como contrabalançar hábitos de trabalho excessivo. Isso envolve três esferas: a individual, com conscientização e mudança de hábitos; a estatal, com políticas públicas e campanhas; a empresarial, já que muitas empresas, mesmo com ganhos de produtividade proporcionados pela IA, acabam exigindo ainda mais trabalho. Pode-se realizar práticas de “detox digital”, momentos off-line e atividades manuais como forma de ajudar na

memorização e ativar diferentes regiões do cérebro, ao contrário de hábitos como apenas fotografar conteúdos.

Para mulheres nordestinas, há mais de um cenário possível: por um lado, ferramentas atuais permitem criar aplicativos e sites sem que se saiba programação, usando voz ou texto, o que pode reduzir desigualdades ao permitir que pessoas com poucos recursos materializem ideias e empreendam. Além disso, a IA pode ampliar o acesso ao conhecimento para além da formação tradicional. Por outro lado, o *gap* pode aumentar, com países como China e Estados Unidos liderando e desafios persistindo em regiões com baixa conectividade. Pode haver um avanço de celulares capazes de processar IA localmente, diminuindo a dependência de internet no futuro, mas provavelmente ainda com custo elevado. Podem coexistir movimentos de aumento de desigualdade e, ao mesmo tempo, oportunidades de inclusão por meio de novas ferramentas.

Qualquer extremo, político ou social, tende a ser indesejável. Em algumas distopias, por exemplo, há vigilância total, como em *1984* (1949). Entretanto, visões de mundo “perfeito” também podem esconder problemas. De fato, sistemas muito rígidos podem até forçar certas práticas de cidadania, mas, se levados ao extremo, não são um caminho desejável.

Outra preocupação é o uso de recursos naturais e a ideia de a IA ter um preço para ser utilizada. Muitos serviços já funcionam em modelo *freemium*: parte é gratuita e, para uso mais intenso, é necessário pagar, muitas vezes em dólar, o que pode não ser acessível para muita gente. Ainda assim, existe uma porta de entrada gratuita em várias plataformas, embora com limites.

A IA é muito eficiente em reconhecer padrões, inclusive em saúde, havendo, por exemplo, pesquisas voltadas a decodificar linguagens de animais mais inteligentes, como golfinhos e primatas. No futuro, talvez seja possível acessar melhor a perspectiva dos animais e identificar com mais precisão sinais de dor ou sofrimento.

Para proteger a população e reduzir o medo do futuro diante do domínio da IA, algumas medidas são necessárias. No nível individual, são essenciais pensamento crítico, reflexão e debate sobre essas questões. Já no nível do Estado, é necessário haver mobilização social e pressão sobre representantes eleitos para que criem políticas públicas e regulamentações, evitando que a sociedade fique à mercê das *big techs* e da

lógica da economia da atenção. Também é importante impor limites de uso, especialmente para crianças e adolescentes, em redes sociais, videogames e outras plataformas, promovendo um uso responsável e ético da tecnologia.

### Referências

CLARKE, A. C. **Profiles of the future: an inquiry into the limits of the possible**. New York: Harper & Row, 1973.

JANNUZZI, R. Jannuzzi. **Canvas de impacto do negócio**. 2020. Disponível em: <https://renatojcec.com/canvas-de-impacto-do-negocio/>. Acesso em: 21 dez. 2025.

ORWELL, G. **1984**. São Paulo: Cia. das Letras, 2009.

TOFFLER, A. **A terceira onda**. Rio de Janeiro: Record, 1980.