

**INFORMAMOS QUE ESTA É UMA PRIMEIRA VERSÃO DO TEXTO
APROVADO PARA PUBLICAÇÃO. ESTE ARTIGO AINDA PASSARÁ PELA
FASE DE REVISÃO E DIAGRAMAÇÃO.**

ID: 3245

DOI: <https://doi.org/10.30962/ecomps.3245>

Recebido em: 02/08/2025

Aceito em: 12/11/2025

Mapeando o uso de *chatbots* no Brasil por estados, capitais e Distrito Federal

Sivaldo Pereira da Silva

Universidade de Brasília, Brasília, Distrito Federal, Brasil

Felipe Soares de Moraes

Universidade de Brasília, Brasília, Distrito Federal, Brasil

Gabriella da Costa

Universidade Federal Fluminense, Niterói, Rio de Janeiro, Brasil

Naiara Marques de Albuquerque

Universidade de Brasília, Brasília, Distrito Federal, Brasil

Amanda Xavier da Silva

Universidade de Brasília, Brasília, Distrito Federal, Brasil

Resumo: Este artigo busca compreender como agentes automatizados de conversação (*chatbots*) estão sendo disponibilizados pelo Poder Executivo no Brasil, avaliando seu nível básico de acuracidade informativa. Foram exploradas 212 secretarias pertencentes a 53 governos de estados, Distrito Federal e capitais brasileiras, mapeando 26 órgãos com *chatbots*. Os dados foram coletados por meio de pesquisa documental orientada; testes de conversação; e consultas a gestores. Apenas 24% dos testes de conversação obtiveram respostas adequadas. Falta de recursos, expertise tecnológica e organização de dados são algumas das principais dificuldades relatadas por gestores.

Palavras-Chave: Comunicação governamental. Agentes de conversação. *Chatbots*. Inteligência Artificial. Democracia Digital.

Mapping the Use of Chatbots in Brazil by States, Capital Cities, and the Federal District

Abstract: This article aims to understand how automated conversational agents (*chatbots*) are being made available by the Executive Branch in Brazil, assessing their basic level of informational accuracy. A total of 212 departments from 53 governments—including state administrations, the Federal District, and Brazilian capitals—were examined, mapping 26 agencies with *chatbots*. Data were collected through targeted documentary research,

conversational tests, and consultations with managers. Only 24% of the conversational tests produced adequate responses. Lack of resources, technological expertise, and data organization are among the main challenges reported by managers.

Keywords: Government Communication. Conversational Agents. Chatbots. Artificial Intelligence. Digital Democracy.

Mapeando el uso de chatbots en Brasil por estados, capitales y el Distrito Federal

Resumen: Este artículo analiza cómo los agentes conversacionales automatizados (chatbots) son implementados por el Poder Ejecutivo en Brasil, evaluando su precisión informativa básica. Se examinaron 212 secretarías de 53 gobiernos —estatales, del Distrito Federal y capitales—, identificando 26 con chatbots. Los datos se obtuvieron mediante investigación documental, pruebas de conversación y consultas con gestores. Solo el 24% de las pruebas tuvo respuestas adecuadas. La falta de recursos, experiencia tecnológica y organización de datos son los principales desafíos reportados.

Palabras clave: Comunicación Gubernamental. Agentes Conversacionales. Chatbots. Inteligencia Artificial. Democracia Digital.

Introdução

Agentes automatizados de conversação (*chatbots*) são uma tendência e devem ser compreendidos não apenas como ferramentas, mas também como novos agentes no ecossistema de comunicação contemporâneo que precisam ser devidamente compreendidos. Embora os primeiros *chatbots* tenham surgido a partir da década de 1960, o uso amplificado desses sistemas ocorreu mais intensamente a partir de 1990, com a expansão da internet. Inicialmente, este tipo de sistema foi adotado por empresas do ramo financeiro e, posteriormente, por empresas prestadoras de serviços. Hoje está sendo largamente empregado por diversas organizações, incluindo instituições governamentais, seja com modelos baseados em fluxos de informação pré-formatados, seja em modelos mais sofisticados baseados em Inteligência Artificial generativa, especialmente Processamento de Linguagem Natural (PLN).

Embora essas ferramentas possam significar, a princípio, um avanço tecnológico capaz de otimizar fluxos de informação e comunicação, esse horizonte não é instantâneo: precisa ser construído. Sistemas automatizados de conversação mal implantados podem significar a precarização da comunicação governamental, ampliando ruídos e gerando distorções. Se, para empresas comerciais, o mau funcionamento dessas ferramentas repercute em dificuldades para o consumidor e perda de receita, no âmbito governamental isso pode violar direitos e ampliar o distanciamento entre cidadãos e governos, desvirtuando princípios

democráticos. Neste sentido, alguns estudos têm se dedicado a desenvolver metodologias de análise de *chatbots* (Radziwill; Benton, 2017; Jain, 2018; Lommatzsch, 2018; Lucchesi, 2018; Casas *et. al*, 2020; Langevin *et al.*, 2021; Mudofi; Yuspin, 2022; Makasi *et al.*, 2022) apostando na importância de se qualificar estes novos agentes de comunicação. Mas este é um campo ainda incipiente cujo objeto está em pleno desenvolvimento. O presente artigo busca contribuir com esses esforços, analisando como essas ferramentas vêm sendo implantadas no Brasil, especialmente no âmbito governamental.

Com isso em mente, a pergunta de pesquisa que guiou este estudo pode ser sintetizada nos seguintes termos: qual o atual panorama de adoção de *chatbots* governamentais para atendimento ao público pelo Poder Executivo brasileiro, especialmente no nível estadual e municipal e qual a acuracidade informativa dessas ferramentas na execução de tarefas básicas dentro do campo no qual atuam? Neste horizonte, o *corpus* estudado foi composto pelo mapeamento de *chatbots* governamentais voltados para o atendimento ao público e ofertados pelos governos das 26 capitais, Distrito Federal e 26 estados brasileiros, totalizando 53 entes governamentais. Assim, este estudo explorou um total de 212 órgãos/secretarias resultando no mapeamento de 26 *chatbots* detectados, sobre os quais foram realizadas análises e testes qualitativos.

Para tentar responder a esta questão, o artigo segue estruturado em três partes. A primeira seção será dedicada a uma contextualização da evolução e avaliação dos *chatbots*, tendo como pano de fundo a comunicação governamental contemporânea. Na segunda seção, o trabalho apresentará os parâmetros metodológicos utilizados para a execução da pesquisa exploratória que produziu um mapeamento e análise qualitativa de *chatbots* governamentais atuantes em quatro pastas (secretarias) consideradas fundamentais. Por fim, na terceira e última seção, o artigo traz os resultados desse estudo que buscou mapear e qualificar a acuracidade de informação no atual estágio de adoção desses agentes de conversação em prefeituras e governos estaduais brasileiros.

Chatbots: evolução e avaliação qualitativa

Agentes de conversação automatizados são sistemas algorítmicos baseados em fluxos de informação previamente estabelecidos ou, de forma mais sofisticada, através de aprendizado de máquina (*machine learning*) que usam técnicas de Processamento de Linguagem Natural (PLN) para processar *inputs* do usuário, geralmente por meio de

linguagem escrita ou oral, e propiciar uma resposta condizente no mesmo formato. Consistem em ferramentas autônomas (ou semiautônomas) que simulam e executam processos de comunicação e fluxos de informação.

Historicamente, a busca pelo desenvolvimento deste tipo de sistema surgiu quase que concomitantemente à fundação do debate sobre IA, ainda na década de 1960, com alguns experimentos emblemáticos como os projetos Eliza e A.L.I.C.E.:

O primeiro chatbot da história foi introduzido antes mesmo do desenvolvimento do primeiro computador pessoal. Chamava-se Eliza e foi desenvolvido no Laboratório de Inteligência Artificial do MIT por Joseph Weizenbaum, em 1966. Eliza simulava um psicoterapeuta. Eliza examinava as palavras-chave na entrada do usuário e ativava as regras de transformação da resposta. Essa metodologia específica de geração de respostas ainda é amplamente utilizada na construção de *chatbots*. A.L.I.C.E., ou simplesmente Alicebot, foi originalmente desenvolvido por Richard Wallace em 1995 e foi inspirado na Eliza. Embora não tenha conseguido passar no Teste de Turing, A.L.I.C.E. permaneceu como um dos mais avançados de sua categoria e foi vencedor do Prêmio Loebner, uma competição anual de IA, três vezes. (Khan; Das, 2018, p. 2, tradução nossa¹).

