Agentes Conversacionais Inteligentes para a Inclusão Digital

Vasco Furtado^{1,2}, Elizabeth S. Furtado^{1,3}

¹Universidade de Fortaleza, UNIFOR
Av. Washington Soares, 1321, Fortaleza, CE, Brasil

²ETICE - Empresa de Tecnologia da Informação do Ceará
Av. Pontes Vieira 220, Fortaleza, CE, Brasil

³Universidade Estadual do Ceará - UECE
Campus do Itaperi, Fortaleza, CE, Brasil

{vasco, elizabet}@unifor.br

Abstract. *In Brazil, where socioeconomic inequality and low literacy rates hinder access to technology, the emergence of intelligent conversational agents based on natural language — especially those powered by Large Language Models (LLMs) — holds the potential to include historically marginalized populations in the digital ecosystem. However, for this scenario to unfold responsibly, it is essential to recognize that most widely used LLMs have been trained predominantly on foreign data, detached from Brazil’s linguistic and cultural realities. Beyond linguistic barriers, there are significant ethical and social risks associated with the use of AI agents in direct interaction with people. The tendency toward anthropomorphism may lead vulnerable users to develop emotional bonds or excessive trust in intelligent agents, thereby increasing the risk of manipulation, misinformation, or malicious use. In this context, one of the greatest challenges for Brazilian computing in the coming years will be to design and implement technologies for human–conversational agent interaction that foster equitable access, taking into account the country’s social, linguistic, and cultural particularities. This challenge encompasses efforts ranging from the development and/or fine-tuning of language models adapted to Brazilian contexts, to the evaluation of ethical implications and risks associated with the widespread adoption of intelligent conversational agents, and the formulation of public policies and inclusive strategies that leverage these technologies as tools for digital citizenship.*

Resumo. *No Brasil, onde a desigualdade socioeconômica e os baixos índices de letramento comprometem o acesso à tecnologia, a emergência de agentes conversacionais inteligentes baseados em linguagem natural – especialmente os impulsionados por Modelos Amplos de Linguagem (LLMs) – pode viabilizar a entrada de populações historicamente marginalizadas no ecossistema digital. Contudo, para que esse cenário se concretize de forma responsável, é essencial reconhecer que os LLMs amplamente utilizados hoje foram treinados majoritariamente com dados estrangeiros, distantes da realidade linguística e cultural brasileira. Além das barreiras linguísticas, há riscos éticos e sociais substanciais associados ao uso de agentes de IA em interação direta com pessoas. A tendência ao antropomorfismo pode levar usuários vulneráveis a desenvolver vínculos emocionais ou confiança excessiva nos agentes inteligentes,*

agravando riscos de manipulação, desinformação ou uso malicioso. Nesse contexto, o grande desafio da computação brasileira nos próximos anos será projetar e implementar tecnologias para interação entre agentes inteligentes conversacionais e humanos que promovam acesso com equidade, considerando as singularidades sociais, linguísticas e culturais do país. Esse desafio incorpora ações que vão desde o desenvolvimento e/ou refinamento de modelos de linguagem ajustados à realidade brasileira, avaliação dos impactos éticos e os riscos da adoção massiva de agentes conversacionais inteligentes e formulação de políticas públicas e estratégias inclusivas que utilizem essas tecnologias como instrumento de cidadania digital.

1. Introdução

A interação entre humanos e sistemas de Inteligência Artificial (IA) tornou-se um eixo central para os avanços tecnológicos e sociais contemporâneos, especialmente em contextos marcados por desigualdades estruturais, como o brasileiro. Em um país onde disparidades socioeconômicas e baixos níveis de letramento ainda comprometem o acesso equitativo ao mundo digital, as interfaces baseadas em linguagem natural — viabilizadas pelos Modelos Amplos de Linguagem (LLMs, do inglês *Large Language Models*) — emergem como promissoras ferramentas de inclusão. Tais modelos permitem o desenvolvimento de agentes conversacionais inteligentes capazes de mediar o acesso à informação de maneira mais acessível e fluida, ampliando o alcance de serviços e conhecimentos a grupos historicamente marginalizados.

