



Ars Ex Machina: Repensando a responsabilidade na era das máquinas criativas

David Gunkel¹.

Northern Illinois University

Resumo: A medida que os computadores e os sistemas relacionados se aproximam cada vez mais das habilidades humanas, o único remanescente do excepcionalismo humano parece ser a criatividade e a arte. Não importa quão inteligentes ou capazes sejam nossos algoritmos, robôs ou sistemas relacionados, a "criatividade" parece ser a única capacidade humana que não é inteiramente computável. Mas agora existem artefatos que podem produzir o que parece ser produção criativa em jogos, geração de textos, artes visuais, música e também no jornalismo. Este texto avalia criticamente as oportunidades, desafios e repercussões de máquinas cada vez mais criativas, concentrando-se na questão da responsabilidade: quem ou o que pode ser responsável pelo trabalho que é gerado; 2) como decidir quem ou o quê se comunica por meio desses artefatos e 3) como nossas respostas a tais questões têm consequências.

Palavras-chave: Inteligência Artificial; Criatividade; Jornalismo; Responsabilidade; Ética.

1. Introdução

Em maio de 2015, a National Public Radio (NPR) encenou uma competição bastante informativa de humano versus máquina. Neste remake da lendária corrida entre

¹ Professor do Departamento de Comunicação da Northern Illinois University. Autor dos livros *Hacking Cyberspace*; *Thinking Otherwise: Philosophy, Communication, Technology*; *The Machine Question: Critical Perspectives on AI, Robots and Ethics*; *Rethinking the Digital Remix: Mash-ups and the Metaphysics of Sound Recording*. Trabalha com estudos sobre Inteligência Artificial, criatividade, ética. dgunkel@niu.edu

John Henry e steam power, um repórter da NPR, Scott Horsley, enfrentou o Wordsmith da Automated Insights, um algoritmo de geração de linguagem natural (NLG), projetado para analisar padrões em big data e transformá-los em narrativas humanas legíveis.

As regras do jogo eram simples: “Os dois contendores esperavam que a Denny’s, a empresa de lanchonetes, apresentasse um relatório de ganhos. Uma vez que foi lançado, o cronômetro começou. Ambos escreveram uma curta história de rádio e foram classificados em velocidade e estilo” (SMITH, 2015). Wordsmith cruzou a linha de chegada em apenas dois minutos com uma composição precisa, mas bastante utilitária. A submissão de Horsley levou mais tempo para ser escrita – sete minutos completos – mas foi considerada uma apresentação mais estilística dos dados.

O que esta pequena experiência demonstrou não é o que se poderia esperar. Não mostrou que a máquina é de alguma forma melhor ou tão boa quanto a do repórter humano. Em vez disso, revelou como esses programas são bons o suficiente para começar a desafiar seriamente as capacidades humanas e a deslocar esse tipo de trabalho. Na verdade, quando a revista Wired pediu a Kristian Hammond, co-fundador da Narrative Science, para prever a porcentagem de artigos de notícias que seriam escritos por algoritmos na próxima década, sua resposta foi um sóbrio 90% por cento (FORD, 2015, p. 85).

Para estudiosos da comunicação, no entanto, esta demonstração também aponta para outra questão relacionada, que está começando a reunir interesse e impulso nos estudos do jornalismo digital (cf. CARLSON, 2015; CLEARWALL, 2014; DÖRR & HOLLNBUCHER, 2016; LEWIS & WESTLUND; 2015 MONTAL E REICH, 2016). O texto escrito é normalmente entendido como o produto de alguém – um autor, repórter, escritor – que, supõe-se, algo a dizer ou a comunicar por meio do documento escrito. É claro, por exemplo, quem “fala” através do instrumento do texto composto pelo repórter humano. É o Scott Horsley. Ele é responsável não apenas por escrever a história, mas também por seu estilo formal e seu conteúdo. Se é uma história bem escrita, é Horsley quem recebe o elogio. Se contiver erros formais ou imprecisões factuais, é Horsley quem é responsabilizado. E se quisermos saber sobre o que o repórter escreveu e por quê, Horsley pode presumivelmente ser consultado e será capaz de responder à consulta. Esta conceituação não é apenas comum, tem o peso da tradição por trás dela.

Na verdade, este pensamento remonta ao Fedro de Platão, onde a escrita – provavelmente a primeira tecnologia da informação – estava situada tanto como produto derivado do discurso falado, quanto como uma criança muda e abandonada, sempre precisando da autoridade do pai para responder por ela e em seu nome (PLATÃO, 1982, p. 275).

Mas e a outra história, a do Automated Insights, do Wordsmith? Quem ou o que fala em um documento que foi escrito – ou montado ou gerado (e a escolha do verbo, verifica-se, importa aqui) -- por um algoritmo? Quem ou o que é, ou pode ser responsabilizado pela escrita? Quem ou o que pode responder em seu nome? É a corporação que fabrica e distribui o software? São os programadores da corporação que foram contratados para escrever as instruções do software? São os dados aos quais o programa teve acesso? É o usuário do aplicativo que o configurou e direcionou para trabalhar nos dados? Ou talvez seja o próprio Wordsmith? O problema, claro, é que essas questões não são tão facilmente resolvidas. Não está totalmente claro quem ou o quê (se algo) fala no e para este texto². Como Montal e Reich (2016) demonstraram, em seu estudo “I Robot. Você, jornalista. Quem é o autor? ”, O desenvolvimento e a implementação do “jornalismo automatizado” resultou em “grandes discrepâncias entre as percepções de autoria e política de creditação, os regimes de atribuição prevalentes e a literatura acadêmica” (2016, p. 1).

