

Mente, máquina e sentido: um ensaio sobre a comunicação no mundo artificial¹

Vander Muniz² – IDP

Ébida Santos³ – IDP

Resumo

Este artigo observa a Comunicação Humano-Máquina (CHM) no contexto da Inteligência Artificial (IA), explorando teoricamente as implicações para a cognição, construção de sentido e a interação social. Com base na Teoria da Mente (ToM) e na cibernética, argumenta-se que os sistemas de IA, ao deixarem de ser agentes passivos, passam a desempenhar um papel ativo na modelagem dos processos de comunicação e cognitivos. A reflexão estrutura-se em quatro eixos centrais: a atribuição de intencionalidade a agentes artificiais e a forma como a ToM influencia essa percepção; o papel da cibernética no ajuste adaptativo da comunicação, a partir de mecanismos de feedback; os impactos da IA na construção da interação social e na percepção da ação dos agentes computacionais; e as tensões éticas, políticas e existenciais, próprias do tempo presente, decorrentes da crescente integração da IA na comunicação humana.

Palavras-chave

Comunicação Humano-Máquina; Teoria da Mente; Cibernética; Inteligência Artificial

Introdução

O crescente desenvolvimento dos sistemas de inteligência artificial (IA) e sua integração à comunicação humana tem feito emergir desafios aos modelos tradicionais de interação e cognição. A Comunicação Humano-Máquina (CHM), impulsionada por modelos de Processamento de Linguagem Natural, redes neurais artificiais profundas e aprendizado por reforço, não está restrita a ser um processo instrumental apenas. Hoje, agentes artificiais alteram, ativamente, a percepção e a construção de sentido enquanto interagem com indivíduos.

Este ensaio instiga a ideia de que a CHM, ao se conectar à Teoria da Mente (ToM) e a cibernética, altera as dinâmicas dos processos de comunicação, tirando os sistemas de IA da posição de agentes passivos e colocando-os como participantes na construção de significado e na adaptação cognitiva dos indivíduos. Porém, precisamos observar que esse fenômeno ocorre em um contexto paradoxal do tempo presente, onde a velocidade do

¹ Trabalho apresentado no GP Teorias da Comunicação, do 25º Encontro dos Grupos de Pesquisas em Comunicação, evento componente do 48º Congresso Brasileiro de Ciências da Comunicação.

² Mestrando do Programa de Pós-Graduação em Comunicação do IDP, e-mail: vanderem@uol.com.br

³ Pós-doutoranda pelo INCT-DSI/UFF, Doutora em Comunicação pela UnB e Docente Orientadora do Programa de Pós-Graduação em Comunicação do IDP, e-mail: ebida.santos@idp.edu.br

progresso tecnológico e a sensação de hiperconexão convivem com incertezas, demandas por reflexão e desafios sociais (Bauman, 2001; Han, 2015). Para explorar essa hipótese, observamos sua construção em quatro seções:

1. A influência da ToM na atribuição de intencionalidade a agentes artificiais;
2. O papel da cibernética no *feedback* adaptativo e na personalização da comunicação;
3. Os impactos dos sistemas inteligentes na construção da interação social e na percepção de ação;
4. O paradoxo do tempo presente e as reflexões sobre humanidade e inteligência artificial, considerando tensões éticas, políticas e existenciais inerentes a essa realidade.

A partir desses pontos centrais, discutimos como a comunicação mediada por sistemas de IA reflete processos mentais humanos e os transforma, conformando padrões de interação, percepção de ação e construção de significado. Essa transformação que influencia a forma como os indivíduos compreendem e processam informações, impactando diretamente a cognição, o discurso no meio social e a própria condição humana.

Teoria da mente e comunicação com sistemas computacionais

A Teoria da Mente (ToM) refere-se à capacidade de inferir estados mentais em outros agentes e antecipar suas intenções (Premack; Woodruff, 1978). Essa característica, essencial para a interação humana, é ativada mesmo em contextos de comunicação com sistemas de IA levando indivíduos a projetarem intencionalidade e emoção em agentes computacionais (Banks, 2020).

