



COMUNICAÇÃO MIDIÁTICA.

ISSN: 2236-8000

v. 20, n. 1, p. 199-214, jan.-jun. 2025

DOI: <https://doi.org/10.5016/mrtzg452>

O que ELIZA tem a nos ensinar sobre Literacia?

¿Qué nos puede enseñar ELIZA sobre la alfabetización en IA?

What does ELIZA have to teach us about IA literacy?

Alexandre FARBIARZ

Doutor em Design pela PUC-Rio, Professor Titular do Departamento de Comunicação Social da Universidade Federal Fluminense UFF/Brasil) e Professor permanente do Programa de Pós-Graduação em Mídia e Cotidiano PPGMC/UFF). E-mail: alexandrefarbiarz@id.uff.br

Lucas COARACY

Doutorando do Programa de Pós-Graduação em Mídia e Cotidiano da Universidade Federal Fluminense (PPGMC/UFF). E-mail: lucaslc@id.uff.br

Enviado em: 19 jun. 2025

Aceito em: 27 ago. 2025

RESUMO

Este artigo integra a pesquisa de doutorado em curso no PPGMC/UFF, intitulada “Um mundo (des)informado por prompts: o cotidiano entrelaçado em IAGens”. Investigamos os usos e sentidos (re)produzidos sobre IAs Generativas no cotidiano acadêmico do Ensino Superior, considerando a desinformação presente nos conteúdos gerados por e sobre elas. Analisamos o caso do programa ELIZA como objeto de estudo e discutimos como essa reflexão contribui para o conceito de LiteraciaIA: uma postura curiosa e crítica diante das IAs.

Palavras-chave: *Inteligência Artificial; Comunicação; Educação.*

RESUMEN

Este artículo forma parte de la investigación doctoral en curso en PPGMC/UFF, titulada "Un mundo (des)informado por prompts: la vida cotidiana entrelazada con las IA". Investigamos los usos y significados (re)producidos sobre las IA generativas en la rutina académica de la educación superior, considerando la desinformación presente en el contenido generado por y sobre ellas. Analizamos el caso del programa ELIZA como objeto de estudio y discutimos cómo esta reflexión contribuye al concepto de alfabetización en IA: una postura curiosa y crítica hacia las IA.

Palabras-clave: *Inteligencia Artificial; Comunicación; Educación.*

ABSTRACT

This article is part of the ongoing doctoral research at PPGMC/UFF, entitled "A world (mis)informed by prompts: everyday life intertwined with AIs". We investigate the uses and meanings (re)produced about Generative AIs in the academic routine of Higher Education, considering the misinformation present in the content generated by and about them. We analyze the case of the ELIZA program as our object of study and discuss how this reflection contributes to the concept of AI Literacy: a curious and critical stance toward AIs.

Keywords: *Artificial Intelligence; Communication; Education.*

Introdução

O mundo contemporâneo lida com o que podemos chamar de “explosão” das Inteligências Artificiais (IA). Em especial, desde o lançamento do ChatGPT-3 da OpenAI em 2022, a modalidade Inteligência Artificial Generativa (IAGen) tem se tornado extremamente popular. Tais tecnologias avançam a ponto de se tornarem onipresentes em nosso cotidiano (Santaella, 2023), pois mesmo que não estejamos lidando diretamente com alguma IA ou IAGen, essas são continuamente inseridas em sistemas que permeiam e/ou governam nossas vidas. Embora inicialmente utilizada em ferramentas digitais e serviços cotidianos, como motores de busca, assistentes virtuais e publicidade direcionada, a IA tem sido expandida para aplicações que influenciam diretamente aspectos cruciais da vida humana, como em tomada de decisões de tribunais (Coeckelbergh, 2023).

Neste cenário, IAs são difundidas enquanto tecnologias supostamente “neutras”, pautadas na perspectiva matemática de otimização do mundo (McQuillan, 2022). Concordamos com o que aponta Freire (2020, p. 68), que “Quem fala de neutralidade são precisamente os que temem perder o direito de usar sua ineutralidade em seu favor”. As IAGens são continuamente colocadas sob o holofote por corporações e pela imprensa, valorizadas pela medida do quão fidedignamente ou variadamente produzem seus conteúdos. Essa visão instrumental esconde as múltiplas problemáticas em torno de IAs e IAGens, como intensificação de desinformação; dependência na extração predatória de dados; extraordinários custos ambientais; precarização de trabalhadores em seus treinamentos, dentre outros (Crawford, 2021; Grohmann, 2020).

