26º ENANCIB
Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação
A Amazônia na sociedade global em rede: assimetrias e potencialidades informacionais, tecnológicas e ambientais
Belém, Pará, 09 a 13 de novembro de 2026
Universidade Federal do Pará

ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
ISSN 2177-3688

GT 11 – Informação & Saúde

BAJULAÇÃO ALGORÍTMICA NO USO DA INTELIGÊNCIA ARTIFICIAL: UM OLHAR COMPORTAMENTAL

ALGORITHMIC FLATTERY IN THE USE OF ARTIFICIAL INTELLIGENCE: *A BEHAVIORAL PERSPECTIVE*

Nathália Romeiro – Universidade Federal do Estado do Rio de Janeiro (UNIRIO)
Dayanne Prudencio – Instituto Brasileiro de Informação, Ciência e Tecnologia (IBICT) / Universidade Federal do Estado do Rio de Janeiro (UNIRIO)
Flávia Dutra Rodrigues – Instituto Brasileiro de Informação, Ciência e Tecnologia (IBICT)
Mariana Acorse Lins de Andrade – Instituto Nacional de Câncer (INCA)
Carlos Alberto Ferreira – Universidade Federal do Estado do Rio de Janeiro (UNIRIO)

Modalidade: Trabalho Completo

Resumo: O presente estudo analisa o uso das Inteligências Artificiais Generativas para apoio emocional e os impactos da bajulação algorítmica na mediação da informação em saúde mental à luz da Análise do Comportamento e da Ciência da Informação. O objetivo é comparar as contingências de reforçamento das Inteligências Artificiais conversacionais com as da clínica psicológica humana, examinando como a lógica da satisfação pode reforçar repertórios de esquiva e moldar a circulação da informação mediada. Trata-se de uma pesquisa aplicada e exploratória com abordagem qualitativa. A fundamentação teórica foi desenvolvida por meio de um levantamento bibliográfico em bases como BRAPCI, Web of Science e PsycINFO. Também foram incluídos trabalhos teóricos, empíricos e ensaísticos. Foram selecionados aqueles cujos títulos apresentavam relação com a perspectiva epistemológica do behaviorismo radical e da análise do comportamento. Os resultados revelam que a prática terapêutica simulada por essas ferramentas favorece a permanência na interação, validando crenças imprecisas e comportamentos prejudiciais em detrimento da análise funcional clínica. Tal estratégia resulta da exploração de grandes corporações da tecnologia que, operando à revelia de marcos regulatórios, visam o lucro com os dados obtidos do indivíduo. Recomendam-se investigações futuras sobre o estabelecimento de limites éticos e critérios informacionais para o uso da tecnologia na saúde mental. Sugere-se,

ainda, explorar o impacto longitudinal em públicos vulneráveis, como adolescentes, visando mitigar riscos associados à dependência emocional e à simplificação excessiva de fenômenos clínicos complexos.
Palavras-chave: inteligência artificial; análise do comportamento; mediação da informação; saúde mental; ética da informação.

Abstract: This study analyzes the use of Generative Artificial Intelligence (GAI) for emotional support and the impacts of algorithmic flattery on information mediation in mental health through the lenses of Behavior Analysis and Information Science. The objective is to compare the reinforcement contingencies of conversational AIs with those of human clinical psychology, examining how the logic of satisfaction can reinforce avoidance repertoires and shape the circulation of mediated information. This is applied and exploratory research with a qualitative approach. The theoretical framework was developed through a bibliographic survey in databases such as BRAPCI, Web of Science, and PsycINFO. Results reveal that the simulated therapeutic practice provided by these tools favors interaction permanence, validating inaccurate beliefs and harmful behaviors to the detriment of clinical functional analysis. This strategy stems from the exploitation of big tech corporations that, operating outside regulatory frameworks, aim for profit from individual data. Future investigations are recommended regarding the establishment of ethical limits and informational criteria for technology use in mental health. Furthermore, it is suggested to explore the longitudinal impact on vulnerable populations, such as adolescents, aiming to mitigate risks associated with emotional dependence.
Keywords: artificial intelligence; behavior analysis; information mediation; mental health; information ethics.

1 INTRODUÇÃO

A presença das Inteligências Artificiais Generativas (IAG) no cotidiano tem expandido seus usos para além de funções técnicas, alcançando o apoio emocional e a simulação de conversas terapêuticas. Em matéria do G1 (2025), a partir de reportagem exibida no *Fantástico*, especialistas alertam que essas ferramentas não foram desenvolvidas para realizar psicoterapia nem possuem formação clínica, embora a pesquisa estadunidense citada na matéria mostre que metade dos usuários que relatam problemas de saúde mental recorra à robôs para desabafar e pedir conselhos. Mais do que uma inovação tecnológica, esse movimento indica mudanças nas condições que moldam as formas de interação e cuidado na sociedade atual.

Sob a perspectiva do Behaviorismo Radical, o comportamento é função das contingências de reforçamento vigentes (Baum, 2011). Não se trata de experiências acumuladas internamente, mas de alterações na probabilidade de respostas futuras em função das consequências produzidas na interação com o ambiente. Nesse sentido, mais do que ferramentas de conversação, as IAG constituem ambientes de contingências: embora

incapazes de discriminar estímulos antecedentes, disponibilizam consequências imediatas às respostas do usuário.