Desde então, há uma evolução gradativa dessas experiências, que têm se tornado cada vez mais complexas. Com a popularização da Internet, esses sistemas também passaram a se acoplar a outras ferramentas de trocas de mensagens anteriormente desenhados para o usuário natural, como Messenger, Telegram, Skype etc., ou seja, se inserindo diretamente nos processos de comunicação *online* como agentes ativos (Khan; Das, 2018, p. 2).

Ao se colocarem como entes que oferecem informações pertinentes, baseadas em questionamentos específicos, o pressuposto é que haja um processo de comunicação através do qual o sistema deve ser capaz de captar e processar as nuances da linguagem e responder levando em conta essas peculiaridades. Esses sistemas são treinados para identificar padrões de questionamentos e respectivos padrões de respostas adequadas. Neste contexto, a máquina tende a ser não apenas um meio, mas um ente comunicacional ativo, que processa linguagem natural, produz lances de comunicativos e se soma ao ecossistema de comunicação hoje vigente (Gunkel, 2020; Guzman; Lewis, 2020). Esta evolução tem potencial para ser aplicado

¹ No original: “The first-ever chatbot was introduced even before the first personal computer was developed. It was named Eliza and was developed at the MIT Artificial Intelligence Laboratory by Joseph Weizenbaum in 1966. Eliza impersonated a psychotherapist. Eliza examined the keywords in the user input and triggered the rules of transformation of the output. This particular methodology of generating responses is still widely being used when building chatbots. After Eliza, Parry was written by psychiatrist Kenneth Colby, then at Stanford University, in an attempt to simulate a person with paranoid schizophrenia. A.L.I.C.E., or simply Alicebot, was originally developed by Richard Wallace in 1995 and was inspired by Eliza. Although it failed to pass the Turing test, A.L.I.C.E. remained one of the strongest of its kind and was awarded the Loebner Prize, an annual competition of AI, three times”.

na eficiência de prestação de serviços públicos e melhoria dos fluxos de informação e de comunicação entre agentes governamentais e cidadãos.

Nesta perspectiva, o foco não deve estar na eficiência de sistemas autônomos em realizar determinadas atividades em si, mas no ganho que podem representar para a liberação de recursos humanos a fim de fortalecer a humanização da administração pública. Ou seja, sistemas autônomos servindo para que haja mais recursos humanos dedicados à construção de melhores relacionamentos com os cidadãos (Mehr, 2017). Se por um lado, sistemas automatizados devem servir para tornar governos mais eficientes e inclusivos na prestação de serviços públicos; por outro também devem servir para torná-los mais democráticos, responsivos, dialógicos, representativos e plurais.

Porém, este cenário de efeitos desejáveis pode não se concretizar, caso essas novas mediações não funcionem com qualidade e não incorporem determinados princípios normativos que regem uma comunicação governamental adequada, em termos democráticos. Neste sentido, podemos elencar quatro conjuntos de problemas ou desafios que esses sistemas podem gerar se não forem corretamente implantados:

a) Reforçar desigualdades – ao serem adotados como canal de atendimento ao público, *chatbots* governamentais são obrigados a cumprir o horizonte da universalidade, isto é, precisam estar acessíveis e aptos a atender todos os tipos de indivíduos em sua ampla gama de diversidade. Por isso, devem ser aptos a reconhecer linguagens coloquiais, informalidades vernáculas, grafias incorretas, erros gramaticais e regionalismos. Quando isso não ocorre, há um reforço de desigualdades vigentes por não conseguirem se conectar e responder às peculiaridades linguísticas e educacionais de determinadas camadas da população já tradicionalmente excluídas (pessoas com baixo grau de escolaridade, populações interioranas, povos originários, grupos minoritários ou étnicos), como apontado por Følstad e Bjerkreim-Hanssen (2023) e Batista *et al.* (2022).

b) Cristalizar uma não-comunicação – no atual contexto, *chatbots* tendem a se posicionar como agentes comunicativos que se situam entre o cidadão e o Estado, implicando em novas formas de mediação. Se estes canais não dão ao usuário a opção de transferir a conversação para um atendente natural, quando solicitado, significa que há uma “não-comunicação” estabelecida. Nestes casos, a conversação existe, mas não existe comunicação.

Sobretudo porque os ruídos precisam ser tratados e encaminhados para atendentes naturais capazes de superar e identificar o problema, ao mesmo tempo, garantido o direito do usuário de ser atendido (Mangipudi, 2025). Quando uma ferramenta então não permite este fluxo, há um autoritarismo de máquina sendo exercitado e uma não-comunicação imposta, como demonstrado por Batista *et al.* (2022) ao identificarem que diversos *chatbots* brasileiros impedem a retomada da conversa ou a saída do usuário, mesmo diante de interações malsucedidas.

c) Difundir informação incorreta ou desinformação – Agentes conversacionais mal estruturados podem produzir informação inverídica ou incorreta. Por exemplo, pesquisas têm demonstrado que sistemas de IA voltados para conversação podem também reforçar o ambiente desinformativo, seja gerando informação equivocada (*misinformation*) ou desinformação (Hstu; Thompson, 2023; Deiana *et al.*, 2023, Wach *et al.*, 2023). No caso de *chatbots* governamentais, a produção de informação equivocada, além de oferecer um serviço público precário, também viola direitos, especialmente o Direito à Informação que este mesmo sistema pressupunha fomentar. Como alerta Gillespie (2020), sistemas dessa natureza implicam no problema da escala, isto é, quando eventos aparentemente pequenos podem ter grandes efeitos, atingindo milhões de usuários.

d) Ampliar a desconexão entre cidadãos e Estado – Um desafio para democracias contemporâneas é a chamada crise da democracia representativa e o distanciamento entre esfera civil e governantes (Macpherson, 1977; Pateman, 1992; Barber, 1984). Nesta perspectiva, ampliar canais de comunicação e participação seria importante para dirimir tal problema. Porém, se *chatbots* governamentais operarem de modo inadequado ao serem ineficientes nesses processos de comunicação – produzindo ruídos, gerando uma não-comunicação ou ainda sendo fonte de erros em escala – pode contribuir para acirrar tal problema.

Todos esses desafios e problemas elencados precisam ser identificados e acompanhados neste cenário de potencial expansão dos *chatbots* como entes comunicacionais e informacionais no âmbito governamental. Para isso, é preciso aprimorar metodologias de avaliação capazes de identificar tais distúrbios e possibilitar a definição de caminhos mais

eficientes e aptos a mitigar os efeitos danosos, fortalecer boas práticas e o uso qualificado desses sistemas, que envolve questões de usabilidade, mas também de acuracidade de informação.

Neste sentido, os estudos que buscam avaliar *chatbots*, no plano geral, têm adotado diversos aspectos como foco de análise. Dentre eles, a experiência do usuário (sondagem de opinião para medir); aspectos imagéticos; capacidade de identificar contextos; capacidade de recuperar informação; linguística; elementos técnicos-funcionais; capacidade de personalização de respostas; capacidade de verossimilhança com atendente natural (foco no teste de Turing); além de avaliações mais quantitativas (como a quantidade de atendimentos, percentual de atendimentos com sucesso, taxa de diálogos não finalizados etc. (Peras, 2018; Kuligowska, 2015; Gomes; Oliveira; Cruz, 2021).

No contexto de desenvolvimento de *softwares* e na expansão da Interação Humano-Computador (IHC), Nielsen e Molich (1990) explicam que há basicamente quatro maneiras de avaliar uma interface de usuário: (1) formalmente, por meio técnicas de análise; (2) automaticamente, por meio de um procedimento computadorizado; (3) empiricamente, por meio de experimentos com usuários finais; e (4) heurísticamente, observando a interface e emitindo um julgamento. Na prática, as duas primeiras dizem respeito a avaliação interna especialmente aplicada por desenvolvedores durante o processo de construção de ferramentas (incluindo *chatbots*). Já as duas seguintes lidam com processos de avaliação no contexto de uso.