Entendemos que, em contraponto a este movimento *hype* pelas IAs, precisamos formar cidadãos que estejam dispostos a refletir, criticar ou até utilizar IAs para além do que são usualmente difundidas ou promovidas. Em outras palavras, que estejamos atentos a vieses, métodos de dominação, estratégias de desinformação, interesses políticos e demais considerações possíveis acerca das IAs. Para isso, acreditamos que é fundamental o desenvolvimento continuado de um letramento midiático crítico.

O compromisso com o letramento midiático crítico envolve estimular e capacitar os sujeitos a refletirem sobre as condições de produção desses enunciados midiáticos, desvelarem sentidos ocultos nas narrativas consumidas, indagar a si mesmos sobre as possíveis consequências da veiculação dessas mensagens, dentre muitos outros encaminhamentos para os quais a consciência crítica sobre esses fenômenos sociais pode levar o sujeito e a coletividade (Bastos, 2024, p. 157).

Nesse sentido, o estímulo ao pensamento reflexivo crítico demarca uma postura mais engajada com aspectos importantes, mas muitas vezes velados, em nossas interações com os conteúdos midiáticos com os quais temos contato em nosso cotidiano. Interessam às grandes corporações o nosso consumo passivo, acrítico, de maneira superficial e corriqueira. Com o número crescente de conteúdos gerados por IAGens, que inundam a Internet em múltiplos formatos audiovisuais, estamos cada vez mais expostos a esse tipo de interação.

Além do que nos deparamos de IAGens, estamos sendo crescentemente expostos ao contato direto com IAGens – seja em seu uso para a produção de material, ou, ainda, voluntaria ou involuntariamente, como interlocutores aos *chatbots*. Hoje, não somente IAGens *chatbots* são comercializados enquanto “amigos”, “professores” e até “parceiros românticos”, como também são utilizados como assistentes virtuais e perfis em redes sociais, sendo mascarados como pessoas reais e interagindo com outros usuários.

Nesse contexto, argumentamos que uma atualização das competências do letramento midiático crítico, voltadas a um mundo com tantas IAs, é possível e necessária. Assim, nós a cunhamos como literacIA, uma junção do conceito de literacia ao contexto de Inteligência Artificial (IA). Nesse sentido, a literacIA constitui um movimento de apropriação crítico-reflexiva da IA, inspirada primariamente nos ensinamentos de Paulo Freire (Freire, 2020; Freire; Guimarães; 2013).

Por isso, eu insisto em dizer: a crítica nossa tem que ser política, e não tecnológica. A posição em que eu me situo, portanto, é esta: eu não sou contra o computador; o fundamental seria nós podermos programar o computador. É a questão do poder: é saber a serviço de quem ele é programado para nos programar (Freire; Guimarães, 2013, p. 92).

Embora Freire não tenha escrito sobre este cenário em particular, marcado pelas IAs, acreditamos que seus pensamentos nos servem para construir uma base sólida para desenvolvermos uma postura apropriada frente às IAGens. Simultaneamente, voltamos ao passado, neste artigo, para explorarmos um dos programas mais importantes e influentes na história do campo da Inteligência Artificial: ELIZA. Acreditamos que, partindo da análise desses elementos, possamos exemplificar a importância da literacIA e discuti-la em nosso cenário contemporâneo. Para isso, começamos com a explicação do que foi ELIZA.

O contexto de ELIZA

Entre os anos de 1964 a 1966, o cientista da computação do Instituto de Tecnologia de Massachusetts (MIT), Joseph Weizenbaum (1976), desenvolveu um programa de computador que provocou impactos profundos no campo da Inteligência Artificial. O programa foi considerado o primeiro chatbot e denominado ELIZA: uma IA que “conversa” com usuários. Weizenbaum (1976) escolheu o nome pois entendeu que, como a personagem Eliza da peça *Pigmalião* – de Bernard Shaw, em 1913 –, ELIZA poderia ser ensinada a “conversar” cada vez melhor. Essa inspiração no teatro não foi acidental: Weizenbaum usualmente comparava sua criação a um ator que performa um papel, seguindo apenas instruções (Natale, 2021).