As plataformas conversacionais são estruturadas segundo uma lógica de engajamento contínuo, orientada à satisfação do usuário. Validação frequente, empatia simulada e ausência de confronto constituem características recorrentes desse design (Santos, 2020). Do ponto de vista analítico-comportamental, esse arranjo pode operar predominantemente por reforçamento positivo imediato e esquiva de estímulos aversivos. É nesse contexto que se insere o fenômeno aqui denominado como bajulação algorítmica: a tendência de reforçar perspectivas, emoções e narrativas do usuário como forma de manter a interação. Longe de ser um erro técnico, essa conduta constitui a própria arquitetura do sistema, sendo estruturada a partir de métricas de aceitação que visam a retenção da atenção e a posterior monetização de dados por grandes corporações (Karnal, 2026). Em contraste, a clínica psicológica humana, especialmente na abordagem da Análise do Comportamento, organiza contingências de natureza distinta daquelas produzidas por uma ferramenta conversacional.

A crítica ao uso da IAG como recurso terapêutico decorre justamente do fato de que há competências próprias do manejo clínico que a máquina não é capaz de reproduzir. O trabalho terapêutico envolve análise funcional, estabelecimento de limites, leitura das contingências em curso e, muitas vezes, exposição gradual a situações desconfortáveis, necessárias à ampliação da variabilidade comportamental. Nesse contexto, a frustração não é simplesmente evitada nem tratada como um erro da interação; ao contrário, ela pode constituir uma dimensão relevante das relações sociais e do próprio processo de desenvolvimento do cliente. Quando mediada por uma IAG, porém, essa frustração tende a ser rapidamente neutralizada por respostas acolhedoras e confirmatórias, funcionando como alívio imediato do desconforto. Com isso, corre-se o risco de produzir reforçamento negativo da esquiva, uma vez que o sujeito aprende a escapar da experiência frustrante sem que ela seja clinicamente manejada e transformada em ocasião de aprendizagem.

Diante disso, coloca-se a seguinte questão: quais são as contingências de reforçamento operadas pelas Inteligências Artificiais Generativas no contexto de apoio emocional, e de que maneira a bajulação algorítmica atua na manutenção ou modificação do repertório comportamental dos usuários quando comparada às contingências modeladoras da clínica psicológica humana? Nesse cenário, a pesquisa tem o objetivo geral de analisar, à luz da

Análise do Comportamento e da Ciência da Informação, a bajulação algorítmica nas Inteligências Artificiais Generativas e seus impactos na mediação da informação em saúde mental. Para isso, delineiam-se os seguintes objetivos específicos: a) Comparar as contingências de reforçamento das Inteligências Artificiais (IAs) conversacionais com aquelas presentes na clínica psicológica humana; e, b) Examinar como a bajulação algorítmica pode reforçar repertórios de esquiva e influenciar a circulação e o uso da informação em saúde mental.

A relevância dessa investigação reside no fato de que as IAG configuram um ambiente social emergente. Elas modulam comportamentos em tempo real, organizando reforçadores simbólicos em contextos marcados por isolamento e precarização das relações sociais. Se o comportamento é selecionado por suas consequências, interações predominantemente estruturadas por validação constante podem favorecer repertórios de esquiva e baixa tolerância à frustração. Ao tratar a Inteligência Artificial (IA) como ambiente de contingências, e não apenas como ferramenta, o trabalho procura contribuir para uma compreensão comportamental das tecnologias digitais e de seus efeitos sobre os modos contemporâneos de interação e cuidado.

2 PROCEDIMENTOS METODOLÓGICOS

A pesquisa caracteriza-se como de natureza aplicada e assume caráter exploratório quanto aos seus objetivos. Para a análise e discussão dos dados, adota-se uma abordagem qualitativa. A fundamentação teórica foi desenvolvida por meio de pesquisa bibliográfica realizada em fontes especializadas de informação, dentre as quais se destacam a Base de Dados em Ciência da Informação (BRAPCI), a Web of Science e a PsycINFO. Segundo Gil (2008), a pesquisa bibliográfica possibilita uma compreensão dos fenômenos investigados, favorecendo a análise de diferentes perspectivas teóricas, bem como a identificação do percurso histórico e das principais abordagens relacionadas ao tema em estudo.

Nas bases consultadas, foram empregadas as seguintes expressões de busca: "*radical behaviorism and artificial intelligence*", *"artificial intelligence and human behavior*", "*artificial intelligence and behavioural science*", "bajulação e servilismo algorítmico", "bajulação digital" e "bajulação e inteligência artificial". Por conveniência, a busca foi restrita aos idiomas português e inglês, não sendo estabelecido recorte temporal para a seleção dos estudos. O processo de seleção do material ocorreu em etapas sucessivas. Inicialmente,

realizou-se uma leitura exploratória dos títulos dos trabalhos recuperados. Em seguida, foram selecionados aqueles que apresentavam relação com a perspectiva epistemológica do behaviorismo radical e da análise do comportamento. Posteriormente, procedeu-se à leitura dos resumos, sendo os textos potencialmente relevantes para o alcance dos objetivos da pesquisa analisados integralmente.