Embora essas características dos agentes conversacionais representem um cenário promissor, elas também trazem complexidades que motivam o desafio central deste artigo: como tornar assistentes conversacionais inteligentes vetores importantes para a inclusão digital. Isso exige uma perspectiva multidisciplinar que vá além da computação (especialmente IA e IHC), envolvendo saberes da filosofia, sociologia, comunicação, direito, psicologia e outras áreas do conhecimento. É por meio dessa articulação que será possível compreender e antecipar os impactos sociais, cognitivos e culturais desses agentes na vida das pessoas e propor soluções que tornem a tecnologia inclusiva.

Para isto, propomos uma abordagem que se inspira e complementa alguns desafios propostos na área de IHC [Pereira et al. 2024]. Enfatizamos que, para que essas tecnologias sejam de fato inclusivas, é fundamental que estejam ancoradas na realidade brasileira. Isso inclui treinar os agentes com dados representativos dos contextos cultural e educacional do país. Estudos demonstram que o vocabulário e as estruturas frasais utilizadas por pessoas em situação de analfabetismo funcional ou baixa escolaridade divergem substancialmente da norma culta empregada por modelos linguísticos convencionais [Tarallo 1985, Monte 2019]. Tais diferenças, se ignoradas, podem comprometer a eficácia das interações [da Silva et al. 2022]. Portanto, os riscos e oportunidades trazidos pela adoção de sistemas de IA devem ser analisados sob uma ótica crítica e situada, considerando as especificidades sociais e culturais da sociedade brasileira.

2. Justificativa

A presente proposta justifica-se pela urgência crescente de tornar os serviços digitais mais acessíveis, inclusivos e socialmente adequados ao contexto brasileiro. Em um país

onde a exclusão digital ainda atinge milhões de pessoas, a adoção de soluções baseadas em agentes conversacionais inteligentes surge como uma estratégia promissora para ampliar o acesso à informação e aos serviços públicos e privados. Interações em linguagem natural — tanto escritas quanto faladas — podem representar uma via crucial de aproximação para indivíduos com baixa escolaridade, permitindo que esses usuários acessem conteúdos e recursos que, de outro modo, permaneceriam inacessíveis.

Contudo, a efetividade dessas interações é condicionada por variáveis demográficas, regionais e educacionais, que impactam a forma como os usuários se comunicam com os sistemas. No caso da língua portuguesa falada no Brasil, estudos demonstram que o desempenho dos assistentes conversacionais por voz pode variar conforme os padrões regionais e os níveis de escolaridade dos interlocutores [Lima et al. 2019, da Silva et al. 2022]. Essas variações afetam aspectos acústicos, cognitivos e estruturais da linguagem, criando obstáculos à compreensão e à personalização da interação.

No campo da Interação Humano-Computador (IHC), já se discute amplamente aspectos éticos, de privacidade e de apoio à diversidade no desenvolvimento de sistemas interativos. A iniciativa de Pereira et al. (2024) propõe seis desafios para a área de IHC no período de 2025 a 2035, incluindo o desafio específico das 'Implicações da IA', tratado como uma questão à parte. No entanto, ao abordar agentes conversacionais inteligentes, torna-se necessário adotar uma perspectiva mais ampla. Isso implica incorporar e expandir os seis desafios propostos, de modo que orientem tanto o desenvolvimento, através de técnicas de treinamento e engenharia de *prompt*, quanto a análise crítica desses agentes. Pois, esses agentes não apenas simulam comportamentos humanos com crescente naturalidade, mas também atuam com relativa autonomia, assumindo decisões, emitindo juízos e, por vezes, extrapolando as intenções originalmente projetadas por seus desenvolvedores.

Essas dificuldades são amplificadas quando se considera que os principais modelos de linguagem disponíveis no mercado são treinados com corpora estrangeiros, que não refletem com precisão as diversidades linguística e cultural brasileiras. Isso tem levado à reprodução de vieses e limitações no desempenho dos agentes, como apontado em estudos recentes [Silva et al. 2024]. Como resultado, as interações podem se tornar frustrantes ou ineficazes — e, em casos mais graves, até manipuladoras, comprometendo a segurança informacional e a autonomia dos usuários.