Essa incerteza em relação à autoria e à atribuição abre uma “lacuna de responsabilidade” significativa que afeta não apenas como pensamos sobre quem ou o que se comunica, mas também como entendemos e respondemos a perguntas relativas à responsabilidade na era de máquinas cada vez mais criativas³.

² Pode-se também perguntar quem ou o que “fala” nesse e por esse texto? O exame da criatividade computacional não é apenas sobre tecnologia; tem consequências importantes para a forma como conceituamos a criatividade (humana), a autoria e a responsabilidade. Para mais sobre este assunto, ver Gunkel (2016).

³ Depende muito de como se define e caracteriza o termo “criativo”. De fato, um método comprovado para proteger o excepcionalismo humano da criatividade da incursão maquínica é definir (e então, se necessário, redefinir) “criativo” de modo que permaneça imune e protegido do processamento computacional. Como Matt Carlson (2015) explica: “O apelo ao jornalismo como atividade criativa também diferenciava as notícias humanas das automatizadas. Rebecca Greenfield (The Wire, 25 de abril de 2012) chegou a questionar se o que a Narrative Science produzia poderia ser considerado jornalismo: “Há empresas inteiras baseadas na ideia de produzir grandes quantidades de notícias, qualidade controlada por fórmulas semelhantes às máquinas. A Narrative Science pode um dia colocar muitos desses jornalistas fora do trabalho. Mas quando a maioria das pessoas fala sobre jornalismo, elas não estão pensando em relatos de ganhos ou repescagens de jogos de basebol. “O jornalismo apropriado foi conceituado como algo mais profundo” (2012, p. 428). A fim de evitar entrar no debate corrente e aparentemente insolúvel

Essas questões são centrais para (se não definitivas ao) projeto de comunicação homem-máquina (HMC). Diferentemente do paradigma dominante de comunicação mediada por computador (CMC), que restringe computadores, robôs e outros tipos de tecnologias à posição intermediária de serem meros instrumentos de expressão humana e transmissão de mensagens (Gunkel, 2012a), a pesquisa do HMC investiga se – e até que ponto – as máquinas são capazes de serem agentes comunicativos por direito próprio.

Este artigo investiga a oportunidade e os desafios que máquinas cada vez mais criativas têm em nossa compreensão de quem ou do que se comunica, quem ou o que pode ser responsável por gerar conteúdo original e quem ou o que ocupa a posição de “Outro” nas interações e relacionamentos sociais. Como essas questões são em grande parte filosóficas, o método do exame também será filosófico em sua orientação, procedimentos e objetivos.

Responsabilidade

O “conceito de responsabilidade”, como Paul Ricœur (2007) apontou em seu ensaio homônimo, é qualquer coisa, exceto claro e bem definido. Embora o uso jurídico clássico do termo, que remonta ao século XIX, pareça bastante bem estabelecido - com “responsabilidade” caracterizada em termos de obrigações civis e penais (seja a obrigação de compensar danos ou a obrigação de se submeter a punição) – o conceito geral é confuso e um pouco vago.

Em primeiro lugar, nos surpreende que um termo com tal sentido firme no plano jurídico deva ser de origem tão recente e não muito bem estabelecido dentro da tradição filosófica. Em seguida, a atual proliferação e dispersão de usos desse termo é intrigante, especialmente porque vai muito além dos limites estabelecidos para seu uso jurídico. O adjetivo “responsável” pode complementar uma grande variedade de coisas: você é responsável pelas consequências de seus atos, mas também é responsável pelas ações dos outros, na medida em que eles foram feitos sob sua responsabilidade ou cuidado... Nes-

do que realmente constitui a “verdadeira criatividade” ou “jornalismo adequado” (que é discutivelmente um esforço exposto ao que Blay Whitby (2011) chama de “No True Scotsman Fallacy”, ver Gunkel (2012a, p. 9). Vou começar com uma definição geral do dicionário, em que “criatividade” é operacionalizada como “o processo de trazer algo à existência”. Esta formulação será mais especificada, desenvolvida e complicada no curso da análise.

ses usos difusos, a referência à obrigação não desapareceu, tornou-se a obrigação de cumprir certas funções, para assumir certos encargos, para cumprir certos compromissos (RICOUER, 2007, pp. 11-12).

Ricœur (2007) traça este sentido da palavra através de sua etimologia (daí o subtítulo para o ensaio “A Semantic Analysis”) até “a polissemia do verbo 'responder’”, que denota “responder por...” ou “para responder a... (uma pergunta, uma apelação, uma liminar, etc.)” (2007, p. 12). Responsabilidade, então, envolve ser capaz de responder e / ou responder por algo – alguma decisão, ação ou ocorrência que eu tenha instituído diretamente por mim ou que tenha sido cobrado ou designado para alguém ou alguma outra coisa sob minha direção ou cuidado.

Essa caracterização é consistente com o desenvolvimento do conceito de autor, que, como Roland Barthes (1978, pp. 142-143) argumentou, não é um fenômeno natural, mas uma figura de autoridade fabricada deliberadamente introduzida e desenvolvida no pensamento europeu moderno. A figura moderna do autor, como Michel Foucault (1984) explica, foi originalmente instituída para responder a uma lacuna percebida na responsabilidade. Como um texto escrito é, como Sócrates inicialmente descreveu (Platão, 1982), cortado de seu progenitor e circula além de seu controle ou supervisão, então as autoridades (governos ou a igreja) precisavam ser capazes de identificar e atribuir responsabilidade a alguém pelo que foi declarado no texto.