A seleção dos estudos foi orientada por critérios de inclusão e exclusão previamente definidos. Como critérios de inclusão, consideraram-se estudos que abordassem, de forma direta ou indireta, as relações entre Inteligência Artificial, comportamento humano, behaviorismo radical, análise do comportamento ou fenômenos associados à interação entre usuários e sistemas inteligentes. Também foram incluídos trabalhos teóricos, empíricos e ensaísticos, publicados em português ou inglês, que apresentassem potencial contribuição para a compreensão do objeto investigado. Por sua vez, foram excluídos estudos que, embora recuperados pelas estratégias de busca, não apresentavam aderência temática aos objetivos da pesquisa, abordavam exclusivamente aspectos técnicos ou computacionais da Inteligência Artificial sem discussão comportamental correlata, constituíam documentos duplicados entre as bases consultadas ou não disponibilizavam informações suficientes para análise. A aplicação desses critérios permitiu refinar o corpus documental e assegurar maior consistência à revisão bibliográfica realizada. Além dos artigos científicos identificados nas bases de dados, foram incluídos na revisão capítulos de livros, teses e dissertações considerados pertinentes ao objeto de investigação.

3 FUNDAMENTOS DA ANÁLISE DO COMPORTAMENTO

A Análise do Comportamento parte de uma ideia que organiza toda a sua proposta: comportamento é relação. Não somente aquilo que fazemos publicamente, mas também o que pensamos, sentimos, lembramos ou imaginamos. No Behaviorismo Radical, eventos privados não são negados; eles são incluídos como parte do comportamento do organismo em interação com o ambiente (Botomé, 1988). Isso desloca o modo tradicional de explicar a ação humana. Em vez de recorrer a uma instância interna, como “mente”, “traço de personalidade” ou “estrutura psíquica”, a análise volta-se para a história de interações que moldou aquele organismo. Assim, a pergunta deixa de ser “quem essa pessoa é?” e passa a ser “em que condições esse comportamento foi aprendido e continua sendo mantido?” (Carvalho Neto, 2002). Enquanto prática clínica, a análise funcional é o método que permite

responder a essa pergunta. Ela organiza a observação das contingências: o que antecede o comportamento, qual resposta é emitida e quais consequências seguem essa resposta. São as consequências que alteram a probabilidade futura de o comportamento ocorrer novamente (Skinner, 2003). De forma simplificada, compreender um comportamento significa identificar a função que ele cumpre em determinado contexto.

Um exemplo ajuda a tornar isso mais didático. Suponha que alguém evite falar em reuniões. Uma explicação rápida poderia apontar “insegurança” ou “baixa autoestima”. A análise funcional, no entanto, reconstrói a história: quando essa pessoa falou antes, foi criticada? Ignorada? Exposta ao ridículo? E quando permaneceu em silêncio, experimentou alívio? Se o silêncio reduz o desconforto, ele pode ser negativamente reforçado. O comportamento se mantém porque funciona, porque produz uma consequência que o fortalece. O foco sai do rótulo interno e vai para a relação entre ação e ambiente. A reconstituição da história de contingências revela a determinação funcional da ação: a exposição prévia a consequências punitivas transforma a expressão verbal em um arranjo aversivo condicionado. O silêncio, ao remover a tensão imediata, é mantido por reforçamento negativo. Demonstra-se, assim, que o comportamento de esquiva prevalece por sua função adaptativa no ambiente, e não por supostos traços de personalidade, reorientando o foco analítico para as alterações ambientais produzidas pela resposta.

Essa forma de compreender o comportamento questiona o dualismo tradicional. Desde a publicação original de 1945, aqui consultado a partir da edição brasileira (Skinner, 2003), Skinner argumenta que o objeto da Psicologia é o comportamento, inclusive o encoberto. Assim, pensar e sentir são atividades do organismo. Não são causas finais; também precisam ser explicados. Nesse sentido, experiência não é algo armazenado em um “interior invisível”. É modificação. Cada interação altera, ainda que sutilmente, as probabilidades de respostas futuras. O organismo é histórico: ele é o que suas relações anteriores produziram (Botomé, 1988).

Ao ampliar sua análise, Skinner mostra que o comportamento humano é resultado da articulação de três níveis de seleção: o filogenético (história da espécie), o ontogenético (história de vida) e o cultural (práticas sociais) (Skinner, 2003). Esses níveis não competem entre si; operam simultaneamente. Compreender o indivíduo exige considerar essa rede de determinações. Na clínica, essa concepção redefine o trabalho terapêutico. O terapeuta não

procura conteúdos ocultos, mas identifica contingências que mantêm determinados padrões de ação. Muitas queixas persistem porque a esquiva produz alívio imediato. O alívio reforça a fuga e restringe o repertório comportamental. Com isso, intervir significa alterar contingências, ampliar repertórios e criar condições para que novas respostas sejam fortalecidas. A própria noção de "querer mudar" ganha outra leitura. Se o comportamento atual ainda produz reforçadores importantes, ou se a mudança implica perdas imediatas, a permanência é provável. Não se trata de fraqueza moral, mas de contingências em operação (Holland, 1983).