Pesquisas indicam que a atribuição de características humanas a sistemas de IAs ocorre com base em padrões previsíveis de resposta (Shank; DeSanti, 2018). A implementação de modelos de redes neurais artificiais profundas amplificou a capacidade destes sistemas de interpretar a linguagem natural e responder de forma contextualizada, aumentando a ilusão de ação e intenção. Tais modelos são exemplificados por soluções de IA Generativa como o ChatGPT da OpenAI (Vaswani et al., 2017).

Porém, há uma distinção entre intencionalidade humana e artificial, constantemente deixada à margem. Agentes artificiais não possuem estados mentais; eles simulam coerência e personalização a partir de enormes volumes de dados pré-processados (Harré; El-Tarifi, 2023). Isso levanta questões epistemológicas e éticas: até que ponto a percepção de ação influencia a confiabilidade e a credibilidade dessas interações? E quais os impactos dessa atribuição de intencionalidade para a tomada de decisões em ambientes mediados por sistemas de IA?

Percepção versus realidade

Aqui, usamos o termo “percepção de ação” para descrever o fenômeno onde indivíduos interpretam comportamentos de sistemas de IA, como respostas personalizadas, recomendações ou tomadas de decisão, como sendo manifestações intencionais de um agente consciente. Essa percepção emerge da simulação de coerência, responsividade e adaptação, frequentemente associadas a interações humanas dotadas de sentido e propósito. Mesmo que esses sistemas não possuam estados mentais, o modo como operam desperta no indivíduo a impressão de que algo ou alguém está “agindo” (Legaspi et al., 2024).

Ao interpretar a “intencionalidade simulada” de um agente artificial com a intencionalidade efetiva, é criado o risco de superestimar a capacidade de compreensão dos sistemas de IA. Mesmo que os indivíduos, frequentemente, percebam nas respostas traços “humanos” como empatia ou afetividade (Banks, 2020), na construção dessas respostas, o que existe é a aplicação de algoritmos probabilísticos que consomem dados massivamente (Vaswani et al., 2017). Identificar a diferença entre percepção e realidade é necessário para evitar uma transferência indevida de confiança a sistemas de IA que, mesmo avançados, não possuem consciência ou emoções (Jeung; Huang, 2023).

Cibernética e *feedback* adaptativo na personalização da comunicação

A cibernética, proposta por Wiener (2019), introduziu o conceito de *feedback* como mecanismo necessário para a adaptação de sistemas, humanos ou não-humanos. Na abordagem da CHM, esse princípio descreve como os sistemas de IA ajustam suas respostas com base em interações passadas, tornando-se mais personalizadas e eficientes (Guzman; Lewis, 2020).

Uma das principais técnicas usadas em sistemas de IA, o aprendizado por reforço, permite que agentes artificiais modifiquem seu comportamento para maximizar a eficiência comunicativa (Noda et al., 2014). Esse tipo de adaptação gera um ciclo de reequilíbrio de respostas contínuo, onde os sistemas de IA adaptam-se adequando a linguagem, o tom e os padrões de cada indivíduo, potencializando a aceitação e a familiaridade (Kaplan et al., 2023). É desta forma que a CHM expande a probabilidade de manipulação dos processos de comunicação, principalmente quando os indivíduos não reconhecem a estrutura algorítmica que constrói essas adaptações.

Sistemas inteligentes e a construção da interação

O efeito dessas mudanças dos processos de comunicação, que introduz os sistemas de IA na dinâmica discursiva, vai além de como as mensagens são estruturadas e entram em como elas reconfiguram as expectativas dos usuários sobre as interações humanas físicas e digitais (Westerman et al., 2020).

Pesquisas demonstram que a previsibilidade e a eficiência dos sistemas de IA podem desencadear fenômenos de adaptação cognitiva onde os indivíduos internalizam padrões interativos, modelados pelos agentes artificiais, como sendo humanos (Kaplan et al., 2023). Isso altera a relação entre linguagem, intenção e compreensão, uma vez que os sistemas de IA ganham esse espaço ativo na formação de sentido em contextos sociais.