Como Foucault (1984) explica, o autor era uma figura de “apropriação penal”. “Textos, livros e discursos realmente começaram a ter autores (além de figuras míticas, 'sacralizadas' e 'sacralizadoras'), na medida em que o autor tornou-se sujeito à punição, isto é, na medida em que os discursos poderiam ser transgressivos” (1984, p. 108). Em outras palavras, os textos vêm a ser organizados sob a figura de um autor para que as autoridades possam identificar quem deve ser responsabilizado por uma declaração publicada, para que se conheça quem pode ser questionado ou quem pode responder em nome da pessoa do texto, e quem poderia, portanto, ser punido por transgressões percebidas.

Acomodar a tecnologia nessa maneira de pensar não é difícil nem complicado. A caneta e o papel, o pincel e a tinta a óleo, a guitarra elétrica e o amplificador são todas tecnologias – essencialmente ferramentas disponíveis e usadas por um artista ou artesão

humano. O que mais importa não é o equipamento usado, mas como esses itens são empregados e por quem produzir o tipo de artefato ou experiência. É, em outras palavras, não a ferramenta, mas o usuário da ferramenta, que é o responsável final pelo que é feito ou não com um instrumento tecnológico específico.

Essa maneira aparentemente intuitiva e de senso comum é persuasiva precisamente porque é estruturada e informada pela resposta tipicamente fornecida à pergunta sobre a tecnologia. “Fazemos a pergunta sobre a tecnologia”, explica Martin Heidegger (1977), “quando perguntamos o que é. Todo mundo conhece as duas afirmações que respondem à nossa pergunta. Uma diz: a tecnologia é um meio para um fim. A outra diz: a tecnologia é uma atividade humana” (1977, pp. 4–5). De acordo com a análise de Heidegger, o papel e função presumidos de qualquer tipo de tecnologia – seja uma ferramenta manual simples, um avião a jato ou um robô sofisticado – é que é um meio empregado por usuários humanos para fins específicos. Heidegger classifica essa particular caracterização da tecnologia como “a definição instrumental” e indica que isso constitui o que se considera como o entendimento “correto” de qualquer artifício tecnológico.

Como Andrew Feenberg (1991) resume, “a teoria instrumentalista oferece a visão mais amplamente aceita da tecnologia. Baseia-se na ideia do senso comum de que as tecnologias são “ferramentas” prontas para atender aos propósitos dos usuários” (1991, p. 5). E porque uma ferramenta ou instrumento “é considerado 'neutro', sem conteúdo valorativo próprio”, um artefato tecnológico é avaliado não por si, mas com base nos empregos particulares que foram decididos por seu projetista ou usuário humano. Consequentemente, a tecnologia é apenas um meio para um fim; não é e não tem um fim por direito próprio. Como Jean-François Lyotard (1993) resumiu com precisão no A condição pós-moderna:

Dispositivos técnicos originados como próteses para os órgãos humanos ou como sistemas fisiológicos cuja função é receber dados ou condicionar o contexto. Eles seguem um princípio, e é o princípio do desempenho ótimo: maximizar a saída (a informação ou modificação obtida) e minimizar a entrada (a energia gasta no processo). A tecnologia é, portanto, um jogo não pertencente ao verdadeiro, ao justo ou ao belo, etc., mas à eficiência: um “movimento” técnico é “bom” quando faz melhor e / ou gasta menos energia que o outro (1993, p. 33).

De acordo com a análise de Lyotard, um dispositivo tecnológico, seja um saca-rolhas, um piano ou um computador, é um mero instrumento de ação humana. Portanto, ela não participa das grandes questões de verdade, justiça ou beleza. É simples e indiscutivelmente sobre eficiência. Uma inovação tecnológica particular é considerada “boa”, se, e somente se, ela se revelar um instrumento (ou meio) mais efetivo para a realização de um fim humanamente definido.

Caracterizado como uma ferramenta ou instrumento do esforço humano, os dispositivos técnicos não são considerados o agente responsável das ações que são realizadas com ou através deles. Essa percepção é uma variante de uma das objeções apontadas por Alan Turing em seu documento sobre inteligência de máquina: “Nossas informações mais detalhadas do Mecanismo Analítico de Babbage”, escreveu Turing (1999), “vêm de um livro de memórias de Lady Lovelace (1842). Nele, ela afirma: “O mecanismo analítico não tem pretensões de originar nada. Ele pode fazer o que quisermos para fazer (itálicos da autora)” (1842, p. 50). Esse esclarecimento – o que Turing chamou de “A objeção de Lady Lovelace” – foi frequentemente usado como base para negar a agência independente ou autonomia para computadores, robôs e outros mecanismos. Tais instrumentos só fazem o que nós os programamos para executar. Por mais aparentemente independente ou autônomo que um sistema técnico seja ou seja projetado para aparecer, sempre existe, em algum lugar e de alguma forma, alguém “atrás da cortina”, puxando as cordas – e, como tal, responsável em última instância e capaz de responder pelo que acontece (ou não) com o instrumento tecnológico.

O novo normal

A teoria instrumental não apenas soa razoável, é obviamente útil. Pode-se dizer que é instrumental por responder às oportunidades e desafios disponibilizados por sistemas e dispositivos tecnológicos cada vez mais complexos. Isso ocorre porque a teoria foi aplicada com sucesso não apenas a dispositivos simples como martelos, pincéis e guitarras elétricas, mas também à sofisticada tecnologia da informação e sistemas, como computadores, aplicativos de inteligência artificial, robôs, etc. Mas tudo isso pode ter acabado precisamente por causa de uma série de inovações recentes, que desafiam as

capacidades explicativas da teoria instrumental, abrindo lacunas significativas na identificação e atribuição de responsabilidade.