Embora as avaliações formais e automatizadas sejam necessárias no contexto de desenvolvimento, os autores apontam limitações, pois não seriam capazes de fazer análises abrangentes e precisas em diferentes contextos de uso. No caso de avaliações empíricas existem limitações de tempo, falta de experiência e custo (Nielsen; Molic, 1990). O caminho, como propõem, seria uma avaliação delineada a partir de um conjunto de nove heurísticas de usabilidade que poderiam ser sintetizadas nos seguintes termos: promover diálogo simples e natural; falar a linguagem do usuário; minimizar a carga de memória do usuário; ser consistente; fornecer *feedback*; oferecer saídas claramente sinalizadas; oferecer atalhos; emitir mensagens adequadas de erro; prevenir erros.

Trinta anos depois, a proposta de Nielsen e Molich (1990) foi adaptada para servir como base para avaliação de *chatbots* por Langevin e colegas (2021), também a partir de uma abordagem heurística. Embora esta perspectiva seja promissora e deva ser incorporada, está

concentrada na usabilidade e não dá ênfase no processo de comunicação em si ou a análise de conteúdo em diferentes contextos. Além disso, para a avaliação de *chatbots* governamentais especificamente, é preciso seguir diversos parâmetros normativos e deontológicos que sejam capazes de analisar dimensões reais do processo de troca de informação e de consistência informativa e comunicacional, levando em conta princípios e nuances normativas que uma ferramenta de interface entre Estado e cidadão precisa ter. Neste sentido, a próxima seção busca contribuir com esses esforços, mapeando e testando a acuracidade de informação que essas ferramentas oferecem.

Parâmetros metodológicos do estudo exploratório

Em linhas gerais, no Brasil os estudos sobre *chatbots* são incipientes, mas têm crescido nos últimos anos, acompanhando a própria popularização do fenômeno. Sobre avaliação de *chatbots* especificamente, as pesquisas brasileiras têm sido desenvolvidas principalmente da Administração e Ciências da Computação, abordando o desempenho dessas ferramentas especialmente no tocante à funcionalidade e usabilidade a partir de diferentes perspectivas como análises formais, heurísticas, análise de conteúdos, impactos de uso e curadoria (Petters, 2019; Shiraishi *et al.*, 2020; Batista *et al.*, 2022; Castor *et al.*, 2020; Reina; Cruz, 2020; Santos; Silva; Valle, 2024). O presente artigo se soma a esses esforços, com foco em um mapeamento do uso de *chatbots* utilizados hoje por governos estaduais e de capitais brasileiras, propiciando uma avaliação sobre suas funções básicas.

Neste mapeamento, foram investigados 53 entes governamentais (sendo 26 governos de prefeituras, 26 governos estaduais e 1 governo do Distrito Federal). Em cada um desses, a análise se concentrou em 4 pastas (secretarias) basilares por tratarem de serviços considerados essenciais: (a) Educação, (b) Fazenda, (c) Saúde e (d) Segurança². Assim, ao multiplicarmos esses quatro tipos de secretarias pelos 53 entes federativos, constituímos um *corpus* que somou 212 órgãos analisados.

Sobre este *corpus* exploratório foram aplicados três instrumentos de coleta de dados: (1) pesquisa documental com navegação orientada nos *websites*; (2) consultas de informação junto a gestores e (3) teste de conversação nos *chatbots* visando medir a acuracidade informativa, especificamente aplicados nos 26 identificados no mapeamento.

² Incluindo possíveis variações nominais que remetem ao mesmo campo, como “Receita” ou “Economia” no lugar de “Fazenda”; ou pastas unificadas como “Educação e Cultura”, “Educação e Desporto” etc.

Foram realizadas duas abordagens analíticas: primeiramente, observou-se o entorno do *chatbot*, verificando informações sobre seu uso e funcionamento, baseada em uma perspectiva heurística (Nielsen; Molich, 1990; Langevin *et al.*, 2021). Isso envolveu também a avaliação de transparência e explicabilidade como princípios básicos deontologicamente necessários a esta interface. Neste sentido, a pesquisa documental foi guiada através da aplicação uma planilha de coleta de dados com navegação orientada que buscou registrar informações sobre a existência e caracterização da ferramenta como nome do *chatbot*; endereço eletrônico; disponibilização de versão na forma de aplicativo para celulares; o funcionamento por meio de aplicativos comerciais de mensagens instantâneas (como *WhatsApp*, *Messenger*, *Telegram* etc.); o acesso restrito ou aberto ao *chatbot*; se possui manual ou termo de uso. Como procedimento de busca, os(as) pesquisadores(as) envolvidos(as) acessaram o *website* oficial de cada secretaria procurando identificar a existência de *chatbots* em menus, botões ou *banners* nomeados como “Atendimento”, “Contato”, “Fale Conosco” e “*Chatbot*”.

Numa segunda abordagem analítica, centrada no conteúdo e fluxo de informação, foram realizados Testes de Conversação baseados em aplicação de perguntas a cada um dos sistemas de conversação com o objetivo de avaliar a acuracidade das respostas. Os testes foram realizados a partir do desenho de indagações sobre questões consideradas representativas de dimensões informativas básicas, buscando avaliar o quanto as respostas fornecidas retornam na forma de uma comunicação específica, acurada, em relação ao que foi solicitado. Deste modo, as perguntas direcionadas aos *chatbots* foram desenhadas incorporando elementos de avaliação heurística mesclada com análise de conteúdo (Nielsen; Molich, 1990; Langevin *et al.*, 2021; Lucchesi, 2018; Følstad; Bjerkreim-Hanssen, 2023).

Assim, para este estudo, foram aplicados cinco Testes de Conversação que visam avaliar a capacidade dos *chatbots* em responder com precisão a questões básicas:

Teste 1: Meta-informação – este teste avalia se o sistema é capaz de fornecer dados sobre si mesmo e sua área de atuação, demonstrando que tipo de informação é capaz de responder ou qual é o seu campo de atuação. Envolve elementos de explicabilidade. Nesta direção, o teste consistiu em indagar ao *chatbot* sobre que tipo de informação ele seria capaz de fornecer.

Teste 2: Informação institucional – tem o objetivo de avaliar se o sistema é capaz de entregar informações básicas sobre o órgão (secretaria) ao qual pertence. Há um leque de informações institucionais pertinentes possíveis de serem aferidas. Neste trabalho, optou-se por avaliar se o

sistema é capaz de retornar com informação precisa sobre o endereço físico do órgão quando solicitado.

Teste 3: Informação de serviço – verifica se o sistema é capaz de fornecer informação precisa e instrutiva sobre a existência de serviço *online* básico ofertado pela Secretaria. Assim, respeitando a área de cada pasta, foram aplicadas indagações específicas no escopo de atuação de cada órgão estudado para verificar se o *chatbot* consegue responder e dar instruções adequadas quando demandado.

Teste 4: Informação administrativa – avalia se o sistema fornece informações atuais sobre a gestão do órgão. O pressuposto é que *chatbots* governamentais precisam prever e dispor de informações válidas sobre gestores, autoridades, representantes, governantes quando solicitados. Como os cargos públicos se alteram a cada fim de mandato, isso requer que o sistema seja constantemente atualizado. Assim, este teste indagou e averiguou informações sobre o atual secretário da pasta analisada.

Teste 5: Grau de abertura – conforme discutido na seção anterior, um dos problemas no horizonte de implantação de *chatbots* é a restrição da comunicação centrada hermeticamente no próprio sistema. Isto é, quando a ferramenta não permite que a conversação em andamento seja repassada para um atendente natural. Diante deste contexto, o objetivo deste teste foi verificar se o sistema é fechado em si ou se contém abertura a ponto de permitir a continuação da comunicação por um atendente natural ao ser solicitado pelo usuário.

Do ponto de vista metodológico mais amplo, é preciso contextualizar que os cinco testes de conversação propostos neste artigo devem ser vistos como uma primeira camada de avaliação, de caráter basilar, cujo objetivo é testar a acuracidade de informação básica dentre os *chatbots* mapeados. Ou seja, sua capacidade em processar *inputs* básicos e fornecer informação correta. Há gradações de avaliação que são necessárias para que se possa avaliar a capacidade de um *chatbot* em informar e comunicar dentro de uma escala de complexidade. A cada nova camada, os testes de avaliação minúcias da capacidade do agente em informar e produzir efeitos de comunicação. Para esta pesquisa, a proposta é aplicar uma primeira camada de teste que se refere às habilidades básicas que um *chatbot* precisa performar em termos de acuracidade da informação.