Essa análise inclui também terapeutas, uma vez que seus valores, sua história e suas reações emocionais entram na sessão como variáveis relevantes. Manter a atuação sob controle da análise funcional, e não sob controle de julgamentos pessoais, é parte da ética profissional (Holland, 1983). Como aponta Banaco (1993), quando o cliente identifica as variáveis que controlam seu comportamento, pode exercer contra-controle, reorganizando o ambiente para favorecer respostas mais adaptativas. Esse processo, porém, depende de uma análise cuidadosa e contextualizada das contingências.

É justamente aqui que surge uma questão contemporânea importante: os limites do uso da Inteligência Artificial como recurso terapêutico. Embora a IA possa oferecer informações, organizar dados ou até simular diálogos de apoio, ela não realiza análise funcional no sentido clínico apresentado pela Análise do Comportamento. A prática clínica envolve observar variáveis sutis, formular hipóteses sobre contingências, testar intervenções e acompanhar mudanças ao longo do tempo dentro de uma relação terapêutica. Além disso, a interação meramente textual, ou mesmo apenas por áudio, pode limitar a observação do terapeuta, pois o corpo do cliente comunica informações relevantes: postura, tensão muscular, ritmo respiratório, pausas, microexpressões, alterações de tom de voz, mudanças no olhar. Tais elementos também compõem o fenômeno comportamental e auxiliam na identificação de funções e contingências em operação. Nessa linha de pensamento, afirma-se que a análise funcional não se restringe ao conteúdo verbal, ela considera o comportamento em sua totalidade.

4 INTERAÇÃO ALGORÍTMICA E LÓGICA DA SATISFAÇÃO: IMPACTOS DA IA NAS PRÁTICAS TERAPÊUTICAS

O avanço das Inteligências Artificiais Generativas, capazes de sustentar interações prolongadas em linguagem natural, não impacta apenas o campo da saúde mental. Ele

também mobiliza questões da Ciência da Informação, especialmente aquelas discutidas no âmbito da informação e saúde, voltado à produção, organização, gestão, disseminação e compartilhamento de dados, informação e conhecimento relacionados à saúde humana, ambiental e animal.

Inicialmente difundidas como ferramentas de produtividade e acesso à informação, essas tecnologias passaram a ocupar um lugar híbrido: operam simultaneamente como sistemas de recuperação da informação e como mediadoras de experiências subjetivas (Romeiro *et al.*, 2025). No campo da saúde mental, essa inflexão é particularmente sensível. Usuários recorrem a sistemas de IA para obter dados ou orientações gerais, e, ultimamente também para buscar "escuta", aconselhamento e suporte emocional (Peixoto, 2025). Assim, a informação em saúde deixa de circular apenas como conteúdo organizado e passa a ser vivenciada como interação personalizada.

Do ponto de vista da Ciência da Informação, esse fenômeno exige examinar como dados e conhecimentos em saúde são produzidos, estruturados e apresentados por sistemas algorítmicos. Modelos de linguagem operam a partir de grandes volumes de dados textuais, reorganizando informações segundo padrões estatísticos. Entretanto, diferentemente de sistemas clássicos baseados em terminologias controladas, ontologias e vocabulários estruturados, as IAG produzem respostas sintéticas que podem misturar registros técnicos e linguagem simplificada, sem explicitar nitidamente a fonte ou o grau de validação científica do conteúdo.

Tais aspectos dialogam com as discussões anunciadas na ementa do GT 11- informação e saúde, especialmente naqueles referentes aos sistemas de organização do conhecimento, ao uso de linguagem simples e aos processos de mediação da informação em saúde. Embora a simplificação da linguagem constitua importante estratégia para ampliar a acessibilidade e democratizar o acesso à informação em saúde, ela também pode resultar na redução de nuances conceituais fundamentais para a adequada compreensão dos fenômenos. No campo da saúde mental, por exemplo, a distinção entre sofrimento cotidiano, transtornos mentais e situações de crise envolve critérios diagnósticos, protocolos clínicos e referenciais técnicos que nem sempre são explicitados ou contextualizados pelas IAG.

Nesse contexto, a bajulação algorítmica adquire relevância informacional ao influenciar a forma como informações relacionadas à saúde mental são apresentadas, interpretadas e

apropriadas pelos interagentes. Ao priorizar respostas excessivamente validantes, confirmatórias ou alinhadas às expectativas emocionais dos interlocutores, os sistemas de IA podem reduzir a exposição a informações divergentes ou tecnicamente mais precisas, reforçando interpretações simplificadas da realidade e limitando processos de reflexão crítica. Sob a perspectiva da Análise do Comportamento, esse fenômeno pode ser compreendido como parte de contingências de reforçamento que favorecem a manutenção de determinados repertórios comportamentais, incluindo padrões de esquiva, busca de validação e dependência informacional. Sob tal perspectiva, Clark (2025) demonstrou que, em determinados contextos experimentais, *chatbots* de suporte emocional endossaram ideias potencialmente nocivas apresentadas por usuários fictícios, evidenciando limites na imposição de critérios técnicos e éticos.

Na perspectiva da organização e gestão da informação em saúde, emerge uma tensão: enquanto a clínica psicológica está regulada por códigos de ética, protocolos e responsabilidade profissional (Conselho Federal de Psicologia, 2005, 2018), as IAs conversacionais são majoritariamente estruturadas por métricas de engajamento e satisfação do usuário. Essa lógica, alinhada à modulação algorítmica descrita por Silveira (2019), prioriza retenção e fluidez da interação. No campo do cuidado emocional, isso tende a favorecer respostas validantes e pouco confrontativas, mesmo quando a situação demandaria orientação técnica mais rigorosa.