Aprendizado de Máquina

As capacidades da máquina são tipicamente testadas e comparadas com jogos, como a corrida entre Scott Horsley e Wordsmith com a qual começamos. Desde o início, na verdade, a condição definidora da inteligência da máquina foi estabelecida com um jogo. Embora a expressão "inteligência artificial" (IA) seja o resultado de uma conferência acadêmica organizada por John McCarthy no Dartmouth College, no verão de 1956, é o trabalho de Alan Turing, intitulado "Computing Machinery and Intelligence", e seu "jogo de imitação" que define e caracteriza o campo. De acordo com Turing, o lugar imediato e aparentemente correto para começar (a saber, com a questão "As máquinas podem pensar?") foi considerado muito ambíguo e mal definido. Por essa razão, Turing mudou o modo de investigação. Ele substituiu a pergunta "As máquinas podem pensar?", com uma demonstração que tomou a forma de uma espécie de jogo de salão envolvendo engano deliberado e identidade equivocada.

A nova forma do problema pode ser descrita em termos de um jogo que chamamos de "jogo de imitação". Ele é jogado com três pessoas, um homem (A), uma mulher (B) e um interrogador (C) que pode seja de qualquer sexo. O interrogador fica em um quarto separado dos outros dois. O objetivo do jogo para o interrogador é determinar qual dos outros dois é o homem e qual é a mulher (TURING, 1999, p. 37).

Turing então faz uma pequena modificação nessa configuração inicial trocando um dos participantes humanos. "O que vai acontecer", Turing (1999) pergunta, "quando uma máquina toma a parte de A neste jogo? O interrogador decidirá erroneamente quantas vezes o jogo for jogado assim, como acontece quando o jogo é jogado entre um homem e uma mulher?" É esta questão, conclui Turing, que "substitui" a questão inicial "As máquinas podem pensar?" (1999, p. 38).

Desde a introdução do "jogo da imitação", por Turing, o desenvolvimento e a realização da IA foram marcados e medidos em termos de jogos e competições entre humanos e máquinas. Já no final dos anos 1950, Arthur Samuel criou uma aplicação

rudimentar de “machine learning” (um termo que Samuel fabricou e introduziu em 1959) que aprendeu a jogar e acabou dominando o jogo de damas. Em 1997, o Deep Blue da IBM derrotou Gary Kasparov no jogo de xadrez, obrigando Douglas Hofstadter (2001), que anteriormente havia rejeitado essa possibilidade, a retratar sua previsão original:

Sabemos agora que a habilidade de jogar xadrez de classe mundial pode realmente ser alcançada por técnicas de força bruta -- técnicas que de modo algum tentam replicar ou emular o que se passa na cabeça de um grande mestre de xadrez. O processo de analogia não é necessário, nem a memória associativa, nem os flashes intuitivos que separam o trigo do joio - apenas uma busca tremendamente ampla e profunda, realizada por hardware super-rápido, especializado em xadrez, usando quantidades ímpias de conhecimento armazenado (HOFSTADTER, 2001, p. 35).

Apesar das aparências iniciais, o xadrez – e este jogo em particular – não era um mero jogo. Muita coisa conta sobre isso, principalmente porque se presumia que o xadrez do grande mestre exigia uma espécie de gênio – o tipo de gênio que é a condição definidora do excepcionalismo humano.

Até certo ponto, essa partida é uma defesa de toda a raça humana. Os computadores desempenham um papel tão grande na sociedade. Eles estão em todo lugar. Mas há uma fronteira que eles não devem atravessar. Eles não devem cruzar a área da criatividade humana. Isso ameaçaria a existência do controle humano em áreas como arte, literatura e música (KASPAROV, 1996, apud HOFSTADTER, 2001, p. 40).

Mas o xadrez foi apenas o começo. Quatorze anos depois, o Watson da IBM limpou o game game Jeopardy. Então, em 2015, houve AlphaGo, um algoritmo Go-playing desenvolvido pelo Google DeepMind, que levou 4 de 5 jogos contra um dos jogadores humanos mais célebres deste notoriamente difícil jogo de tabuleiro.

O AlphaGo é único na medida em que empregou uma arquitetura híbrida que combina aspectos da programação GOFAI⁴, como a metodologia de busca em árvore

⁴ O GOFAI é um acrônimo para “Good Old Fashioned Artificial Intelligence”, que John Haugeland (1985, p. 112) implantou pela primeira vez para distinguir abordagens “clássicas” do desenvolvimento de IA de outras arquiteturas como o connencionalismo. O GOFAI operacionaliza a inteligência como computação e formaliza a computação “como a manipulação governada por regras de cadeias de símbolos interpretáveis” (Walmsley, 2012, p. 48). Nas abordagens GOFAI, há etapas de programação explícitas e discretas que são codificadas e executadas linha a linha. Se um programa em particular produz um resul-

que foi utilizada tanto pelo Deep Blue quanto pelo Watson, com capacidades de aprendizado de máquina de redes neurais profundas derivadas do trabalho pioneiro de Arthur Samuel. Como o Google DeepMind (2016) explicou, o sistema “combina pesquisa de árvore Monte-Carlo com redes neurais profundas que foram treinadas por aprendizado supervisionado, de jogos de especialistas em humanos e por aprendizado por reforço de jogos de auto-jogo”. AlphaGo não joga simplesmente seguindo um conjunto de movimentos inteligentemente projetados e alimentados por programadores humanos. Ele é projetado para formular suas próprias instruções e agir de acordo com essas “decisões”. Como Thore Graepel, um dos criadores do AlphaGo, explicou:

embora tenhamos programado esta máquina para jogar, não temos ideia de como ela virá com seus movimentos, são um fenômeno emergente do treinamento. Nós apenas criamos os conjuntos de dados e os algoritmos de treinamento. Mas os movimentos que surgem estão fora de nossas mãos (METZ, 2016c).