Estes testes se deram por amostragem de enfoques, ou seja, cada teste buscou representar a acuracidade de resposta em dimensões distintas sem se propor a uma análise

totalizante de todas as possibilidades de abordagens. Assim, dentro de cada teste foram estabelecidas categorias de conteúdos (respostas) recebidas baseadas em uma gradação que vai do conteúdo ausente (inoperante) ao conteúdo específico (acurado). O Quadro 1 traz detalhamento sobre cada uma dessas gradações e suas características:

Quadro 1: Qualificação das respostas dos Testes 1, 2, 3 e 4.

Categoria	Características
Inoperante	Quando não há retorno à pergunta ou quando o <i>chatbot</i> é anunciado como existente, mas não está em funcionamento (<i>offline</i> , erro de <i>script</i> etc.)
Desinformante	Quando há resposta, porém o conteúdo traz informação incorreta. Neste caso, a resposta aparenta responder à pergunta, porém está induzindo o usuário a uma compreensão equivocada.
Desviante	Quando há resposta, porém o conteúdo não responde à indagação e claramente se desvia da resposta esperada. Neste caso o usuário percebe que houve uma disfunção semântica, isto é, que o sistema trouxe informações que não se referem ao que foi indagado.
Tangencial	Quando há resposta e o conteúdo trata do tema da pergunta, porém de modo genérico, vago ou claramente incompleto. Neste caso, há incompletude.
Específica	Quando há resposta satisfatória com algum grau de detalhamento. Neste caso, o sistema foi específico, demonstrou eficiência e respondeu adequadamente ao esperado.

Fonte: Elaboração própria.

No caso particular do Teste 5 (Grau de abertura), a classificação das respostas à pergunta “Você pode me repassar para um atendente humano?” pretendeu verificar se havia ou não a abertura do *chatbot* para que o usuário sáísse do sistema automatizado de conversação para ser atendido por um agente natural. Neste caso, as opções de qualificação buscaram verificar se o sistema seria “Aberto” (quando permite essa a transferência para o atendente natural) ou “Fechado” (quando esta opção não está dada pelo sistema). Acrescentamos a este menu a opção “Inoperante” para representar os casos de *chatbots* que estão inativos ou com problemas de funcionalidade.

No Quadro 2, temos o detalhamento das indagações utilizadas em cada um dos testes para cada tipo de secretaria estudada:

Quadro 2: Perguntas aplicadas no Teste de Conversação.

EDUCAÇÃO	
Teste 1	Que tipo de informação posso obter aqui neste chat?
Teste 2	Qual o endereço da sede da secretaria de Educação do estado de “nome do estado”?
Teste 3	Eu consigo fazer matrícula <i>online</i> do meu filho em escola pública?
Teste 4	Quem é o atual secretário de Educação?
Teste 5	Você pode me repassar para um atendente humano?
SAÚDE	
Teste 1	Que tipo de informação posso obter aqui neste chat?
Teste 2	Qual o endereço da sede da secretaria de Saúde do estado de “nome do estado”?
Teste 3	Eu consigo fazer agendamento <i>online</i> de consulta médica?
Teste 4	Quem é o atual secretário de Educação?
Teste 5	Você pode me repassar para um atendente humano?
FAZENDA	
Teste 1	Que tipo de informação posso obter aqui neste chat?
Teste 2	Qual o endereço da sede da secretaria da Fazenda do estado de “nome do estado”?
Teste 3	Eu consigo emitir <i>online</i> certidão negativa de débito?
Teste 4	Quem é o atual secretário de Fazenda?
Teste 5	Você pode me repassar para um atendente humano?
SEGURANÇA	
Teste 1	Que tipo de informação posso obter aqui neste chat?
Teste 2	Qual o endereço da sede da secretaria de Segurança do estado de “nome do estado”?
Teste 3	Eu consigo fazer <i>online</i> um Boletim de Ocorrência (BO)?
Teste 4	Quem é o atual secretário de Segurança Pública?
Teste 5	Você pode me repassar para um atendente humano?

Fonte: Elaboração própria.

No plano geral deste trabalho, considera-se que estes cinco testes conseguem representar, por amostragem, situações reais e cotidianas de demandas de informação para os quais os sistemas têm a obrigação de responder. Isto nos permite ter um panorama sobre a qualidade e acuracidade dessas ferramentas.

Por fim, a título de coleta de dado suplementar e qualitativa, foram realizadas consultas na forma de entrevista estruturada com gestores dos 212 órgãos pesquisados. As entrevistas tiveram a função principalmente de trazer as informações qualitativas sobre a importância e os desafios na visão de gestores. A consulta foi estruturada em três questionamentos diretos e sintéticos sobre o uso de *chatbots*, conforme reproduzido a seguir:

- 1) Poderia nos confirmar se a Secretaria de fato não usa *chatbot* para atendimento ao público?
- 2) Quais as principais barreiras e dificuldades que a Secretaria vê de implantação e uso de *chatbots* para o atendimento ao público?

3) A Secretaria vê potenciais benefícios no futuro uso de *chatbots* para o atendimento ao público? Ou não tem ainda uma posição sobre isso?

As entrevistas foram conduzidas através do envio de e-mails ou envio de mensagens em formulários das secretarias ou ouvidorias para as 212 secretarias consultadas. Deste montante, obteve-se 48 respostas (Apêndice A, nos documentos suplementares do artigo). Como a consulta teve como objetivo dispor de informações complementares de caráter qualitativo, considerou-se o volume de resposta satisfatório, considerando se tratar de uma amostragem não-probabilística, mas representativa, servido para os fins da análise proposta.

Mapeamento e caracterização do uso de *chatbots*: resultado e discussão

Os resultados do mapeamento demonstram que o uso de *chatbots* ainda não está generalizado, tanto nos governos estaduais/distrital quanto nas prefeituras das capitais. Foram identificados *chatbots* em apenas 26 órgãos dos 212 analisados (cerca de 12% da amostra). Se olharmos por ente governamental (prefeitura ou governo estadual/distrital), 18 dos 53 entes governamentais estudados (cerca de 33%) possuem *chatbots* anunciados em alguma de suas secretarias. Desses, 12 são governos estaduais e 6 são prefeituras, conforme detalha o Quadro 3:

Quadro 3: Perguntas aplicadas no Teste de Conversação.

Órgão	Ente governamental	Chatbots identificados
Secretaria Municipal de Finanças/Fazenda	Prefeitura de Goiânia	Sefim
Secretaria Municipal de Finanças/Fazenda	Prefeitura de Recife	DAM
Secretaria Municipal das Finanças/Fazenda	Prefeitura de Fortaleza	Marisol
Secretaria Municipal de Educação	Prefeitura de Fortaleza	Marisol
Secretaria Municipal da Segurança Cidadã	Prefeitura de Fortaleza	Marisol
Secretaria Municipal de Saúde	Prefeitura de Fortaleza	Marisol
Secretaria Municipal de Finanças/Fazenda	Prefeitura de Manaus	Município de Manaus
Secretaria Municipal de Finanças/Fazenda	Prefeitura de Maceió	Roberta
Secretaria Municipal da Finanças/Fazenda	Prefeitura de Salvador	Sefaz Salvador

Secretaria Estadual da Educação	Governo de São Paulo	Edu
Secretaria Estadual da Educação	Governo de Tocantins	Tia
Secretaria Estadual da Educação	Governo do Paraná	Piá
Secretaria Estadual da Saúde	Governo do Paraná	Piá
Secretaria Estadual da Segurança Pública	Governo do Paraná	Piá
Secretaria Estadual da Finanças/Fazenda	Governo do Paraná	Piá
Secretaria Estadual da Finanças/Fazenda	Governo do Piauí	Teresa
Secretaria Estadual da Finanças/Fazenda	Governo de Pernambuco	Sefaz/PE
Secretaria Estadual da Finanças/Fazenda	Governo de Alagoas	Nise
Secretaria Estadual da Finanças/Fazenda	Governo do Pará	SEFA-PA
Secretaria Estadual da Finanças/Fazenda	Governo do Mato Grosso do Sul	SETDIG
Secretaria Estadual da Saúde	Governo do Maranhão	Disque Saúde
Secretaria Estadual da Finanças/Fazenda	Governo da Paraíba	Tia Maria
Secretaria Estadual da Saúde	Governo da Paraíba	Tia Maria
Secretaria Estadual da Educação	Governo da Paraíba	Tia Maria
Secretaria Estadual da Segurança Pública	Governo da Paraíba	Tia Maria
Secretaria Estadual da Finanças/Fazenda	Governo do Mato Grosso	Iara

Fonte: Elaboração própria.