Esse processo leva à consolidação conceitual da percepção de ação. Uma tendência de os indivíduos atribuírem agência e intencionalidade às respostas de sistemas de IA, como se estes fossem oriundos de escolhas conscientes. Essa percepção reconstrói os conceitos tradicionais da comunicação. Isso ocorre quando se transfere para máquinas atributos evolutivamente vinculados à interação humana, como empatia, julgamento e intenção. O resultado é uma mudança nas normas cognitivas e sociais. Passa-se a esperar coerência, responsividade e até moralidade de sistemas que operam com base em regras probabilísticas e aprendizado estatístico. Ao naturalizar essa presença artificial como interlocutor legítimo, os limites entre humano e não humano se tornam mais difusos, e as fronteiras da linguagem e da ação, mais ambíguas (Legaspi et al., 2024).

Humanidade e IA

Os dias presentes, que abrigam a rápida evolução da CHM, são marcados por uma combinação de hiperconectividade, aceleração do tempo social e busca por significados profundos (Bauman, 2001; Han, 2015). No mesmo tempo que sistemas de IA entregam facilidade na troca de informações e otimizam processos, eles podem, também, gerar ansiedade, fragilidade nas relações e a sensação de que a sociedade se reconfigura em um ritmo superior do que a capacidade humana tem de assimilá-la (Han, 2015). Essa percepção aumenta quando observamos que os sistemas de IA, na simulação de intencionalidade, moldam a forma como os indivíduos se relacionam com a informação e com o próprio sentido do “ser humano” (Banks, 2020; Harré; El-Tarifi, 2023).

A velocidade com que os sistemas de IA processam dados e fornecem respostas reforça a expectativa de instantaneidade. No nosso contexto, a “pressa” digital pode restringir o espaço-tempo para a reflexão crítica e aprofundada, uma vez que a interação com agentes artificiais se baseia em respostas imediatas e eficientes, onde uma compreensão mais ampla sobre como e por que determinados conteúdos são gerados não tem validade (Jeung; Huang, 2023; Vaswani et al., 2017).

A intencionalidade expressa na velocidade de resposta, que pressupõe uma ação benéfica, contrasta com a dependência, progressiva, que é desenvolvida em relação aos sistemas de IA (Bauman, 2001). Mesmo que o aprendizado por reforço permita o aperfeiçoamento contínuo dos agentes artificiais, esse aprimoramento pode restringir a capacidade de autonomia dos indivíduos, que confiará nas recomendações e respostas produzidas sem questionar a lógica subjacente (Guzman; Lewis, 2020).

Aqui emerge um aparente paradoxo. Quanto mais conectadas os indivíduos estão por meio de sistemas de IA, maior pode ser a sensação de distanciamento humano. Essa possibilidade surge posto que as interações mediadas por sistemas de IA, que simulam empatia e afetividade (Banks, 2020), não compartilham experiências afetivas ou contextuais concretas (Harré; El-Tarifi, 2023). Em ritmo a isto, a hiperconexão digital pode aprofundar a solidão e a superficialidade dos vínculos, em lugar de efetivamente aproximar as pessoas (Han, 2015).

Se, por um lado, os sistemas de IA oferecem possibilidades de otimizar processos e solucionar problemas complexos, por outro, a velocidade de seu desenvolvimento exige um reajuste de nossa compreensão dos processos de comunicação, responsabilidade social e do significado de “ser humano”.

Quando observamos isto à luz da ToM, da cibernética e da hiperconectividade, vemos que o paradoxo do tempo presente não é uma contradição entre progresso e inércia, mas sim o modo como a aceleração tecnológica repercute nas formas de cognição, percepção de ação e construção de significado da própria sociedade. Reconhecer os potenciais e os riscos advindos dessa dinâmica é fundamental para preservar a autonomia discursiva, a diversidade de perspectivas e a profundidade das relações humanas.

Essas observações gerais, demonstram a necessidade de uma abordagem empírica que desenvolva estudos multi, inter e pluridisciplinares. Estes estudos podem mostrar se e como a autonomia discursiva dos indivíduos se modifica diante de sistemas responsivos e adaptativos. Westerman et al. (2020) destacam que a forma pela qual as pessoas interpretam as respostas de sistemas de IA depende do contexto de uso e do grau de familiaridade com a tecnologia, isso reforça a necessidade dos estudos em diferentes coortes.