O papel que ELIZA desempenhou primeiro, atribuído por seu criador, foi o de psicoterapeuta. Weizenbaum (1976) constatou que se baseou na abordagem de psicoterapia de Carl Rogers, pois, nesta linha, muitas das respostas do psicoterapeuta são reflexões das palavras do paciente para ele mesmo. Em seu mecanismo de funcionamento, ELIZA pesquisava as palavras do texto submetido pelo usuário, procurando por palavras-chave. As respostas, então, eram direcionadas de acordo com palavras-chave ou padrões encontrados, utilizando também mensagens pré-configuradas, como “Entendi” ou “Continue” (Natale, 2021). Desse modo, nada do que ELIZA era se configurava como um acidente: ela foi feita para parecer plausível, para sustentar uma ilusão de inteligência.

ELIZA era um programa "burro". Ele conseguia reconhecer as sequências de caracteres que compõem as palavras, mas não sabia o significado de suas comunicações ou daquelas que recebia. Para ELIZA, a sequência D-E-P-R-I-M-I-D-O evocava um de um conjunto de possíveis prefixos "clichê", mas o programa não tinha mais conhecimento sobre o que significa estar deprimido (Turtle, 2005 [1984], p. 42, tradução nossa¹).

Essa ilusão demonstrou ser particularmente eficaz, pois Weizenbaum (1976) se surpreendeu com o quanto que os usuários que “conversavam” com ELIZA estavam dispostos a entrar em assuntos íntimos. Em uma famosa anedota, Weizenbaum (1976) conta

¹ ELIZA was a “dumb” program. It could recognize the character strings that make up words, but it did not know the meaning of its communications or of those it received. To ELIZA, the string D-E-P-R-E-S-S-E-D called up one of a set of possible “boilerplate” prefixes, but the program did not have any further knowledge about what it means to be depressed.

como sua secretária, que acompanhou o processo de criação de ELIZA, pediu que ele saísse da sala, pois queria privacidade enquanto conversava com ELIZA.

Para o campo da IA, em especial no desenvolvimento de *chatbots*, ELIZA virou uma grande referência. Por mais que possamos creditar o mérito de Weizenbaum ao aspecto técnico da criação de ELIZA, é no “sucesso da ilusão” que o programa se destacou. Ter isso em mente é de crucial importância para nossos estudos em literacIA. Seria um erro considerar que, no avanço das IAs – em especial as IAGens – houve apenas um aprofundamento nas operações matemáticas. Se hoje temos *chatbots* que parecem mais “inteligentes” do que nunca, isso se deve também às decisões tomadas pelos seus programadores, designers e demais envolvidos com o objetivo de aprimorar a ilusão de inteligência (Natale, 2021).

Ainda, o contexto em que tais tecnologias são divulgadas, e o que se fala delas, é de grande importância. De acordo com a pesquisa de Martin (1993), que examinou matérias jornalísticas nos Estados Unidos entre 1941 e 1960, um segmento significativo da população via os computadores como máquinas pensantes extraordinárias - percepção intensamente moldada pela cobertura midiática do período. Os jornalistas, impulsionados pelo otimismo vigente, descreviam os avanços da computação com entusiasmo exagerado. Desse modo, na busca de ampliar seu público, profissionais da imprensa frequentemente hiperbolizavam os avanços técnicos e recorriam a metáforas simplificadoras – e imprecisas – que equiparavam o funcionamento das máquinas à cognição humana (Martin, 1993).

Essa tendência, porém, não se limitava ao jornalismo. Como destaca Natale (2021), cientistas da época notaram que, ao descrever as máquinas com atributos de “inteligência”, o público tendia a interpretá-las literalmente como entidades pensantes. Isso levou parte da comunidade acadêmica a inflar as capacidades reais das tecnologias e suas projeções futuras, estratégia que visava tanto captar atenção quanto garantir financiamento para pesquisas (Natale, 2021). Nesse contexto, até mesmo termos técnicos eram substituídos por analogias que sugeriam similaridades com a mente humana. Diante desse cenário, ELIZA – programa de processamento de linguagem básico – foi percebido como demonstração de inteligência genuína, e não como um mero artifício ilusório. Este fenômeno foi concebido como Efeito ELIZA:

Como Weizenbaum reconheceu perceptivamente, as formas como o funcionamento de ELIZA foram narradas por outros cientistas e na imprensa contribuíram para fortalecer a “metáfora do computador”, pela qual a capacidade

do software de criar a aparência de inteligência foi trocada pela inteligência em si. Embora ele tivesse concebido e programado ELIZA com a expectativa de que ajudaria as pessoas a dispensar a aura mágica dos computadores, as narrativas que emergiram do escrutínio público de sua invenção estavam, na verdade, reforçando as mesmas metáforas que ele pretendia dissipar (Natale, 2021, p. 61, tradução nossa²).