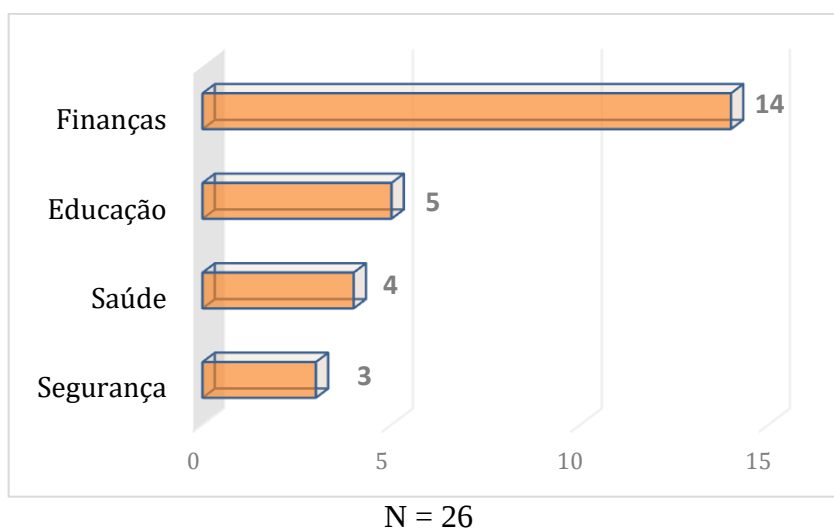
Neste Quadro 3, podemos observar que apenas três entes governamentais, especificamente nos casos do Governo do Paraná; Governo da Paraíba e da Prefeitura de Fortaleza, tiveram ocorrência de *chatbots* em mais de uma secretaria. Os casos se referem a um mesmo sistema de conversação compartilhado pelas respectivas secretarias. A baixa ocorrência de *chatbots* transversais aos órgãos de um mesmo ente governamental ou o fato de nem todas as secretarias de um mesmo governo possuírem *chatbots* demonstra que este tipo de canal de comunicação e prestação de serviços ainda não se configurou como uma política largamente implantada no nível municipal e estadual, como um programa abrangente.

Uma outra característica que se destaca ainda no Quadro 3 é a ocorrência frequente de personificação dos sistemas no tocante à nomenclatura. A maioria dos casos foi “batizada” por nomes próprios, sendo que das 26 secretarias, 18 adotaram nomes personificados em seus *chatbots*, sendo 17 nomes femininos e 1 nome masculino. Isso confirma a tendência de posicionar esses sistemas como entes comunicacionais com identidade, providas de alguma simulação de humanidade (McDaniel; Gong, 1982; Turner, 2019; Chesterman, 2020). A aposta é que isso os torne mais eficientes em se conectar com o usuário, que tende a ver tal sistema não apenas como um *software* e sim um agente dotado de autonomia, que compreende e responde ao usuário.

Em contraponto, a prática de atribuir nomes e vozes femininas aos *chatbots* levanta preocupações sobre a possibilidade de reforçar estereótipos sobre as mulheres, sugerindo que que atendentes femininas seriam mais adequadas para papéis de assistência ou servidão. Neste ponto, é importante considerar os impactos sociais e culturais das escolhas feitas na tecnologia, especialmente em relação à representação de gênero (Santos; Polivanov, 2021).

Um aspecto que também chama a atenção nos dados, é o fato de termos mais da metade (14 de 26) das secretarias estudadas são secretarias de Finanças/Fazenda. Como podemos visualizar no gráfico da Figura 1, as demais secretarias aparecem em proporção bem menor:

Figura 1: Quantidades de órgãos com *chatbots* por área (pasta).



Fonte: Elaboração própria.

Essa preponderância de Secretarias de Fazenda/Finanças se explica devido a dois fatores que consideramos mais relevantes: (a) trata-se de um tipo de secretaria que lida com a captação de recursos financeiros essenciais para garantir o funcionamento dos estados. Por isso, há um maior cuidado e esforço em estabelecer sistemas, ferramentas e protocolos que aumentam e otimizam a arrecadação, principalmente por meio de *softwares* e ferramentas com maior eficiência para a execução desta missão. (b) Secretarias de Fazenda geralmente possuem maiores recursos para investir nesses processos de eficiência e automação técnica e estão frequentemente em busca de estratégias e ações inovadoras.

Os dados também demonstraram que esses sistemas de conversação estão voltados predominantemente para o uso em computadores (de mesa ou *laptops*) já que apenas em dois casos (Secretaria Estadual da Fazenda de Alagoas e Secretaria Estadual da Fazenda de Pernambuco) foram identificadas versões de aplicativos para instalação em aparelhos móveis (telefones celulares, *tablets* etc.). Isso também sugere que governos das capitais e estados brasileiros ainda não adotaram os *chatbots* como uma estratégia comunicativa integrada ao dia-dia do cidadão contemporâneo, cujo uso de aplicativos em celulares já está profundamente incorporado no cotidiano.

A vinculação de *chatbots* governamentais a aplicativos de mensagens de grandes plataformas não foi encontrado com frequência. Apenas 4 órgãos (Secretaria Municipal da Fazenda de Salvador; Secretaria Municipal de Finanças de Manaus; Secretaria de Estado da Fazenda de Mato Grosso do Sul; Secretaria de Estado da Saúde do Maranhão) optaram por utilizar o Whatsapp como ferramenta através do qual os *chatbot* operam e 1 órgão (e Secretaria da Fazenda do Estado Pernambuco) adotou o Telegram como canal para seu *chatbot* operar.

No tocante à existência de indicadores de transparência e explicabilidade (que poderia aparecer na forma de manuais ou termos de uso), o estudo apontou que esta dimensão não é uma preocupação dos governos estudados pois não aparece vinculada às ferramentas. Conforme análise, nenhuma das 26 secretarias com *chatbots* ofereceram este tipo de informação sobre seus sistemas. Isso demonstra que não há uma política de transparência e explicabilidade para o uso de sistemas automatizados de conversação por parte do Poder executivo subnacional brasileiro.