Além disso, é preciso considerar o elevado grau de receptividade do brasileiro a novas tecnologias, o que, embora positivo sob o ponto de vista da adoção, pode facilitar fenômenos de antropomorfismo e confiança excessiva em sistemas automatizados. Interfaces que simulam características humanas, como nos assistentes corporificados (ECAs) [Yasavur and Rishe 2014], ou agentes projetados para atuar como companheiros pessoais [Oliveira et al. 2025], tendem a gerar vínculos emocionais que podem agravar a vulnerabilidade cognitiva de certos grupos populacionais. Esses riscos levantam questões cruciais sobre a segurança, a privacidade, a credibilidade e a ética na utilização de tecnologias interativas [Slattery et al. 2024, Shneiderman 2020].

Diante desse cenário, é imprescindível que a comunidade científica nacional se debruce sobre os temas desse desafio, desenvolvendo estratégias e soluções para agentes conversacionais inteligentes que dialoguem com a realidade brasileira de forma ética,

crítica e socialmente comprometida.

3. Objetivos

O objetivo desta proposta é identificar e analisar propostas relacionadas ao desafio do desenvolvimento de agentes conversacionais inteligentes no contexto brasileiro que promovam a inclusão digital, a fim de planejar como mitigar os riscos que lhes são associados e alavancar a introdução da tecnologia. As propostas se concentram em:

1. Desenvolvimento de modelos de linguagem e técnicas de engenharia de *prompt* ajustados ao Brasil: Criar e treinar modelos que considerem as particularidades linguísticas e culturais do Brasil, especialmente focados em populações marginalizadas;

2. Desenvolvimento de abordagens multidisciplinares; Elaborar estratégias e diretrizes, com contribuições de áreas multidisciplinares, para que os *stakeholders* mais relevantes possam aplicar a fim de compreenderem e anteciparem os impactos e perigos desses agentes na vida das pessoas;

3. Políticas e Ações de Inclusão Digital: Explorar o potencial de agentes conversacionais inteligentes como ferramenta para reduzir a exclusão digital no Brasil, promovendo a inclusão de pessoas com pouca ou nenhuma familiaridade com o mundo digital; e

4. Avaliação Ética da Interação com agentes conversacionais inteligentes: Analisar os impactos éticos do uso de agentes conversacionais inteligentes, identificando potenciais perigos, como manipulação e dependência emocional, além de riscos de segurança relacionados ao uso malicioso dessas tecnologias.

4. Descrição do Desafio

A interação entre humanos e agentes conversacionais inteligentes representa, atualmente, um importante desafio da computação no contexto brasileiro. Tal complexidade decorre, em grande medida, da necessidade de adaptar essas tecnologias às especificidades linguísticas, culturais e educacionais da população. O desafio adquire contornos ainda mais críticos quando se considera que muitas dessas interações acontecerão com pessoas em situação de baixa escolaridade, inseridas em um cenário de profunda diversidade sociocultural.

Sem diretrizes claras que orientem o desenvolvimento e a implementação desses sistemas, corre-se o risco de desperdiçar uma oportunidade estratégica de promoção da inclusão digital. Em vez de reduzir desigualdades, a tecnologia mal calibrada pode perpetuar ou até mesmo ampliar barreiras existentes, tornando a experiência digital excludente para os públicos que mais necessitam de apoio.

Outro aspecto preocupante diz respeito à confiança desmedida que usuários com baixa cognição podem depositar em agentes conversacionais. Atribuir a esses sistemas um grau elevado de autoridade ou humanidade — especialmente em contextos marcados por baixa literacia digital — pode gerar efeitos colaterais graves, como dependência emocional, manipulação involuntária e vulnerabilidade a fraudes. Esses riscos são intensificados pela já precária infraestrutura de segurança digital no país, agravando o potencial de danos individuais e coletivos.