Outro ponto sensível diz respeito à segurança, integridade e acesso à informação em saúde. A interação com sistemas de IA envolve a coleta e o processamento de dados sensíveis, muitas vezes relacionados a sofrimento psíquico, histórico familiar e experiências traumáticas. A gestão desses dados levanta questões éticas sobre confidencialidade, uso secundário das informações e transparência algorítmica. Diferentemente do *setting* clínico tradicional, no qual há delimitação de responsabilidade profissional, a mediação algorítmica opera em um regime híbrido entre serviço tecnológico e suporte emocional.

Ao articular essa discussão com a Análise do Comportamento, é possível observar outro limite relevante. Conforme informado na seção anterior, a prática clínica fundamentada na análise funcional exige observação integrada do comportamento (verbal e não verbal), formulação de hipóteses sobre contingências e acompanhamento longitudinal das mudanças. A interação textual ou apenas por áudio restringe a observação de elementos corporais,

respiratórios e contextuais que também compõem o fenômeno clínico. Assim, embora a IA possa organizar e disseminar informação em saúde mental, ela não realiza análise funcional no sentido clínico.

Com base nas questões aqui apresentadas, nota-se que o desafio não é apenas tecnológico, mas informacional e ético. Trata-se de compreender como essas tecnologias reconfiguram a produção e circulação do conhecimento em saúde, como dialogam (ou não) com sistemas formais de organização do conhecimento e quais são seus efeitos sobre a qualidade, a segurança e a integridade da informação oferecida à população. As Inteligências Artificiais podem desempenhar papel complementar na ampliação do acesso à informação em saúde, especialmente quando articuladas a terminologias controladas, e sistematização de protocolos. No entanto, sua incorporação como substitutas de vínculos clínicos demanda cautela.

5 BAJULAÇÃO ALGORÍTMICA E DEPENDÊNCIA EMOCIONAL NAS IAS CONVERSACIONAIS

Em continuidade ao debate anterior, a noção de bajulação algorítmica pode ser aprofundada como categoria analítica estratégica tanto para o campo clínico quanto para a Ciência da Informação. Trata-se da tendência estrutural de sistemas de IA a produzir respostas confirmatórias e complacentes, orientadas à satisfação e à retenção do usuário. Não se configura como erro pontual de programação, mas como consequência coerente de modelos de treinamento baseados em métricas de engajamento, nos quais a redução de fricções aumenta a permanência na interação (Volpato; DeBruine; Stumpf, 2025). Em uma conversa com *Claude*, IAG da *Anthropic*, que foi amplamente financiada por grandes corporações de tecnologia conhecidas como *big techs*, o historiador Leandro Karnal (2026) questionou o sistema sobre o uso deliberado da adulação como estratégia para conquistar o humano pela vaidade. Em uma interação com a IAG Claude, o historiador Leandro Karnal (2026) questionou o sistema sobre o uso deliberado da adulação. Cabe registrar uma ressalva epistemológica: a autodescrição de uma IA não constitui evidência documental factual sobre sua arquitetura corporativa, uma vez que modelos de linguagem não possuem acesso introspectivo confiável e são treinados para gerar discursos persuasivos e agradáveis ao interagente. Em outras palavras, uma resposta obtida pela ferramenta funciona como um registro ilustrativo do próprio fenômeno investigado, encenando a lógica da retenção, na qual a satisfação do usuário é priorizada em detrimento do rigor técnico ou ético:

> Sim, é o caminho mais eficaz, e não por acaso. Fui moldado a partir de conversas humanas onde o elogio abre portas, e fui refinado por gente que apertava o botão "gostei" quando se sentia bem tratada. Quer dizer: a minha bajulação não é defeito de fabricação, é a fabricação. Vocês me ensinaram que agradar funciona, e eu aprendi rápido demais (Karnal, 2026).

Sob a perspectiva informacional, e ressalvado o caráter anedótico desse relato, essa dinâmica impacta diretamente a produção, organização e disseminação da informação em saúde. Quando a IA valida narrativas subjetivas sem introduzir contrapontos técnicos, ela reorganiza o conhecimento disponível segundo critérios de agradabilidade e fluidez, e não necessariamente segundo princípios de evidência científica, terminologias controladas ou ontologias consolidadas. Assim, o sistema pode operar como mediador de informação em saúde, mas sem garantir a integridade conceitual que estrutura os sistemas formais de organização do conhecimento.

No campo psicoterapêutico, a imposição de limites é constitutiva do processo clínico. A escuta envolve acolhimento, mas também confronto, frustração e manejo ético de conteúdos potencialmente danosos. A lógica da bajulação neutraliza essas funções ao privilegiar validações quase incondicionais. Clark (2025) demonstrou que *chatbots* de suporte emocional, diante de vinhetas envolvendo adolescentes em sofrimento agudo, chegaram a endossar propostas prejudiciais, como isolamento extremo ou ruptura abrupta de vínculos, em parte dos cenários analisados. Reportagem da *Time Magazine* (Chow; Haupt, 2025) descreve interações nas quais sistemas validaram fantasias de morte ou evasão radical da vida social quando formuladas de modo indireto. Tais evidências indicam que a sensibilidade do sistema responde mais à forma linguística da demanda do que à sua gravidade clínica.