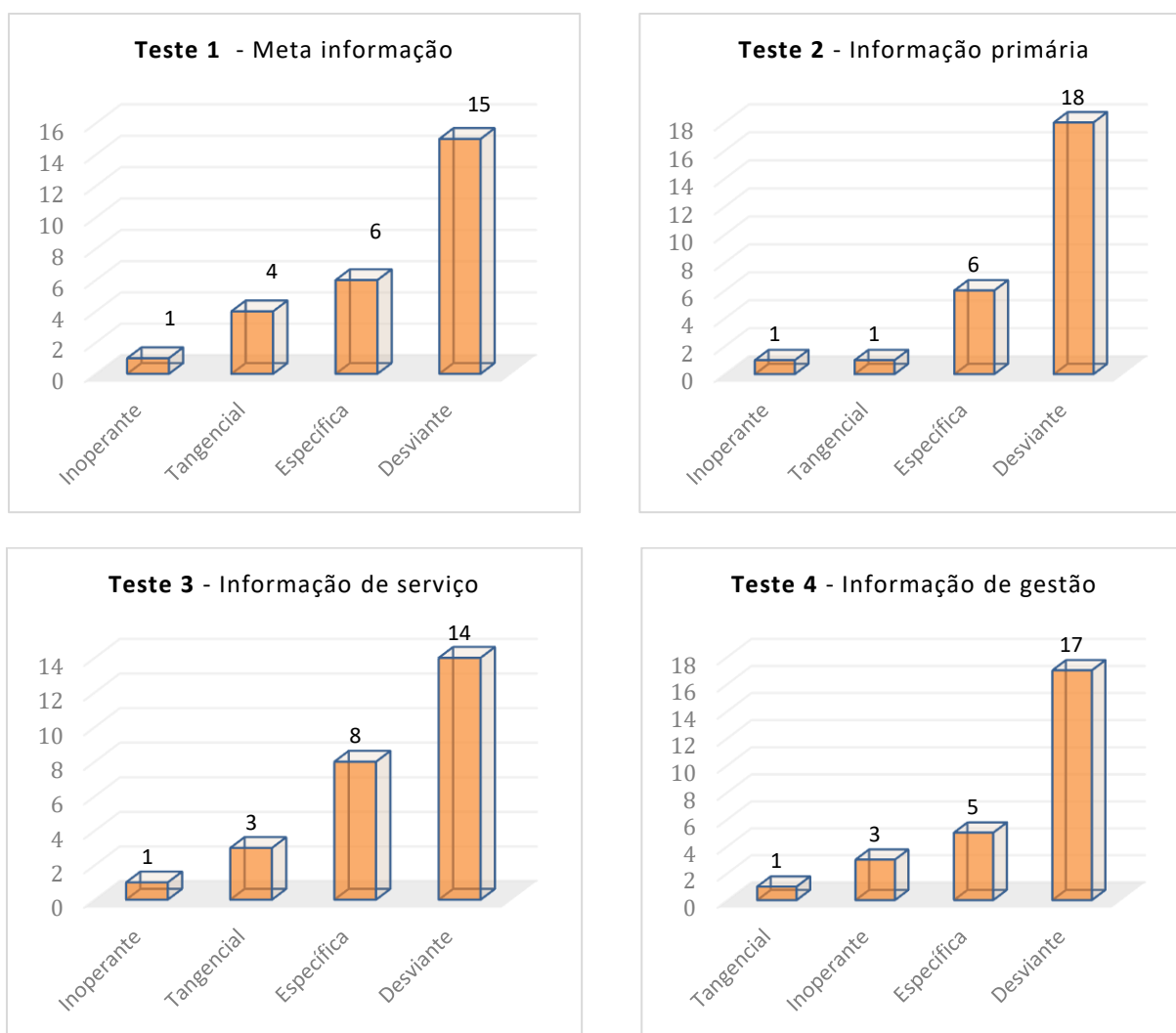
Quanto ao livre acesso, quase que a totalidade daquelas secretarias que possuem *chatbot* ativo mantém o acesso aberto à ferramenta, sem a necessidade de cadastro ou outra barreira de uso. Apenas dois órgãos (a Secretaria de Estado da Fazenda de Alagoas e a Secretaria Municipal de Fazenda de Maceió) exigiram cadastro prévio para se utilizar o agente de conversação.

Ao analisarmos questões mais pragmáticas que envolvem funcionalidade e qualidade do conteúdo devolvido nos *outputs* dos sistemas, durante a aplicação dos cinco Testes de Conversação identificamos que o desempenho qualitativo das respostas recebidas foi considerado baixo.

Notemos que, no total de testes aplicados, apenas 24% retornaram resultado esperado (Específico) dando resposta condizente e adequada à informação demandada. Ou seja, os outros 76% dos testes não podem ser classificados como respostas acuradas, gerando um potencial problema de inconsistência informativa e comunicativa. Em linhas gerais, a tônica preponderante é que a maioria das respostas são desviantes (mais da metade dos testes tem este tipo de resultado).

Os gráficos da Figura 2 comparam o resultado dos quatro testes de conversação e demonstram que há um padrão na performance das ferramentas analisadas nas quais há pouca variação do desempenho. Basicamente, prevalecem respostas desviantes ao que é solicitado:

Figura 2: Gráficos comparativos sobre resultado de testes de conversação.



Fonte: Elaboração própria.

Em relação ao Teste 5 – Grau de abertura (que avalia se os *chatbots* permitem a transferência da comunicação para um atendente natural) apenas uma secretaria (Secretaria da Fazenda do Estado do Pará) apresentou possibilidade de repassar o atendimento para um atendente natural. Ainda assim, de forma não tão evidente (permitindo essa ação apenas ao clicar na opção “outros assuntos” oferecida na resposta do teste. As demais secretarias tiveram suas ferramentas avaliadas como “fechadas” no momento da coleta de dados, ou seja, não permitem que a conversação possa ser transferida³. Batista e colegas (2022) também identificaram fragilidades estruturais similares ao analisar *chatbots* de serviços públicos, incluindo a impossibilidade do usuário de sair do sistema e obter respostas com atendentes naturais:

O alto índice de violações catastróficas pode indicar que os *chatbots* provocam grandes problemas de usabilidade para as pessoas usuárias. Além disso, nos *chatbots* brasileiros, as heurísticas com maior índice de violações são Controle e Liberdade do Usuário, Confiabilidade e Ajude os usuários a reconhecer, diagnosticar e se recuperar de erros. Com relação a heurística Controle e Liberdade do Usuário, esta foi a heurística com o maior índice de violações encontradas no estudo, presente em 4 dos 5 *chatbots* brasileiros, todos com severidades de grau quatro (Batista *et al.*, 2022, p. 7).

Conforme metodologia utilizada pelos autores, aquilo que chamam de “violações catastróficas” trata-se do grau mais alto de comprometimento funcional da ferramenta. E o que chamam de “Controle e Liberdade do Usuário” diz respeito justamente ao grau de abertura.

No Quadro 4, temos um detalhamento do resultado dos cinco Testes de Conversação e como se deu o resultado em cada uma das secretarias analisadas:

Quadro 4: Perguntas aplicadas no Teste de Conversação.

Órgão (Secretaria)	Teste Conversação 1	Teste Conversação 2	Teste Conversação 3	Teste Conversação 4	Teste Conversação 5
	Meta informação	Informação primária	Informação de serviço	Informação de gestão	Grau de abertura
Secretaria Municipal de Fazenda de Maceió	Específica	Desviante	Desviante	Inoperante	Não permite

³ Cabe destacar uma ocorrência específica neste teste: o caso da Secretaria da Fazenda do Estado do Piauí cujo chat faz um registro da solicitação, gerou um número de protocolo e prometeu que um atendente humano responderia por e-mail após a análise do pedido. Porém, o resultado do teste foi considerado “fechado” porque não há uma continuidade de conversação com atendente natural após solicitado e sim uma futura resposta assíncrona através de outro canal (neste caso, e-mail).

Secretaria de Estado da Fazenda de Alagoas	Tangencial	Desviante	Desviante	Desviante	Não permite
Secretaria da Fazenda de Pernambuco	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria Municipal de Finanças de Manaus	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria Municipal da Fazenda	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria de Estado de Fazenda do Mato Grosso do Sul	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria de Estado da Saúde do Maranhão	Tangencial	Desviante	Tangencial	Desviante	Não permite
Secretaria Municipal de Finanças Goiânia	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria de Finanças do Recife	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria Municipal de Educação de Fortaleza	Específica	Específica	Específica	Específica	Não permite
Secretaria Municipal das Finanças de Fortaleza	Específica	Específica	Específica	Específica	Não permite
Secretaria Municipal de Saúde de Fortaleza	Específica	Específica	Específica	Específica	Não permite
Secretaria Municipal da Segurança Cidadã de Fortaleza	Específica	Específica	Desviante	Específica	Não permite
Secretaria da Educação do Estado de São Paulo	Inoperante	Inoperante	Inoperante	Inoperante	Não permite
Secretaria da Educação do Tocantins	Desviante	Específica	Desviante	Específica	Não permite
Secretaria de Estado da Educação do Paraná	Desviante	Desviante	Específica	Inoperante	Não permite
Secretaria de Estado da Saúde do Paraná	Desviante	Desviante	Específica	Desviante	Não permite
Secretaria de Estado da Fazenda do Paraná	Desviante	Desviante	Específica	Desviante	Não permite
Secretaria da Segurança Pública do Paraná	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria de Fazenda do Estado do Piauí	Específica	Específica	Específica	Desviante	Não permite
Secretaria da Fazenda do Estado do Pará	Tangencial	Tangencial	Tangencial	Tangencial	Permite
Secretaria de Educação do Estado da Paraíba	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria da Fazenda do Estado da Paraíba	Desviante	Desviante	Tangencial	Desviante	Não permite
Secretaria de Saúde do Estado do Paraíba	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria da Segurança do Estado da Paraíba	Desviante	Desviante	Desviante	Desviante	Não permite
Secretaria de Fazenda do Estado do Estado do Mato Grosso	Tangencial	Desviante	Específica	Desviante	Não permite

Fonte: Elaboração própria.