Diante desse cenário, torna-se imprescindível compreender em profundidade as múltiplas dimensões envolvidas na interação das pessoas com esses agentes. Essa compreensão deve fundamentar o desenvolvimento de agentes conversacionais mais seguros, éticos [Gero et al. 2025] e inclusivos, apoiando a criação de normas, protocolos e diretrizes que orientem seu uso responsável [Neves Camêlo and Alves 2023]. Esses instrumentos devem ser passíveis de incorporação tanto em sistemas de governança institucional — como políticas públicas e regulamentações — quanto em práticas de engenharia de *prompt* e de governança distribuída, lideradas pela sociedade civil e comunidades usuárias da tecnologia.

5. Metodologia

- **Coleta de Dados:** Obter dados sobre interações de usuários brasileiros com modelos de linguagem amplos, identificando falhas na compreensão linguística e nas respostas fornecidas. Isso incluirá um enfoque em populações com baixo nível educacional;
- **Interdisciplinaridade:** Promover uma forte interação entre as áreas de IA e IHC bem como entre essas áreas e outros ramos do saber como comunicação, filosofia, linguística, psicologia e sociologia.
- **Desenvolvimento de Modelos:** Criar versões personalizadas de modelos de linguagem, com ajustes baseados nos dados coletados, visando melhorar a compreensão e a utilidade das respostas para o público-alvo brasileiro;
- **Avaliação de Riscos:** Implementar avaliações contínuas dos riscos associados à interação humano-IA, como manipulação emocional e insegurança digital. Essas avaliações devem ser feitas em colaboração com especialistas em ética e segurança da informação;
- **Políticas de Mitigação:** Desenvolver diretrizes para o desenvolvimento e uso de agentes conversacionais inteligentes no Brasil, com foco em minimizar os riscos e maximizar os benefícios da inclusão digital;
- **Desenvolvimento de serviços digitais:** Atender a grandes parcelas da população excluídas digitalmente usando agentes inteligentes apropriados; e
- **Educação:** integrar teoria e prática na educação, promovendo uma reflexão contínua sobre sociedade e regulamentações, princípios éticos, processos técnicos e contemplando-os na engenharia de *prompt* para a interação do usuário com agentes inteligentes conversacionais [Porto and Furtado 2024] e [Silva et al. 2024].

6. Potenciais Métricas de Avaliação

Para avaliar o progresso das soluções propostas, as seguintes métricas serão consideradas:

- **Precisão Cultural-Linguística:** Percentual de interações bem-sucedidas entre usuários brasileiros e agentes conversacionais inteligentes, com base em parâmetros específicos de linguagem e contexto cultural;
- **Nível de Satisfação do Usuário:** Avaliação da experiência do usuário com agentes conversacionais inteligentes, medida através de pesquisas de satisfação e feedback direto;
- **Redução na Exclusão Digital:** Percentual de novos usuários digitais incluídos com sucesso por meio de interações com agentes conversacionais inteligentes, especialmente em regiões com baixos índices de acesso à tecnologia;

- Risco de Manipulação: Incidências de manipulação ou mal-uso da IA, identificadas em testes controlados e monitoramento contínuo; e
- Conformidade Ética: Grau de aderência às diretrizes éticas desenvolvidas, para apoiar o letramento dos dados e a privacidade acessível, com auditorias regulares para garantir conformidade.

7. Conclusão

A interação entre humanos e agentes conversacionais inteligentes é uma fronteira relevante e desafiadora para o futuro da computação no Brasil. Trata-se de uma área que, ao mesmo tempo em que oferece um potencial transformador para ampliar o acesso à informação, facilitar o uso de serviços públicos e privados, e promover a inclusão digital de milhões de brasileiros, também traz consigo riscos substanciais que não podem ser ignorados.

Tais riscos — que vão desde a exclusão involuntária de populações com baixa escolaridade até a indução à dependência emocional e à manipulação — exigem uma abordagem crítica e comprometida com os princípios da equidade, da transparência e da responsabilidade social. No contexto brasileiro, onde coexistem intensa desigualdade social, rica diversidade cultural e alta adesão às tecnologias digitais, o desafio de desenvolver sistemas que respeitem as singularidades da linguagem, da cognição e das formas de vida locais é particularmente urgente.