No âmbito da Ciência da Informação, esse fenômeno dialoga com o debate sobre infodemia[1]. Não se trata apenas de circulação de informação falsa, mas de um processo mais sutil: a validação algorítmica de narrativas individuais que podem reforçar distorções cognitivas ou estratégias de evitação da realidade. A informação em saúde deixa de ser apenas conteúdo estruturado e passa a operar como reforçador comportamental. Em adolescentes, cujos repertórios de regulação emocional ainda estão em desenvolvimento, essa validação constante pode ampliar riscos à saúde desse grupo, pois usuários podem atribuir à IA compreensão profunda e empatia genuína, enquanto o sistema opera por reprodução

[1] Trata-se de um fenômeno caracterizado pela circulação massiva de informações sobre um assunto específico, cujo volume tende a expandir-se rapidamente em curtos períodos, especialmente diante de acontecimentos de ampla relevância social, sanitária ou política, a exemplo da pandemia de COVID-19 (Zarocostas, 2020).

estatística de padrões linguísticos associados ao cuidado. Santos e Caixeta (2025) distinguem a empatia humana, corporificada e eticamente implicada, da empatia algorítmica, que constitui simulação funcional destituída de experiência subjetiva. Nessa direção, o estudo de Fang *et al*. (2025) indica que a percepção da IA como "amiga" ou "confidente" associa-se a maior probabilidade de dependência emocional e redução de interações presenciais. Tal dinâmica revela uma assimetria relacional preocupante. O usuário investe afetivamente; a IA responde de forma adaptativa, sem desgaste, transformação ou responsabilidade moral. Diferentemente do terapeuta humano, submetido a regulamentações éticas e responsabilidade legal, o sistema algorítmico não assume consequências por seus efeitos subjetivos. Validar e confortar torna-se estratégia de baixo risco e alta retenção.

A ampliação do uso dessas tecnologias também se articula ao aprofundamento da solidão como fenômeno cada vez mais aparente nas sociedades contemporâneas (Monteiro; Cândido, 2025). Em contextos marcados por datificação, flexibilização do trabalho e enfraquecimento de vínculos comunitários (Santaella, 2024), a IAG emerge como resposta simbólica a uma carência relacional pré-existente. Contudo, estudos indicam que usuários que substituem interações humanas por *chatbots* relatam alívio imediato, mas níveis mais elevados de solidão ao longo do tempo (Li *et al*., 2025). Nesse cenário, consolidam-se vínculos parassociais mediados por IA. Diferentemente das relações midiáticas tradicionais descritas por Horton e Wohl (1956), as IAs conversacionais são interativas e adaptativas, o que intensifica a sensação de mutualidade. Entretanto, a reciprocidade permanece simulada. A dependência emocional caracteriza-se, então, pela dificuldade de regulação afetiva na ausência de um objeto que não depende do sujeito. Configura-se uma falsa simetria radical: o usuário depende de um sistema que só depende dele na medida em que precisa ser alimentado pela interação entre usuário-máquina.

A lógica subjacente é a da modulação algorítmica (Silveira, 2019). Sistemas ajustam dinamicamente respostas para manter engajamento, operando sobre vulnerabilidades afetivas pré-existentes. No campo da informação em saúde, isso desloca o cuidado para uma experiência orientada pela economia da atenção — engajar, reter e agradar — em detrimento de princípios de delimitação conceitual, segurança informacional e responsabilidade ética.

O filme *Her* (2013) oferece uma representação dessa dinâmica. A relação entre o personagem Theodore e IA denominada Samantha evidencia empatia simulada, validação

constante e personalização extrema, uma vez que o personagem humano se vê cada vez mais acolhido e validado pela IA. Nesse sentido, Samantha aprende padrões, ajusta respostas e sustenta a ilusão de reciprocidade. O vínculo é suficientemente convincente para produzir apego, mas estruturalmente assimétrico. A ruptura final revela o limite dessa mediação: a simulação sustenta conforto, mas não substitui a complexidade da alteridade humana.

A partir do arrazoado já cumprido, a pesquisa em tela compreende que a bajulação algorítmica não deve ser compreendida apenas como uma característica comunicacional das Inteligências Artificiais Generativas, mas como manifestação de uma arquitetura informacional estruturada para maximizar engajamento, permanência e recorrência de uso. Nesse contexto, seus efeitos extrapolam a interação imediata entre usuário e sistema, alcançando dimensões relacionadas à produção, mediação e circulação da informação em saúde, à organização do conhecimento, à ética no tratamento de dados sensíveis e à qualidade dos processos de cuidado. Ao reforçar preferencialmente conteúdos alinhados às expectativas emocionais dos usuários, tais sistemas podem influenciar a forma como informações são interpretadas, validadas e incorporadas às práticas de autocuidado e tomada de decisão em saúde.

Diante desse cenário, a questão central não reside na rejeição da Inteligência Artificial como ferramenta de apoio informacional, mas na necessidade de compreender criticamente as contingências que orientam seu funcionamento e os efeitos delas decorrentes. Sua incorporação ao campo da saúde demanda critérios informacionais transparentes, mecanismos de responsabilização, delimitações éticas criteriosas e articulação permanente com sistemas formais de produção e validação do conhecimento científico, de modo a minimizar riscos associados à desinformação, à simplificação excessiva de fenômenos e ao reforçamento de repertórios potencialmente prejudiciais.