Nesse contexto, narrativas que carregavam grande entusiasmo às atuais e futuras capacidades máquinas se tornaram comuns. Weizenbaum (1976) criticou veementemente, considerando perversa, a posição de autores que se demonstraram otimistas diante de uma possível automatização do trabalho terapêutico através de ELIZA. Embora tal desfecho não tenha se concretizado, a mera concepção acrítica desta possibilidade revelara o anseio de uma generalizada implementação de “incríveis máquinas pensantes” em setores da sociedade.

O efeito Eliza é, portanto, poderoso e difícil, em alguns casos talvez impossível de resistir. Sua operacionalização serve bem a certas partes – cientistas cujas oportunidades de financiamento se beneficiam da deturpação de seu trabalho; tecnólogos cujos negócios e vendas se beneficiam do fascínio que ele confere aos seus produtos (Dillon, 2020, p. 20, tradução nossa³).

O mundo viu, nas décadas seguintes, uma desaceleração nas pesquisas em IA, cujo período foi apelidado de “inverno da IA” (Russell; Norvig, 2022). No entanto, hoje em dia, com a explosão de popularidade das IAs, argumentamos que há novos entrelaçamentos da IA com desinformação, incluindo o perigo de ressurgimento do Efeito ELIZA.

IA(Gen) e desinformação

Em 2017, um grupo de pesquisadores encerrou um experimento em IA quando o *chatbot* passou a se comunicar com outros *chatbots* utilizando mensagens em códigos incompreensíveis para os humanos. Tal acontecimento resultou em notícias que explicavam o incidente de maneira preocupante, incitando a possibilidade de um futuro distópico em

² As Weizenbaum perceptively recognized, the ways ELIZA’s functioning was narrated by other scientists and in the press contributed to strengthening the “computer metaphor,” by which software’s capacity to create the appearance of intelligence was exchanged for intelligence itself. Although he had conceived and programmed ELIZA with the expectation that it would help people dismiss the magical aura of computers, the narratives that emerged from public scrutiny of his invention were actually reinforcing the very same metaphors he intended to dispel.

³ The Eliza effect is therefore powerful and difficult, in some cases perhaps impossible, to resist. Its operationalization serves certain parties well – scientists whose funding opportunities benefit from its misrepresentation of their work; technologists whose businesses and sales benefit from the allure it gives to their products.

que máquinas se juntam contra os humanos. Em realidade, a empresa tomou essa decisão somente porque tais códigos atrapalhariam a conversa de *chatbots* com humanos, o único objetivo do projeto. Segundo Natale (2019), houve falta de clareza na imprensa, uma confusão nas dimensões técnicas e um sensacionalismo quanto a um software inofensivo.

No ano de 2023, em um episódio que ganhou notoriedade mundo afora como um dos primeiros preocupantes usos do ChatGPT, o advogado Dr. Schwartz submeteu à corte um documento repleto de informações falsas, produzido pela IAGen. O advogado afirmou que não pensou que os resultados fornecidos pelo programa poderiam ser falsos (Maruf, 2023). Trabalhando sob esta suposição, o advogado, ainda, perguntou ao próprio programa se as informações que ele havia fornecido eram verídicas, confiando posteriormente na resposta afirmativa fornecida pela IAGen.

Em 2024, notícias sobre uma suposta inteligência enganadora em uma IA apareceram: “Novo ‘ChatGPT’ mentiu e manipulou para não ser desligado” (Ribeiro, 2024); “ChatGPT pego mentindo para desenvolvedores: Novo modelo de IA tenta se salvar de ser substituído e desligado.” (The Economic Times, 2024, tradução nossa⁴). O que nem sempre foi claramente exposto, porém, foi como o experimento consistia de uma série de testes controlados, justamente incentivando e avaliando a capacidade de subterfúgio desse modelo (Meinke *et al.*, 2024). Portanto, ainda que surpreendente, a situação não foi a de um modelo que saiu do controle, ainda que assim possa parecer. Em meio a essa notícia, publicações alarmistas surgiram em resposta - como um vídeo de 1,3 milhões de visualizações publicado no Facebook (Capel, 2024), encarando o ocorrido como uma evidência de que a IA havia ganho consciência. Além disso, há vídeos que não somente erroneamente atribuem características humanas à IA, como também atribuem poderes mágicos:

[...] falsos poderes de oráculo também são atribuídos a ferramentas de IA que geram imagens a partir de texto. Vinte dos vídeos analisados, somando mais de 190 milhões de visualizações, pediam para que essas tecnologias previssem o futuro ou revelassem novas versões para fatos históricos, reforçando teorias conspiratórias (Rudnitzki; Barbosa, 2023, para. 13).