Consequentemente, o AlphaGo é intencionalmente projetado para fazer coisas que seus programadores não podem antecipar ou mesmo entender. E isso é, pelo menos para Hofstadter, o ponto em que as máquinas começam a abordar o que é tipicamente chamado de “criatividade”. “Quando os programas deixam de ser transparentes para seus criadores, a abordagem à criatividade já começou” (Hofstadter, 1979, p 670).

O o AlphaGo é um sistema de computador autônomo (ou pelo menos semi-autônomo) que parece ter algo de “mente própria”. E é aí que as coisas ficam interessantes, especialmente quando se trata de perguntas sobre responsabilidade. O AlphaGo foi projetado para jogar Go e provou suas habilidades ao vencer um jogador humano experiente. Então, quem ganhou? Quem recebe o elogio? Quem realmente venceu Lee Sedol? Seguindo os ditames da teoria instrumental da tecnologia, as ações realizadas com o computador precisariam ser atribuídas aos programadores humanos que inicialmente projetaram o sistema e são capazes de responder pelo que ele faz ou não faz. Mas essa explicação não se encaixa necessariamente em um aplicativo como o AlphaGo, que foi deliberadamente criado para fazer coisas que excedem o conhecimento e o controle de

tado surpreendente ou imprevisto, o programador pode rever o código fonte, identificar exatamente onde ocorreu o “erro” e instituir uma alteração.

seus designers humanos. De fato, na maior parte das reportagens sobre este evento marcante, não são os engenheiros da Google ou da DeepMind que são creditados com a vitória. É o AlphaGo. Em rankings publicados, por exemplo, é AlphaGo que é nomeado como o número dois do mundo (Go Ratings, 2016).

Criatividade computacional

O AlphaGo é apenas um exemplo do que pode ser chamado de criatividade computacional. “Criatividade computacional”, como definido por Simon Colton e Geraint A. Wiggins (2012), “é um subcampo da pesquisa em Inteligência Artificial (IA) ... onde construímos e trabalhamos com sistemas computacionais que criam artefatos e idéias” (p. 21). Wordsmith e o produto concorrente Quill, da Narrative Science, são bons exemplos desse tipo de esforço na área da narrativa e da escrita de narrativas. Inovações semelhantes foram desenvolvidas no campo da composição e do desempenho da música, em que algoritmos e robôs produzem o que se costuma chamar (ou pelo menos tentado a chamar) de “trabalhos originais”.

Na música clássica, por exemplo, há os Experimentos de David Cope. em Musical Intelligence (EMI, pronunciado Emmy) e sua sucessora Emily Howell, que são compositores algorítmicos capazes de analisar composições existentes e gerar novas pontuações originais que são comparáveis e, em alguns casos, indistinguíveis das obras canônicas de Mozart, Bach, e Chopin (Cope, 2001). Na performance musical, há Shimon, uma marimba tocando jazz-bot da Georgia Tech University que não só é capaz de improvisar com músicos humanos em tempo real, mas “é projetada para criar interações musicais significativas e inspiradoras com humanos, levando a novas experiências musicais e resultados ”(Georgia Tech, 2013; Hoffman & Weinberg, 2011). E na área de arte visual, há The Painting Fool, de Simon Colton, um pintor automatizado que aspira a ser “levado a sério como um artista criativo por direito próprio” (Colton, 2012, p. 16).

Mas projetar sistemas para serem criativos imediatamente se depara com um problema semelhante ao encontrado originalmente por Turing. Como Amílcar Cardoso, Tony Veale e Geraint A. Wiggins (2009) reconhecem explicitamente, “a criatividade é um fenômeno elusivo” (2009, p. 16). Por esta razão, pesquisadores no campo da criatividade computacional introduziram e operacionalizaram uma formulação bastante espe-

cífica para caracterizar seus esforços: “A filosofia, ciência e engenharia de sistemas computacionais que, assumindo responsabilidades particulares, exibem comportamentos que observadores imparciais considerariam ser criativo” (Colton & Wiggins, 2012, p. 21).

O termo operativo nesta caracterização é responsabilidade. Como Colton e Wiggins (2012) explicam, “a palavra responsabilidades destaca a diferença entre os sistemas que construímos e as ferramentas de apoio à criatividade” (2012, p. 21, ênfase no original). Com um aplicativo de software como o Photoshop, “o programa é uma mera ferramenta para melhorar a criatividade humana” (Colton, 2012, pp. 3–4); é um instrumento usado por um artista humano que é e permanece responsável pelas decisões criativas. A pesquisa em criatividade computacional, em contraste, “esforça-se para construir software que é independentemente criativo” (Colton, 2012, p. 4).