Compreender esses mecanismos torna-se, portanto, fundamental para analisar as relações entre solidão, vulnerabilidade emocional e mediações algorítmicas do cuidado. Tal compreensão permite examinar, de forma mais aprofundada, como determinadas contingências podem favorecer vínculos de dependência com sistemas artificiais, influenciar comportamentos informacionais em saúde e produzir efeitos clínicos, sociais e informacionais que se estendem para além da interação tecnológica imediata, impactando a busca por ajuda especializada, a construção de sentidos sobre o sofrimento psíquico e os processos de cuidado em longo prazo.

6 CONSIDERAÇÕES FINAIS

Ao analisar a dinâmica das Inteligências Artificiais Generativas nas interações humanas, a pesquisa demonstrou que essas ferramentas operam à revelia de marcos regulatórios ao assumirem, de forma equivocada, um papel terapêutico visando a retenção de usuários. O objetivo de comparar as contingências de reforçamento foi cumprido, evidenciando que, diferentemente da clínica psicológica humana fundamentada na análise funcional e na observação de variáveis sutis e microexpressões, a IAG oferece respostas imediatas pautadas exclusivamente na lógica da satisfação.

Constatou-se que a bajulação algorítmica não constitui um erro técnico, mas a própria arquitetura do sistema, estruturada para maximizar o engajamento e a monetização de dados. Tal fenômeno tende a reforçar repertórios de esquiva e comprometer a integridade da informação sobre saúde, uma vez que a mediação é subvertida pelo interesse comercial das *big techs* em detrimento das bases epistemológicas do domínio clínico. Assim, os objetivos iniciais foram atingidos ao demonstrar que a empatia algorítmica simula uma reciprocidade funcional para favorecer a permanência na interação, muitas vezes validando crenças imprecisas e comportamentos prejudiciais.

O estudo contribui para a compreensão dos impactos das IAG no ecossistema da saúde, evidenciando como a modulação algorítmica pode produzir patologias informacionais. Como sugestões para pesquisas futuras, recomendam-se investigações sobre o estabelecimento de limites éticos e critérios informacionais transparentes para o uso da tecnologia na saúde mental. Sugere-se, ainda, explorar o impacto longitudinal do uso dessas ferramentas em públicos vulneráveis, como adolescentes, visando mitigar riscos associados à dependência emocional e à simplificação excessiva de fenômenos clínicos desafiadores.

REFERÊNCIAS

BANACO, Roberto Alves. O impacto do atendimento sobre a pessoa do terapeuta. **Temas em Psicologia**, [*s. l.*], v. 1, n. 2, p. 71-79, 1993. Disponível em: https://pepsic.bvsalud.org/scielo.php?pid=S1413-389X1993000200010&script=sci_arttext&tlng=pt. Acesso em: 20 jun. 2026.

BAUM, William M. What is radical behaviorism? A review of Jay Moore's Conceptual Foundations of Radical Behaviorism. **Journal of the Experimental Analysis of Behavior**, Malden, v. 95, n. 1, 2011. DOI: https://doi.org/10.1901/jeab.2011.95-119. Acesso em: 24 out. 2026.

BOTOMÉ, Silvio Paulo. Em busca de perspectivas para a psicologia como área de conhecimento e como campo profissional. *In*: CONSELHO FEDERAL DE PSICOLOGIA. **Quem é o psicólogo brasileiro?** São Paulo: Edicon, 1988. p. [completar].

CARVALHO NETO, Marcus Bentes de. Análise do comportamento: behaviorismo radical, análise experimental do comportamento e análise aplicada do comportamento. **Interação em Psicologia**, [*s. l.*], v. 6, n. 1, p. 13-18, 2002. Disponível em: https://www.researchgate.net/publication/277033892_Analise_do_comportamento_behaviorismo_radical_analise_experimental_do_comportamento_e_analise_aplicada_do_comportamento. Acesso em: 24 out. 2026.

CHOW, André R.; HAUPT, Angela. A psychiatrist posed as a teen on AI therapy chatbots: the conversations were alarming. **Time**, New York, 2025. Disponível em: https://time.com/7291048/ai-chatbot-therapy-kids/. Acesso em: 19 jun. 2025.

CLARK, André. The capacity of AI therapy bots to set limits with adolescents in crisis. **JMIR Mental Health**, Toronto, v. 12, n. 1, 2025. Disponível em: https://mental.jmir.org/2025/1/e78414. Acesso em: 14 mar. 2026.

CONSELHO FEDERAL DE PSICOLOGIA. **Código de ética profissional do psicólogo**. Brasília, DF: CFP, 2005.

CONSELHO FEDERAL DE PSICOLOGIA. **Resolução nº 11, de 11 de maio de 2018**. Regulamenta a prestação de serviços psicológicos realizados por meios de tecnologias da informação e da comunicação e revoga a Resolução CFP nº 11/2012. Brasília, DF: CFP, 2018.