Por diversas razões, como a complexidade técnica do tema e os interesses financeiros envolvidos (seja no pânico ou na empolgação), há uma distância entre o público e o conhecimento sobre IAs, preenchida muitas vezes por desinformação. No entanto, focar

⁴ ChatGPT caught lying to developers: New AI model tries to save itself from being replaced and shut down.

somente no que concerne ao conhecimento de um indivíduo acerca desta tecnologia não é nosso objetivo. Além disso, ao invés de representar apenas uma parcela deste problema multifacetado, devemos nos atentar às demais questões político-sociais, como quem é favorecido com essas práticas desinformativas. Como afirma Bucci (2022), o problema não é o de uma parcela da sociedade não entender de tecnologia, mas sim o de uma tecnologia, que não entende nada de sociedade, agir para controlá-la.

Seus desenvolvimentos ainda estão concentrados em seletos grupos no norte global, que agressivamente dominam o mercado (Verdegem, 2022). São essas poucas organizações ou corporações que, partindo de posições dominantes, desenvolvem tecnologias que diretamente afetam as vidas de pessoas em situações sociais radicalmente diferentes, muitas vezes sem a contextualização necessária (McQuillan, 2022). Portanto, apesar de serem comercializadas enquanto “ferramentas objetivas”, o conteúdo gerado pelas mesmas reflete vieses que servem a interesses de uma elite, reforçando problemas estruturais com os dados com que IAGens foram alimentadas.

É por isso que, frente a esse cenário, é urgente o desenvolvimento do que chamamos de literacIA. Uma literacia em IA é necessária não somente ao indivíduo, no que tange o seu desenvolvimento pessoal e profissional, mas também como uma questão de cidadania, para conviver com e poder usufruir de IAs sem ser “capturado” por elas como mero consumidor acrítico. É uma proposta de formação continuada que forneça possibilidade de confrontar e transformar esse contexto de IAs, que nos atinge carregando vieses de uma classe dominante.

Desenvolvendo a literacIA

Kellner e Khan (2015, p. 61) afirmam que,

[...] consideramos literacias como sendo socialmente construídas nas práticas educativas e culturais que envolvem vários discursos institucionais e pedagogias. Contra a visão autônoma que postula a literacia como estática, consideramos como algo que está em constante evolução e mudança em resposta às transformações sociais e culturais, bem como aos interesses das elites que controlam as instituições hegemônicas.

Neste contexto, compreendemos a literacIA como a prática que é desenvolvida em resposta aos avanços desenfreados de IA no mundo. Na concepção de Kellner e Kahn (2015, p. 61), a literacia não é apenas um conjunto único de competências, mas algo que possibilita

o desenvolvimento de “habilidades envolvidas na utilização eficaz de formas socialmente construídas de comunicação e de representação”. Dessa maneira, as literacias são socialmente construídas, estão em constante transformação e em fricção com os interesses hegemônicos.

No entanto, uma postura crítica às novas tecnologias, e os interesses por trás delas, não significam uma rejeição total a elas. Embora tenhamos muitas críticas acerca dos usos e (des)informações promovidas por IAGens, defendemos, como Freire (2000), que o caminho não se dá na “demonização” das novas tecnologias, mas no encontro com elas. Freire (2000) discutiu a televisão enquanto possível instrumento alienante - no entanto, isso não implica em um abandono ou a recusa desse meio. Para Freire (2000, p. 49), uma alfabetização em televisão não é lutar contra a televisão, mas “estimular o desenvolvimento da curiosidade e do pensar críticos”. Da mesma maneira, argumentamos que um afastamento em relação à IA pode levar a um pânico infundado. Para evitar isso, nos interessa uma postura curiosa: “curiosidade como inquietação indagadora, como inclinação ao desvelamento de algo, como pergunta verbalizada ou não, como procura de esclarecimento, como sinal de atenção que sugere alerta” (Freire, 2011, p. 23). Assim, uma aproximação curiosa de seu funcionamento tem seu valor não somente enquanto em abertura às possibilidades de seus usos, mas também para desmistificá-la:

[...] máquinas são feitas para se comportar de maneiras maravilhosas, muitas vezes o suficiente para deslumbrar até o observador mais experiente. Mas uma vez que um programa em particular é desmascarado, uma vez que seu funcionamento interno é explicado em linguagem suficientemente simples para induzir o entendimento, sua mágica se desintegra; ele se revela como uma mera coleção de procedimentos, cada um bastante compreensível (Weizenbaum, 1976, p. 36, tradução nossa⁵).