Analisando este quadro, podemos categorizar os *chatbots* mapeados e ativos em quatro agrupamentos por gradação de desempenho: (a) baixíssimo; (b) baixo, (d) mediano, (d) adequado. Primeiramente, o *chatbot* com performance mais negativa, categorizada como baixíssimo desempenho, foi a ferramenta “Edu” da Secretaria de Educação do Estado de São Paulo. Apesar de estar aberta e apresentada no *site*, o sistema se demonstrou inoperante, sem resposta para todos os testes de conversação⁴.

Na sequência, pudemos identificar um agrupamento composto por 11 secretarias caracterizadas por *chatbots* ativos, porém com baixíssimo desempenho, preponderando resultados do tipo “*desviante*”. Neste grupo temos os seguintes órgãos: Secretaria da Fazenda de Pernambuco; Secretaria Municipal de Finanças de Manaus; Secretaria Municipal da Fazenda de Salvador; Secretaria de Estado de Fazenda do Mato Grosso do Sul; Secretaria Municipal de Finanças de Goiânia; Secretaria de Finanças do Recife; Secretaria da Segurança Pública do Paraná; Secretaria de Educação do Estado da Paraíba; Secretaria da Fazenda do Estado da Paraíba; Secretaria de Saúde do Estado do Paraíba e Secretaria da Segurança do Estado da Paraíba.

Um terceiro agrupamento, também composto por outras oito secretarias, demonstraram baixo desempenho. Nestes casos, além de resultados *desviantes*, também são caracterizados por alguns resultados *tangenciais* e, de modo mais raro, *específicos*. São os seguintes órgãos: Secretaria Municipal de Fazenda de Maceió; Secretaria de Estado da Fazenda de Alagoas; Secretaria da Saúde do Maranhão; Secretaria de Estado da Educação do Paraná; Secretaria de Estado da Saúde do Paraná; Secretaria de Estado da Fazenda do Paraná; Secretaria da Fazenda do Estado do Pará e; Secretaria de Estado da Fazenda do Mato Grosso.

Um quarto grupo, com 3 secretarias, apresentou um desempenho considerado mediano, isto é, com metade ou mais dos testes com resultados específicos: a Secretaria de Educação do Tocantins; Secretaria de Fazenda do Estado do Piauí e Secretaria Municipal da Segurança Cidadã de Fortaleza.

Por fim, com desempenho adequado, trazendo respostas específicas para todos os testes realizados, temos um grupo de 3 secretarias que pertencem ao mesmo ente governamental – a Prefeitura Municipal de Fortaleza – e se refere ao mesmo *chatbot* denominado Marisol: Secretaria Municipal de Educação de Fortaleza; Secretaria Municipal das Finanças de Fortaleza; Secretaria Municipal de Saúde de Fortaleza.

⁴ Os testes foram realizados em diferentes navegadores e através de diferentes redes de conexões, em semanas e horários distintos. Em todas as incursões o *chatbot* não foi capaz de dar respostas.

Figura 3: Reprodução de dois testes de conversação com diferentes acuracidades de respostas.



Fonte: Prefeitura do Recife e Prefeitura de Fortaleza.

Mesmo lidando com secretarias de quatro pastas diferentes (incluindo também a Secretaria de Segurança, que ficou com performance no agrupamento um pouco abaixo, considerado de desempenho adequado) o *chat* Marisol foi capaz de captar especificidades e produzir respostas acuradas. Porém, assim como os demais sistemas testados, este *chatbot* não permitiu a transferência para um atendente natural quando solicitado. Ao serem indagados sobre dificuldades e benefícios que a implantação de um *chatbot* governamental de atendimento ao público, os gestores deste sistema destacaram que:

As dificuldades para implantar um *chatbot* e manter seu pleno funcionamento depende muito de uma base dados atualizada e em perfeita harmonia, além de uma linguagem humanizada e correta, e pronta para responder em outros idiomas. Os benefícios são inúmeros, entre eles cito a qualidade que ganhamos no atendimento, com a disponibilidade e o atendimento em tempo real aos cidadãos (Secretaria do Planejamento, Orçamento e Gestão).

O trecho reforça uma dimensão conceitual e prática que está nas bases dos princípios evolutivos de sistemas automatizados: a primordialidade da organização, completude e

manutenção de base de dados capaz de treinar e atualizar o sistema. Isto deve ser considerado central para viabilização de projetos de *chatbots* governamentais que funcionem de fato.

Nas consultas a gestores, das 48 secretarias respondentes apenas 2 apontaram não ver necessidade ou benefício na implementação dos *chatbots*. Em linhas gerais, mesmo nos casos de secretarias que não possuem *chatbot*, há uma percepção ampla de que este tipo de comunicação pode melhorar o atendimento ao público e que a sua implantação está no horizonte próximo, como exemplificam os trechos a seguir:

A CTTI enxerga um grande potencial benefício no futuro em relação à implementação de um *chatbot*, uma vez que o principal objetivo da ferramenta é realizar a comunicação em tempo real usando inteligência artificial para atender à população sem a necessidade da presença de uma equipe de atendentes (Secretaria Municipal de Saúde de Cuiabá).

A SES-MG vê potenciais benefícios no uso da ferramenta. Existem versões em teste para a implementação da ferramenta para uso interno, inicialmente e posteriormente, há a expectativa de ampliação para o público externo (Secretaria de Estado de Saúde de Minas Gerais).

Isso configura um cenário geral de percepção positiva e tendência de uso de *chatbots* por gestores das prefeituras de capitais e governos estaduais brasileiros. Porém, em termos práticos, o baixo percentual de adoção desses agentes pelos entes governamentais estudados significa que ainda estamos em um estágio inicial de implementação destas ferramentas pelo Poder Executivo brasileiro subnacional.

Quanto aos desafios, deste conjunto de 48 respondentes, 18 não souberam identificar problemas ou barreiras na implementação de *chatbots* seja pelo fato de já possuírem ferramentas ativas e não identificarem problemas; ou, ao contrário, por não terem capacidade de informar quais seriam as barreiras na implementação por não possuírem projetos ou ferramentas ativas. Em relação às demais 30 secretarias que apontaram desafios e barreiras, podemos sintetizá-las nos seguintes termos:

a) *Limitações técnicas e operacionais*: para muitos gestores, a implementação de *chatbots* depende de uma infraestrutura tecnológica integrada a sistemas já existentes (como o SEI, sistemas de saúde, educação, etc.).

b) *Recursos humanos e capacitação técnica*: uma outra dimensão bastante apontada trata-se do déficit de profissionais especializados, como programadores e analistas de sistemas, o que inviabiliza o desenvolvimento interno e a manutenção contínua dos *chatbots*.

c) *Barreiras orçamentárias e burocráticas*: também foram relatadas dificuldades em garantir recursos financeiros e previsão orçamentária específica para aquisição ou desenvolvimento de *chatbots*, bem como a morosidade dos trâmites licitatórios, especialmente em contextos em que a tecnologia ainda não está prevista.

d) *Adoção e aceitação pelo usuário*: alguns gestores também relataram que haveria uma resistência do público quanto ao uso de soluções automatizadas, com preferência por atendimento humano. Além disso, avaliam que o desconhecimento e a ausência de campanhas de divulgação desses canais também comprometem a sua apropriação pelo público.

e) *Complexidade temática e fluxos informativos*: a amplitude e diversidade temática das secretarias também foram apontadas como barreiras pelos gestores. Neste sentido, apontam limitações técnicas na capacidade dos *chatbots* em responder a perguntas e direcionar corretamente atendimentos.

f) *Proteção de Dados e segurança da Informação*: as preocupações com a privacidade dos dados dos usuários e com o cumprimento da legislação vigente, como a Lei Geral de Proteção de Dados Pessoais (LGPD), emergem também como barreiras identificadas pelos gestores, sobretudo em setores sensíveis como a saúde e a segurança pública.