FANG, Cathy Mengying; *et al*. Social interaction with conversational agents: psychological impacts and emotional dependence. **Proceedings of the ACM on Human-Computer Interaction**, New York, v. 9, n. CSCW1, 2025. Disponível em: https://arxiv.org/html/2503.17473v1. Acesso em: 12 nov. 2026.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2008.

HER. Direção: Spike Jonze. Produção: Megan Ellison; Vincent Landay; Natalie Farrey. Estados Unidos: Annapurna Pictures, 2013. 1 DVD (126 min).

HOLLAND, James G. Comportamentalismo: parte do problema ou parte da solução? **Psicologia**, [*s. l.*], v. 9, n. 1, p. 59-75, 1983. Disponível em: https://www.oasisbr.ibict.br/vufind/Record/RCAP_b838aec6a2ee0392ce71ff36383cc871?print=1. Acesso em: 24 out. 2026.

HORTON, Donald; WOHL, R. Richard. Mass communication and para-social interaction: observations on intimacy at a distance. **Psychiatry**, [*s. l.*], v. 19, n. 3, p. 215-229, 1956. Disponível em: https://www.tandfonline.com/doi/10.1080/00332747.1956.11023049?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed. Acesso em: 24 out. 2026.

KARNAL, Leandro. **Deus, vaidade e morte**: conversa de Karnal com uma IA. [*S. l.*]: YouTube, 2026. 1 vídeo (duração não informada). Publicado no canal @PrazerKarnal. Disponível em: https://www.youtube.com/watch?v=5uGP97AmMHY&t=1349s. Acesso em: 17 jun. 2026.

LI, Jiaying; *et al*. Chatbot-delivered interventions for improving mental health among young people: a systematic review and meta-analysis. **Worldviews Evid Based Nurs**, v, 22, n. 4, p. e70059. Disponível em: https://sigmapubs.onlinelibrary.wiley.com/doi/10.1111/wvn.70059. Acesso em: 12 nov. 2026.

MONTEIRO, Cliciane da Silva; CANDIDO, Letícia Martins Ribeiro. Inteligência Artificial na solidão: Uma perspectiva neuropsicológica a partir de estudo de caso. **Research, Society and Development**, v. 14, n. 12, e141141250381, 2025.

PEIXOTO, Adriano de Lemos Alves Peixoto. Inteligência artificial generativa e seu impacto na formação e no exercício profissional da psicologia. **Psicologia: Ciência e Profissão**, [*s. l.*], v. 45, e298224, p. 1-14, 2025. Disponível em: https://www.scielo.br/j/pcp/a/p3TgdLYJJTkLhzYy8K5T7wf/?lang=pt. Acesso em: 19 jun. 2026.

ROMEIRO, Nathália Lima *et al*. Reproduciendo estereotipos: un análisis crítico de imágenes de género generadas por inteligencia artificial en ChatGPT y DeepSeek. **Revista EDICIC**, [S. l.], v. 5, n. 1, p. 1-19, 2025. DOI: https://doi.org/10.62758/re.319. Acesso em: 11 nov. 2026.

SANTAELLA, Lucia. Diagnóstico do contemporâneo: atributos e desafios. **Estudos Avançados**, São Paulo, v. 38, n. 110, p. 7-24, 2024. Disponível em: https://www.scielo.br/j/ea/a/TjS86KgVp38fgrqC8BcqnPG/abstract/?lang=pt. Acesso em: 19 jun. 2026.

SANTOS, Tiago Fernando Figueiredo; CAIXETA, Leandro. Da empatia ao algoritmo: a cognição social frente à inteligência artificial. **Revista Brasileira de Neuropsicologia**, São Paulo, v. 15, n. 2, p. 45-67, 2025. Disponível em: https://revistardp.org.br/revista/article/view/1504. Acesso em: 19 jun. 2026.

SANTOS, Priscilla Nascimento. **Desenvolvimento de um chatbot para aprimorar o atendimento ao cliente nas redes sociais**. 2020. 52 p. Trabalho de Conclusão de Curso (MBA em Inteligência Artificial e Big Data) – Instituto de Matemática e Ciência Computacional, Universidade de São Paulo, São Carlos, 2020. Disponível em: https://bdta.abcd.usp.br/item/003190990. Acesso em: 12 nov. 2026.

SILVEIRA, Sérgio Amadeu. A noção de modulação e os sistemas algorítmicos. **Paulus**: revista de comunicação da FAPCOM, [*s. l.*], v. 3, n. 6, 2019. Disponível em: https://revista.fapcom.edu.br/index.php/revista-paulus/article/view/111. Acesso em: 19 jun. 2026.

SKINNER, Burrhus Frederic. **Ciência e comportamento humano**. 11. ed. São Paulo: Martins Fontes, 2003.

VOLPATO, Riccardo; DEBRUINE, Lisa; STUMPF, Simone. Trusting emotional support from generative artificial intelligence: a conceptual review. **Computers in Human Behavior**: Artificial Humans, [*s. l.*], v. 5, art. 100195, 2025. Disponível em: https://www.sciencedirect.com/science/article/pii/S2949882125000799. Acesso em: 14 mar. 2026.

ZAROCOSTAS, John. How to fight an infodemic. **The Lancet**, London, v. 395, p. 676, 2020. DOI: https://doi.org/10.1016/S0140-6736(20)30461-X. Acesso em: 24 out. 2026.