Isso requer mudar cada vez mais a responsabilidade do usuário humano para o mecanismo. Como Colton (2012) descreve, “se podemos pedir repetidamente, responder e codificar software para assumir quantidades crescentes de responsabilidade, ele eventualmente (...) começará a criar de forma autônoma para um propósito, com pouca ou nenhuma envolvimento humano” (Colton, 2012, p. 13). Indicativo desta mudança na posição e atribuição de responsabilidade é o site de The Painting Fool, que foi deliberadamente concebido para que seja o programa de computador que assume a responsabilidade de responder em seu próprio nome.

A pergunta que falta em Colton é: quem compôs essa explicação? Foi gerado por The Painting Fool, que foi projetado para oferecer alguma explicação sobre seus próprios esforços criativos? Ou é o produto de um ser humano, como Simon Colton, que assume a responsabilidade de responder por e em nome do programa?

Embora o grau de atribuição de responsabilidade artística a esses mecanismos continue sendo uma questão contestada e indecisa, o que não é debatido é o fato de que as regras do jogo parecem estar em fluxo e que há cada vez mais evidências de uma lacuna de responsabilidade. Mesmo que isso seja, neste momento, o que Mark Riedl e outros chamaram de mera “imitação” e não criatividade real (Simonite, 2016) - que é, devemos observar, apenas outra versão ou uma imitação de John Searle (1984) Argumento da sala chinesa - o trabalho da máquina nos obriga a reconsiderar como a respon-

sabilidade é atribuída e, no processo, desafia a forma como normalmente respondemos às questões relativas à responsabilidade.

Conclusões

No final, o que temos é uma situação em que nossa teoria da tecnologia - uma teoria que tem uma história considerável por trás disso e que foi determinada como aplicável a ferramentas manuais simples como é a sistemas tecnológicos complexos - parece ser incapaz de responder ou responder por desenvolvimentos recentes em aprendizado de máquina e criatividade computacional, onde a responsabilidade é cada vez mais atribuível e atribuída à máquina. Embora isso certamente faça diferença ao decidir questões de obrigação legal e moral, também é crucial em situações relacionadas à criatividade e à inovação. A criatividade, na verdade, parece ser a última linha de defesa para impedir o "apocalipse robótico" iminente. E não é só Kasparov que acha que há muito a perder nas máquinas. De acordo com Colton e Wiggins (2012), a pesquisa da IA convencional também marginalizou os esforços na criatividade computacional. “Talvez”, escrevem eles, “a criatividade seja, para alguns defensores da IA, o lugar que não se pode ir, como a inteligência é para os oponentes da IA. Afinal, a criatividade é uma das coisas que nos torna humanos; nós valorizamos muito, e nós a guardamos com inveja”(p. 21). Portanto, a questão que resta a ser respondida é como podemos ou devemos responder às oportunidades / desafios de *ars ex machina*.

Podemos, por um lado, responder como normalmente temos, dispensando essas inovações tecnológicas recentes como apenas mais um instrumento ou ferramenta da ação humana. Essa abordagem foi modelada e implantada com sucesso em situações relativas à responsabilidade moral e legal e é a condição definidora da ética computacional. “Sistemas de computador”, escreve Deborah Johnson (2006).

são produzidos, distribuídos e utilizados por pessoas envolvidas em práticas sociais e atividades significativas. Isso é tão verdadeiro nos atuais sistemas de computadores quanto nos futuros sistemas de computadores. Independentemente de quão independentes, automáticos e interativos os sistemas computacionais do futuro se comportem, eles serão os produtos (diretos ou indiretos) do comportamento humano, das instituições sociais humanas e da decisão humana. (p. 197)

Entendido dessa maneira, os sistemas de computação, independentemente de quão automáticos, independentes ou aparentemente autônomos possam se tornar, não são e nunca poderão ser agentes autônomos e independentes (Johnson, 2006, p. 203). Eles serão, como todos os outros artefatos tecnológicos, sempre e para sempre instrumentos de valor humano, tomada de decisão e ação. Quando algo ocorre por meio de uma máquina - seja para o bem ou para o mal - há sempre alguém - alguma pessoa ou pessoas humanas - que podem responder por isso e ser responsabilizados.⁶

O mesmo argumento poderia ser feito para aplicativos aparentemente criativos como AlphaGo, Emily Howell ou The Painting Fool. Quando AlphaGo vence uma grande competição, quando uma partitura atribuída a Emily Howell é executada por uma orquestra sinfônica, ou quando The Painting Fool gera um impressionante trabalho de arte visual que é exibido em uma galeria, ainda há alguma pessoa humana (ou pessoas) quem é o responsável final e pode responder ou responder pelo que foi produzido. As linhas de atribuição podem se tornar cada vez mais complicadas e demoradas, mas há, pode-se argumentar, sempre alguém por trás dos bastidores quem é o responsável. E a evidência disso já está disponível nas situações em que foram feitas tentativas de transferir a responsabilidade para a máquina. Considere o lance decisivo de AlphaGo 37 no jogo dois contra Lee Sedol. Se quisermos saber mais sobre a mudança e sua importância, o AlphaGo certamente pode ser questionado sobre isso. Mas o algoritmo não terá nada a dizer em resposta. Na verdade, era responsabilidade dos programadores e observadores humanos responder em nome da AlphaGo e explicar o significado e o impacto do movimento. Como a tecnologia de escrever no Fedro de Platão, se você perguntar, o texto diz apenas uma e a mesma coisa, e precisa sempre da ajuda do pai, quando questionado ou injustamente injuriado (Platão, 1982, p. 275d-e). Consequentemente, como Colton (2012) e Colton et al. (2015) reconhecem explicitamente, se o projeto de criatividade computacional é para ter sucesso, o software precisará fazer mais do que produzir artefatos e comportamentos que tomamos e respondemos como saída criativa. Também precisará assumir a responsabilidade pelo trabalho, explicando o que fez e como o fez. “O software”, como Colton e Wiggins (2012) afirmam, “deve estar disponível para questionamento sobre suas motivações, processos e produtos” (p. 25), eventualmente não apenas gerando títulos e explicações e narrativas sobre o trabalho, mas também

sendo capaz de responder às questões entrando em diálogo crítico com seu público (Colton et al., 2015, p. 15). Embora Colton não o reconheça explicitamente como tal, esse esforço situa a criatividade computacional diretamente dentro do paradigma da pesquisa do HMC.