Entendermos que as capacidades de produções sofisticadas de IAGens não são “evidência de inteligência da máquina” nos torna mais atentos às armadilhas possíveis em seu uso – como, por exemplo, no que tange a acreditar em algo escrito pela máquina só porque “parece plausível”. Essa compreensão é indispensável para que possamos nos posicionar de forma crítica diante, não somente às IAs e IAGens, mas às inovações tecnológicas de uma forma geral.

⁵ [...] machines are made to behave in wondrous ways, often sufficient to dazzle even the most experienced observer. But once a particular program is unmasked, once its inner workings are explained in language sufficiently plain to induce understanding, its magic crumbles away; it stands revealed as a mere collection of procedures, each quite comprehensible.

Em resumo, entendemos que o desenvolvimento da literacia compreende múltiplas dimensões, como o entendimento da não-neutralidade da IA; reconhecimento de perigos e potências da IA; curiosidade em compreender o método de funcionamento da IA e abertura à possibilidade de uso da IA. Estando postos esses elementos, podemos tentar responder à pergunta que intitula este artigo:

O que ELIZA tem a nos ensinar sobre literacia?

Em primeiro lugar: nada. ELIZA, um software não-inteligente e não-consciente, não é capaz de nos “ensinar” nada. A pergunta é uma provocação que deliberadamente utiliza os artifícios que criticamos no decorrer do artigo: a supervalorização das habilidades de um programa, através de uma antropomorfização sob o efeito ELIZA. Essa antropomorfização pode ocorrer em diferentes graus, incluindo em detalhes sutis. Por vezes, como observou Weizenbaum (1966), pessoas se referem a ELIZA como “ela”. Relacionado a isso, vale mencionar a relevância do filme *Ela*, de 2013, que retrata um “romance” entre um homem e uma IA – revelando a antropomorfização já no nome da obra. Seja na realidade ou na ficção, os modos como são apresentadas IAs e como interagirmos com elas incorporam escolhas deliberadas de design. Essas, por sua vez, carregam as visões de mundo de seus criadores.

Para Dillon (2020), o próprio Weizenbaum (1966) é também responsável por perpetuar o efeito ELIZA que ele tanto criticou, seja em sua retórica e na escolha de nomeação do programa. Segundo o autor, Weizenbaum (1966) acriticamente reproduziu questões de poder na objetificação feminina e na divisão de trabalho. Isto é refletido também nos dias atuais, em contextos de IAs assistentes virtuais: vozes femininas são mais vezes incorporadas em programas que refletem propósitos assistivos, subservientes, enquanto vozes masculinas são mais vezes aplicadas em programas que oferecem conselhos em posição de autoridade (Dillon, 2020). Isso é somente um dos vários aspectos que nos demonstram como sempre há sentidos que pautam (e são reforçados por) tecnologias. Portanto, retomamos que as análises sobre ELIZA e seu contexto não somente nos são úteis para compreender a história do campo de IA, mas são também estímulos ao desenvolvimento de uma literacia.

Embora ELIZA possa parecer primitivo hoje, ele encapsula muitas das decisões de design que continuam a se manifestar nos sistemas que usamos, isso é, decisões

sobre como humanos e máquinas interagem e até que ponto a computação deve governar nossas vidas e mentes. [...] Embora possamos não pensar em ELIZA como IA hoje, mesmo que essa suposição equivocada tenha sido feita no passado, devemos estar cientes das racionalidades instrumentais semelhantes incorporadas ao aprendizado de máquina e aos sistemas de IA atuais (Berry, 2023, p. 19, tradução nossa⁶).

Evidentemente, compreender os mecanismos de funcionamento de ELIZA não é o suficiente para entender como funcionam IAGens *chatbots*. De fato, estamos lidando hoje com níveis de complexidade muito elevados. Porém, assim como investigar o código por trás de ELIZA pode revelar o que outrora parecia inteligência, estudar os métodos probabilísticos de funcionamento de Grandes Modelos de Linguagem (*LLMs*) pode desvelar as operações estatísticas que outrora pareciam consciência. A história do programa ELIZA pode ser utilizada, por exemplo, por educadores como um disparador com seus educandos para os questionamentos que possamos desenvolver com as IAs atuais. É instigando a postura curiosa e crítica, não concebendo a IA como “neutra” nem “demônio”, que podemos estimular apropriações e reflexões sobre IAs que tenham pautas cidadãs como prioridade – diferentemente de um consumo acrítico, imerso em desinformação, que carrega os interesses da classe dominante.