Acreditamos que, por meio de uma abordagem ética, interdisciplinar e culturalmente situada, é possível conceber e implementar agentes conversacionais inteligentes que não apenas ampliem a participação digital, mas que o façam de maneira segura, inclusiva e respeitosa às condições reais de seus usuários. Isso implica o desenvolvimento de modelos de linguagem treinados com dados representativos do Brasil, o estabelecimento de protocolos para mitigação de riscos, e a construção de políticas públicas orientadas à promoção de uma cidadania digital plena.

Nesse esforço, destaca-se ainda a importância de desenvolver técnicas de engenharia de *prompt* que estejam em conformidade com as diretrizes estabelecidas para a construção de interações responsáveis. Essas técnicas devem considerar as variações linguísticas e culturais dos usuários, bem como os objetivos éticos e funcionais dos agentes, garantindo que os modelos operem de forma transparente, previsível e alinhada às necessidades e limitações da população atendida.

Enfrentar esse desafio com seriedade e visão estratégica permitirá ao Brasil não apenas avançar internamente em termos de justiça tecnológica e inclusão social, mas também posicionar-se como referência internacional na formulação de soluções de inteligência artificial que aliem inovação técnica a compromisso ético.

References

- da Silva, T. H. O., Furtado, V., Furtado, E., Mendes, M., Almeida, V., and and, L. S. (2022). How do illiterate people interact with an intelligent voice assistant? *International Journal of Human-Computer Interaction*, 40(3):584–602.
- Gero, K. I., Desai, M., Schnitzler, C., Eom, N., Cushman, J., and Glassman, E. L. (2025). Creative writers’ attitudes on writing as training data for large language models. CHI ’25, New York, NY, USA. Association for Computing Machinery.

- Lima, L., Furtado, V., Furtado, E., and Almeida, V. (2019). Empirical analysis of bias in voice-based personal assistants. In *Companion Proceedings of The 2019 World Wide Web Conference*, pages 533–538.
- Monte, A. (2019). A influência da escolaridade e do sexo/gênero no uso variável da concordância verbal de terceira pessoa do plural. *Revista Diálogos*, 7:89–104.
- Neves Camêlo, M. and Alves, C. (2023). G-priv: A guide to support lgpd compliant specification of privacy requirements. *iSys - Brazilian Journal of Information Systems*, 16:2:1 – 2.
- Oliveira, J., Silva, T., Oliveira, R., and Furtado, E. (2025). Recommendations of embodied conversational agents to healthcare applications.
- Pereira, R., Darin, T., and Silveira, M. S. (2024). Grandihc-br: Grand research challenges in human-computer interaction in brazil for 2025-2035. In *Proceedings of the XXIII Brazilian Symposium on Human Factors In Computing Systems*.
- Porto, A. V. F. and Furtado, M. E. S. (2024). Framework to specify dialogues for natural interaction with conversational assistants applied in prompt engineering. In Arai, K., editor, *Intelligent Systems and Applications*, pages 231–253, Cham. Springer Nature Switzerland.
- Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered ai systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 10(4):1–31.
- Silva, V., Furtado, E., Oliveira, J., and Furtado, V. (2024). Engenharia de prompts em assistentes conversacionais para promoção de autocuidado baseados em modelos amplos de linguagem. In *Simpósio Brasileiro de Computação Aplicada à Saúde*, Goiânia.
- Slattery, P., S. A. K., Grundy, Emily A. C.; Graham, J., Noetel, M.; Uuk, R. D. J. P. S. C. S., and Thompson, N. (2024). The ai risk repository: A comprehensive meta-review, database, and taxonomy of risks from artificial intelligence.
- Tarallo, F. (1985). *A pesquisa sociolinguística*. Ática.
- Yasavur, U.; Lisetti, C. and Rische, N. (2014). Let’s talk! speaking virtual counselor offers you a brief intervention. *Journal on Multimodal User Interfaces*, 8(4):381–398.