Ao mesmo tempo, e por outro lado, não devemos ser rápidos demais para descartar ou explicar as oportunidades abertas por essas incursões maquínicas e intervenções no que tem sido um domínio protegido e exclusivamente humano.⁷ A questão, na verdade, não é simplesmente se computadores, algoritmos de aprendizado ou outras aplicações podem ou não ser responsáveis pelo que fazem ou não fazem; a questão também tem a ver com a forma como determinamos, descrevemos e definimos a responsabilidade em primeiro lugar. Isso significa que há um componente forte e fraco para esse esforço, o que Mohammad Majid al-Rifaie e Mark Bishop (2015, p. 37) chamam, seguindo a distinção original de Searle em relação aos esforços em IA, formas fortes e fracas de criatividade computacional. . Esforços no que seria a variedade “forte” envolvem os tipos de desenvolvimento de aplicativos e demonstrações introduzidas por indivíduos e organizações como Simon Colton, Google DeepMind ou David Cope. Mas esses esforços também têm um aspecto de "inteligência artificial fraca" na medida em que simulam, operacionalizam e testam várias conceituações de responsabilidade e expressão artística, levando a reavaliações críticas e potencialmente perspicazes de como caracterizamos esse conceito em nosso próprio pensamento. Como Douglas Hofstadter (2001) admitiu, nada o fez repensar seu próprio pensamento sobre pensar mais do que a tentativa de lidar com o sentido do EMI de David Cope (p. 38). Em outras palavras, desenvolver e experimentar novas capacidades de máquina não necessariamente tira nada de seres humanos e o que (presumivelmente) os torna especiais, mas oferece novas oportunidades para ser mais preciso e científico sobre essas características distintivas e seus limites. Não é, portanto, um jogo de soma zero onde um lado ganha e o outro necessariamente perde.

Referências

- Barthes, R. (1978). *Image, music, text* (S. Heath, Trans.). New York, NY: Hill and Wang.
- Cardoso, A., Veale, T., & Wiggins, G. A. (2009). Converging on the divergent: The history (and future) of the international joint workshops in computational creativity. *AI magazine*, 30(3), 15–22. Retrieved from <http://www.aaai.org/ojs/index.php/aimagazine/article/view/2252>
- Carlson, M. (2015). The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. *Digital Journalism*, 3(3), 416–431.
- Clearwall, C. (2014). Enter the robot journalist: Users' perceptions of automated content. *Journalism practice*, 8(5), 519–531.
- Colton, S. (2012). The painting fool: Stories from building an automated painter. In J. McCormack and M. d'Inverno (Eds.), *Computers and creativity* (pp. 3–38). Berlin: Springer Verlag.
- Colton, S. & Wiggins, G. A. (2012). Computational creativity: The final frontier. In L. De Raedt, C. Bessiere, D. Dubois, P. Doherty, P. Frasconi, F. Heintz, & P. Lucas (Eds.), *Frontiers in artificial intelligence and applications*, (Vol. 242, pp. 21–26). Amsterdam: IOS Press Ebooks. <http://ebooks.iospress.nl/volume/ecai-2012>
- Colton, S., Pease, A., Corneli, J., Cook, M., Hepworth, R., & Ventura, D. (2015). Stakeholder groups in computational creativity research and practice. In T. R. Besold, M. Schorlemmer, & A. Smaill (Eds.), *Computational creativity research: Towards creative machines* (pp. 3–36). Amsterdam: Atlantis Press.
- Cope, D. (2001). *Virtual music: Computer synthesis of musical style*. Cambridge, MA: MIT Press.
- Dörr, K. N., & Hollnbuchner, K. (2016). Ethical challenges of algorithmic journalism. *Digital journalism*, 5(4), 404–419.
- Feenberg, A. (1991). *Critical theory of technology*. New York: Oxford University Press.
- Ford, M. (2015). *Rise of the robots: Technology and the threat of a jobless future*. New York: Basic Books.