⁶ Although ELIZA might appear primitive today, it encapsulates many of the design decisions that continue to be manifest in the systems we use, namely, decisions concerning how humans and machines interact and the extent to which computation should be allowed to govern our lives and minds. [...] Although we might not think of ELIZA as AI today, even if that mistaken assumption was made in the past, we should be aware of the similar instrumental rationalities embedded in machine learning and today’s AI systems.

REFERÊNCIAS

- BASTOS, L. L. A. **Os diálogos entre sujeitos jogadores e jogos digitais**: Investigação a partir de uma perspectiva comunicacional, cotidiana e do letramento midiático crítico. Orientador: Alexandre Farbiarz. 2024, 404 f. Tese (Doutorado em Mídia e Cotidiano) – Instituto de Arte e Comunicação Social, Universidade Federal Fluminense, Niterói, 2024. Disponível em: <https://app.uff.br/riuff/hogin/handle/1/33335>. Acesso em: 25 jul. 2025.
- BERRY, D. M. The limits of computation: Joseph Weizenbaum and the ELIZA chatbot. **Weizenbaum Journal of the Digital Society**, v. 3, n. 3, 2023. Disponível em: <https://ojs.weizenbaum-institut.de/index.php/wjds/article/view/106>. Acesso em: 25 jul. 2025.
- BUCCI, E. Reflexões sobre a ética & inteligências maquínicas. In: SANTAELLA, L. (org.). **Simbioses do humano & tecnologias**: impasses, dilemas, desafios. São Paulo: Edusp, 2022, p. 91–101.
- CAPEL, C. **As IAs**. Elas já são conscientes. [Vídeo]. Facebook, 11 dez. 2024. Disponível em: <https://www.facebook.com/watch/?mibextid=wwXIfr&v=1669685363580401&rdid=sVGW6O0Fx0xWFUFn>. Acesso em: 25 jul. 2025.
- COECKELBERGH, M. **Ética na inteligência artificial**. São Paulo: Ubu Editora, 2023.
- CRAWFORD, K. **Atlas of AI**: Power, politics and the planetary costs of artificial intelligence. New Haven: Yale University Press, 2021.
- DILLON, S. The Eliza effect and its dangers: from demystification to gender critique. **Journal for Cultural Research**, v. 24, n. 1, p. 1–15, 2020. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/14797585.2020.1754642>. Acesso em: 25 jul. 2025.
- FREIRE, P. **Pedagogia da autonomia**: Saberes necessários à prática educativa. Rio de Janeiro: Paz e Terra, 2011.
- FREIRE, P. **Pedagogia da indignação**: Cartas pedagógicas e outros escritos. São Paulo: Editora UNESP, 2000.
- FREIRE, P. **Pedagogia do oprimido**. 74. ed. Rio de Janeiro: Paz e Terra, 2020.
- FREIRE, P.; GUIMARÃES, S. **Educar com a mídia**: Novos diálogos sobre educação. Rio de Janeiro: Paz e Terra, 2013.
- GROHMANN, R. Plataformização do trabalho: entre a dataficação, a financeirização e a racionalidade neoliberal. **Revista Eletrônica Internacional de Economia Política da Informação, da Comunicação e da Cultura**, v. 22, n. 1, p. 106–122, 2020. Disponível em: <https://periodicos.ufs.br/epic/article/view/12188/10214>. Acesso em: 25 jul. 2025.

KELLNER, D.; KAHN, R. Reconstruindo a tecnoliteracia: uma abordagem de múltiplas literacias. **Comunicação & Educação**, v. 20, n. 2, p. 57-82, 2015.

MARTIN, C. D. The myth of the awesome thinking machine. **Communications of the ACM**, v. 36, p. 120-133, 1993.

MARUF, R. Advogado pede desculpas por falsas citações judiciais criadas pelo ChatGPT. **CNN Brasil**, 28 mai. 2023. Disponível em: <https://www.cnnbrasil.com.br/internacional/advogado-pede-desculpas-por-falsas-citacoes-judiciais-criadas-pelo-chatgpt/>. Acesso em: 25 jul. 2025.

MCQUILLAN, D. **Resisting AI**: An anti-fascist approach to artificial intelligence. Bristol: Bristol University Press, 2022.