Considerações finais

As reflexões apresentadas ao longo deste ensaio indicam que a CHM, na articulação dos princípios da ToM e da cibernética, ultrapassa a mediação instrumental e assume um papel ativo na construção do significado e na transformação cognitiva dos indivíduos. Se, ao usarmos um prisma, pelo ponto de entrada da luz temos a atribuição de intencionalidade e o *feedback* adaptativo impulsionando a eficácia dos processos de comunicação, na saída, a refração desmembra em questionamentos éticos, políticos e epistemológicos que demandam um olhar científico atento e capaz de acompanhar, teórica e temporalmente, essas transformações.

O uso intenso de sistemas de IA pode reforçar uma cultura de instantaneidade e uma sobrecarga informacional que acentua o risco de manipulação discursiva e de subestimação das limitações tecnológicas. É neste sentido que mecanismos capazes de administrar essas tensões se fazem necessários. Entre eles, políticas de transparência e

regulação tornam-se mandatórias, assim como debates públicos sobre governança algorítmica e *accountability*.

No vasto e interdisciplinar campo da comunicação, futuras investigações empíricas são essenciais para trazer luz sobre como, em diferentes contextos culturais e socioeconômicos, os sistemas de IA moldam padrões de interação e a percepção de ação. Concomitantemente, é urgente fomentar discussões interdisciplinares sobre ética, sentido e humanidade na era digital. É necessário reconhecer que a comunicação mediada por sistemas de IA afetam nossa cognição e discurso e redefinem dimensões fundamentais da condição humana.

Em última instância, a forma como integramos os sistemas de IA em nossas vidas será determinante para suavizar ou amplificar as contradições do tempo presente.

Referências

- BANKS, Jaime. Theory of Mind in Social Robots: Replication of Five Established Human Tests. **International Journal of Social Robotics**, v. 12, n. 2, p. 403–414, 12 maio 2020.
- BAUMAN, Zygmunt. **Modernidade Líquida**. 1. ed. [S.l.]: Zahar, 2001.
- GUZMAN, Andrea L.; LEWIS, Seth C. Artificial intelligence and communication: A Human–Machine Communication research agenda. **New Media & Society**, v. 22, n. 1, p. 70–86, 4 jan. 2020.
- HAN, Byung-Chul. **Sociedade do Cansaço**. 1. ed. [S.l.]: Vozes, 2015.
- HARRÉ, Michael S.; EL-TARIFI, Husam. Testing Game Theory of Mind Models for Artificial Intelligence. **Games**, v. 15, n. 1, p. 1, 28 dez. 2023.
- JEUNG, Jun Li; HUANG, Janet Yi-Ching. Correct Me If I Am Wrong: Exploring How AI Outputs Affect User Perception and Trust. *In*: New York, NY, USA: ACM, 14 out. 2023.
- KAPLAN, Alexandra D. *et al.* Trust in Artificial Intelligence: Meta-Analytic Findings. **Human Factors: The Journal of the Human Factors and Ergonomics Society**, v. 65, n. 2, p. 337–359, 28 mar. 2023.
- LEGASPI, Roberto *et al.* The sense of agency in human–AI interactions. **Knowledge-Based Systems**, v. 286, p. 111298, fev. 2024.

NODA, Kuniaki *et al.* Multimodal integration learning of robot behavior using deep neural networks. **Robotics and Autonomous Systems**, v. 62, n. 6, p. 721–736, jun. 2014.

PREMACK, David; WOODRUFF, Guy. Does the chimpanzee have a theory of mind? **Behavioral and Brain Sciences**, v. 1, n. 4, p. 515–526, 4 dez. 1978.

SHANK, Daniel B.; DESANTI, Alyssa. Attributions of morality and mind to artificial intelligence after real-world moral violations. **Computers in Human Behavior**, v. 86, p. 401–411, set. 2018.

VASWANI, Ashish *et al.* Attention Is All You Need. **Advances in Neural Information Processing Systems**, v. 30, 12 jun. 2017.

WESTERMAN, David *et al.* I-It, I-Thou, I-Robot: The Perceived Humanness of AI in Human-Machine Communication. **Communication Studies**, v. 71, n. 3, p. 393–408, 26 maio 2020.

WIENER, Norbert. **Cybernetics or Control and Communication in the Animal and the Machine**. [S.l.]: The MIT Press, 2019.