- Foucault, M. (1984). What is an author? (J. V. Harari, Trans.). In P. Rabinow (Ed.) *Foucault reader* (pp. 101–122). New York, NY: Pantheon.
- Georgia Tech. (2013). Robotic musicianship group: Shimon.
<http://www.gtcmt.gatech.edu/researchprojects/shimon>
- Go Ratings. (20 November 2016). <https://www.goratings.org/>
- Google DeepMind. (2016). AlphaGo. <https://deepmind.com/alpha-go.html>
- Green, P., & Wei-Haas, L. (1985). <https://deepblue.lib.umich.edu/handle/2027.42/174>
- Gunkel, D. J. (2012a). Communication and artificial intelligence: Opportunities and challenges for the 21st century. *Communication +1, 1*(1), 1–27.
<http://scholarworks.umass.edu/cpo/vol1/iss1/1>
- Gunkel, D. J. (2012b). *The machine question: Critical perspectives on AI, robots and ethics*. Cambridge, MA: MIT Press.
- Gunkel, D. J. (2016). *Of remixology: Ethics and aesthetics after remix*. Cambridge, MA: MIT Press.
- Gunkel, D. J., & Taylor, P. A. (2014). *Heidegger and the media*. Cambridge: Polity Press.
- Haugeland, J. (1985). *Artificial intelligence: The very idea*. Cambridge, MA: MIT Press.
- Heidegger, M. (1977). *The question concerning technology and other essays* (W. Lovitt, Trans.). New York: Harper & Row.
- Hoffman, G. & Weinberg, G. (2011). Interactive improvisation with a robotic marimba player. *Autonomous robots, 31*(2–3), 133–153.
- Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An eternal golden braid*. New York, NY: Penguin.
- Hofstadter, D. R. (2001). Staring Emmy straight in the eye—And doing my best not to flinch. In D. Cope (Ed.), *Virtual music: Computer synthesis of musical style* (pp. 33–82). Cambridge, MA: MIT Press.
- Hui, F. (2016). AlphaGo games—English. *DeepMind*.
<https://deepmind.com/research/alphago/alphago-games-english/>
- Johnson, D. G. (2006). Computer systems: Moral entities but not moral agents. *Ethics and Information Technology, 8*(4), 195–204.

- Lewis, S. C., & Westlund, O. (2015). Big data and journalism: Epistemology, expertise, economics, and ethics. *Digital journalism*, 3(3), 447–466.
- Lyotard, J. F. (1993). *The postmodern condition: A report on knowledge* (G. Bennington & B. Massumi, Trans.). Minneapolis, MN: University of Minnesota Press.
- Majid al-Rifaie, M., & Bishop, M. (2015). Weak and strong computational creativity. In T. R. Besold, M. Schorlemmer & A. Smail (Eds.), *Computational creativity research: Towards creative machines* (pp. 37–50). Amsterdam: Atlantis Press.
- McFarland, M. (2016). What AlphaGo’s sly move says about machine creativity. *The Washington Post* (15 March).
https://www.washingtonpost.com/news/innovations/wp/2016/03/15/what-alphagos-sly-move-says-about-machine-creativity/?utm_term=.0c8281af53c9
- Metz, C. (2016a, March 11). The sadness and beauty of watching Google’s AI play go. *Wired*. <https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/>
- Metz, C. (2016b, March 16). In two moves, AlphaGo and Lee Sedol redefine the future. *Wired*. <https://www.wired.com/2016/03/two-moves-alphago-lee-sedol-redefined-future/>
- Metz, C. (2016c, March 10). Google’s AI wins a pivotal second game in match with go grandmaster. *Wired*. <http://www.wired.com/2016/03/googles-ai-wins-pivotal-game-two-match-go-grandmaster/>
- Montal, T., & Reich, Z. (2016). I, robot. you, journalist. Who is the author?: Authorship, bylines and full disclosure in automated journalism. *Digital journalism*, 5(7), 829–849. <http://dx.doi.org/10.1080/21670811.2016.1209083>
- Painting Fool. (2017). *The painting fool*. <http://www.thepaintingfool.com/>
- Plato. (1982). *Phaedrus* (H. N. Fowler, Trans.). Cambridge, MA: Harvard University Press.
- Ricœur, P. (2007). *Reflections on the just* (D. Pellauer, Trans.). Chicago: University of Chicago Press.
- Searle, J. (1984). *Mind, brains and science*. Cambridge, MA: Harvard University Press.

- Simonite, T. (2016, June 8). Ok computer, write me a song. *MIT technology review*.
<https://www.technologyreview.com/s/601642/ok-computer-write-me-a-song/>
- Smith, S. V. (2015, May 20). An NPR reporter raced a machine to write a news story. Who won? *Morning Edition*.
<http://www.npr.org/sections/money/2015/05/20/406484294/>
- Steiner, C. (2012, July 27). Can creativity be automated? Computer algorithms have started to write news stories, compose music, and pick hits. *MIT Technology Review*. <https://www.technologyreview.com/s/428437/can-creativity-be-automated/>
- Turing, A. (1999). Computing machinery and intelligence. In P. A. Meyer (Ed.), *Computer media and communication: A reader* (pp. 37–58). Oxford: Oxford University Press.
- Walmsley, J. (2012). *Mind and machine*. New York: Palgrave Macmillan.
- Whitby, B. (2011). On computable morality: An examination of machines as moral advisors. In M. Anderson & S. L. Anderson (Eds.), *Machine ethics* (pp. 138–150). Cambridge: Cambridge University Press.
- BENJAMIN, W. O narrador: considerações sobre a obra de Nikolai Leskov. In: **Obras Escolhidas**: magia, técnica, arte, política. V 1. 8. ed. São Paulo: Brasiliense, 2012.
- HALL, Stuart et al. A produção social das notícias: o “mugging” nos media. In: TRAQUINA, Nelson (org.). **Jornalismo**: questões, teorias e estórias. 2.ed. Lisboa: Vega, 1999.
- LAGO, C.; BENETTI, M. (Org.). **Metodologia da Pesquisa em Jornalismo**. 3a. ed. Petrópolis: Vozes, 2010.
- LIMA, Edvaldo Pereira. O jornalismo literário e a academia no Brasil: fragmentos de uma história. **Famecos**, v. 23, n. 1, p. 1-19, 2016.
- TRAQUINA, N. (Org.). **Jornalismo**: questões, teorias e “estórias”. Florianópolis: Insular, 2016.