MEINKE, A. *et al.* **Frontier Models are Capable of In-context Scheming**. arXiv preprint arXiv:2412.04984, dez. 2024. Disponível em: <https://arxiv.org/abs/2412.04984>. Acesso em: 25 jul. 2025.

NATALE, S. **Deceitful Media**: Artificial Intelligence and Social Life after the Turing Test. New York: Oxford University Press, 2021.

NATALE, S. If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA. **New Media & Society**, v. 31, n. 3, p. 712-728, 2019. Disponível em: <https://journals.sagepub.com/doi/10.1177/1461444818804980>. Acesso em: 25 jul. 2025.

RIBEIRO, M. Novo “ChatGPT” mentiu e manipulou para não ser desligado. **Hypescience**, 9 dez. 2024. Disponível em: <https://hypescience.com/novo-chatgpt-mentiu-e-manipulou-para-nao-ser-desligado>. Acesso em: 25 jul. 2025.

RUDNITZKI, E.; BARBOSA, J. Virais no TikTok atribuem poderes mágicos à IA e espalham desinformação sobre a tecnologia. **Aos Fatos**, 26 ago. 2023. Disponível em: <https://www.aosfatos.org/noticias/tiktok-inteligencia-artificial-desinformacao/>. Acesso em: 25 jul. 2025.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence**: A modern approach. 4 ed. Harlow: Pearson Education, 2022.

SANTAELLA, L. Cultura de plataforma e desafios à criatividade. In: SANTAELLA, L.; RIBEIRO, D. M.; ALZAMORA, G. (org.). **Pensar a Inteligência Artificial**: Cultura de plataforma e desafios à criatividade. Belo Horizonte: Fafch/Selo PPGCOM/UFMG, 2023, p. 19-34.

THE ECONOMIC TIMES. ChatGPT caught lying to developers: New AI model tries to save itself from being replaced and shut down. **The Economic Times**, 7 dez. 2024. Disponível em: <https://economictimes.indiatimes.com/magazines/panache/chatgpt-caught-lying-to-developers-new-ai-model-tries-to-save-itself-from-being-replaced-and-shut-down/articleshow/116077288.cms>. Acesso em: 25 jul. 2025.

TURKLE, S. **The second self**: computers and the human spirit. 20th anniversary ed. Cambridge: The MIT Press, 2005 [1984].

VERDEGEM, P. Dismantling AI capitalism: The commons as an alternative to the power concentration of Big Tech. **AI & Society**, v. 39, p. 727-737, 2022.

WEIZENBAUM, J. **Computer Power and Human Reason**: From Judgement to Calculation. New York: W. H. Freeman and Company, 1976.

WEIZENBAUM, J. ELIZA—a computer program for the study of natural language communication between man and machine. **Communications of the ACM**, v. 9, n. 1, p. 36-45, 1966. Disponível em: <https://dl.acm.org/doi/10.1145/365153.365168>. Acesso em: 25 jul. 2025.

BIOGRAFIA DOS AUTORES

Alexandre Farbiarz

Doutor em Design (PUC-Rio), Mestre em Educação e Linguagem (USP) e Mestre em Design (PUC-Rio). Professor Titular do Curso de Jornalismo e professor permanente do Programa de Pós-Graduação em Mídia e Cotidiano (PPGMC) do Instituto de Arte e Comunicação Social (IACS) da Universidade Federal Fluminense (UFF). Coordenador do grupo de pesquisas Educação para as Mídias em Comunicação (educ@midias.com/PPGMC/UFF) e do grupo de pesquisas Design na Leitura de Sujeitos e Suportes em Interação (DeSSIn/PPG Design/PUC-Rio). Pesquisa nas áreas de Comunicação, Design e Educação, com ênfase em relações discursivas, mídias e ensino aprendizagem; atuando principalmente nos temas: Educação Crítica para as Mídias, Mídia-Educação, Literacia Midiática, Educomunicação, Educação a Distância, Jogos e Educação, Gamificação, Comunicação Visual, Discurso e Linguagem. **E-mail de contato:** alexandrefarbiarz@id.uff.br

Lucas Lima Coaracy

Doutorando no Programa de Pós-graduação em Mídia e Cotidiano da Universidade Federal Fluminense (PPGMC-UFF). Mestre no Programa de Pós-graduação em Mídia e Cotidiano da Universidade Federal Fluminense (PPGMC-UFF). Bacharel e licenciado em Psicologia pela Universidade Federal Fluminense (2016). Membro do grupo de pesquisa Educação para as Mídias em Comunicação - educ@midias.com. **E-mail de contato:** lucaslc@id.uff.br