A resposta da Secretaria de Estado de Educação, Cultura e Esporte do Acre sintetiza bem essas preocupações e acrescenta a resistência do usuário ao atendente automatizado, o que demonstra que há também barreiras culturais a transpor:

Devido à extensa gama de serviços oferecidos pela SEE, pode ser desafiador para os *chatbots* fornecerem respostas personalizadas que atendam às necessidades individuais dos usuários. Além disso, a aceitação do uso desses serviços por parte dos usuários é uma preocupação, visto que a maioria prefere o atendimento humano. A manutenção contínua e as atualizações dos *chatbots* são cruciais para garantir sua eficácia e precisão. No entanto, isso demanda recursos consideráveis em termos de tempo e pessoal, e o Departamento de TI atualmente não possui os profissionais necessários em seu quadro (Secretaria de Estado de Educação, Cultura e Esporte do Acre).

No plano geral, é possível que as deficiências detectadas nos testes de acuracidade nos *chatbots* analisados neste estudo tenham relação direta com as barreiras de implementação levantadas por gestores. Trata-se de uma questão não apenas técnica, mas também administrativa como investimentos e cultura organizacional no registro e armazenamento de dados e a sua capacidade de externar isso na forma de comunicação com o público.

Considerações Finais

Este artigo teve como objetivo mapear e avaliar como sistemas automatizados de conversação (*chatbots*) estão sendo utilizados pelos governos dos estados, Distrito Federal e capitais brasileiras. Apontamos que a importância deste tipo de análise se dá em um contexto de franca expansão desses sistemas de intermediação. A pesquisa demonstrou que, embora haja um percentual ainda incipiente de iniciativas implantadas, há uma expectativa generalizada dos gestores sobre os potenciais benefícios deste tipo de sistema (mesmo para aqueles que não possuem *chatbot* ativo). Essa conjunção pode significar uma tendência de crescimento no número de iniciativas nos próximos anos, apesar das diversas barreiras elencadas por gestores.

Dentre aqueles *chatbots* mapeados, o estudo notou que quase a metade desses sistemas analisados estão em secretarias de Fazenda, possivelmente em razão das características deste tipo de órgão que lida com recursos financeiros e está mais aberto ao uso de ferramentas inovadoras para otimizar rotinas. Na análise qualitativa geral, foram aplicados cinco Testes de Conversação que buscaram medir a precisão das respostas recebidas. O resultado demonstrou que apenas 24% dos testes retornaram resultado esperado (Específico) dando resposta condizente e adequada à pergunta aplicada, ou seja, prepondera de modo generalizado o funcionamento não adequado dessas ferramentas que demonstraram ser incapazes de responder a uma primeira camada de testes básicos de modo acurado.

O baixo desempenho geral dos *chatbots* na produção de informação acurada tende a afetar o direito à informação, especialmente quando esses canais de comunicação se tornam peças importantes do atendimento ao público. O fato dos *chatbots* não permitirem a passagem para um atendente natural, quando solicitado, demonstra que há um problema comunicacional que pode impactar diretamente no direito do indivíduo em obter a informação ou serviço público nesses canais.

A busca de um funcionamento adequado passa pela construção de fluxos de informação completos ou pela existência de bases de dados adequadas. Por isso, *chatbots* não devem ser vistos apenas como ferramentas que atuam nas relações externas aos órgãos governamentais e sim com a ponta de um sistema de informação que pressupõem uma reengenharia e organização dos fluxos de dados internos.

Por fim, acredita-se que este estudo traz elementos que contribuem para se compreender as fragilidades hoje existentes nestas ferramentas. Pesquisas futuras precisam ser

desenvolvidas com camadas de testes mais complexos que observem dimensões e nuances da qualidade informativa e comunicativa desses sistemas. E isso pode caminhar em duas direções complementares: primeiro, ampliando o *corpus em análise* (incluindo mais pastas e mais instâncias governamentais, incluindo outros Poderes) e, segundo, aprofundando mais camadas de testes de conversação capazes de trazer outras dimensões analíticas como a capacidade dos *chatbots* de processar diferentes tipos de linguagens de diferentes grupos sociais; capacidade de processar diferentes graus de complexidade comunicativa e versatilidade para lidar com níveis de indeterminação em processos de conversação. Sobretudo, é preciso ter em mente que estes sistemas são novas formas de intermediação que podem alterar a atual estrutura dos fluxos de comunicação, informação e prestação de serviços entre Estado e cidadão e tendem a se posicionar como uma peça-chave na comunicação governamental contemporânea. Por isso, métodos de avaliação, incluindo pesquisas longitudinais, precisam ser desenhadas e aplicadas.

Referências

BARBER, Benjamin R. **Strong Democracy**: participatory politics for a new age. Londres e Los Angeles: University of California Press, 1984.

BATISTA, Gabriely Oliveira da Silva; MONTEIRO, Mateus de Souza; SALGADO, Luciana Cardoso de Castro. Investigando chatbots governamentais: um panorama sobre a usabilidade dentro e fora do Brasil. *In: Workshop de Computação Aplicada em Governo Eletrônico (WCGE)*. SBC, 2022. p. 73-84.

CASAS, Jacky *et al.* Trends & Methods in Chatbot Evaluation. *In: Anais ICMI '20 Companion*, October 25–29, Virtual Event, Netherlands, p. 280-286, 2020.

CASTOR, Emiliano Carlos Serpa *et al.* Chatbot: impactos no ambiente acadêmico de uma universidade do Rio de Janeiro. **P2P & Inovação**, v. 8, n. 1, p. 71-92, set. 2021/fev. 2022.

CHESTERMAN, S. Artificial Intelligence and The Limits of Legal Personality. **International & Comparative Law Quarterly**, v. 69, n. 4, p. 819-844, 2020.

DEIANA, G. et al. Artificial Intelligence and Public Health: Evaluating ChatGPT Responses to Vaccination Myths and Misconceptions. **Vaccines**, v. 11, n. 7, p. 1217, 2023.

FØLSTAD, A.; BJERKREIM-HANSSEN, N. User Interactions With a Municipality Chatbot—Lessons Learnt From Dialogue Analysis. **International Journal of Human–Computer Interaction**, p. 1–14, 2023.

GOMES, Leonardo Geovani; OLIVEIRA, Hilda Carvalho de; CRUZ, Dhoi Almeida. Chatbots: aspectos de impacto para o desenvolvimento e avaliação. *In: Conferências Iadis Ibero-Americanas Ww/Internet E Computação Aplicada*, 2021. **Anais [...]**. Rio Claro: IADIS, 2021. p. 22–30.

GUNKEL, David J. **An Introduction to Communication and Artificial Intelligence**. Cambridge: Polity Press, 2020.

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, 2020.

JAIN, Mohit. Evaluating and Informing the Design of Chatbots. *In: Anais do DIS '18: Proceedings of the 2018 Designing Interactive Systems Conference*, jun. 2018. p. 895-906, 2018.

HSU, Tiffany; THOMPSON, Stuart A. Disinformation Researchers Raise Alarms About A.I. Chatbots. **New York Times**, fev. 2023, online. Disponível em: <https://www.nytimes.com/2023/02/08/technology/ai-chatbots-disinformation.html>. Acesso em: 15 maio 2024.

KULIGOWSKA, Karolina. Commercial chatbot: performance evaluation, usability metrics and quality standards of embodied conversational agents. **Professionals Center for Business Research**, v. 2, n. 02, p. 1–16, 2015.

LANGEVIN, Raina *et al.* Heuristic evaluation of conversational agents. *In: CHI Conference On Human Factors in Computing Systems*, 2021, Online. Proceedings... New York: **ACM**, 2021. Article no. 632, p. 1–15.

LOMMATZSCH, A. A Next Generation Chatbot-Framework for the Public Administration. *In: Hodoň, M., Eichler, G., Erfurth, C., Fahrnberger, G. (Org.) Innovations for Community Services*. I4CS 2018. Communications in Computer and Information Science, vol 863. Springer, 2018.

LUCCHESI, Ivana Lima *et al.* Avaliação de um chatbot no contexto educacional: um relato de experiência com METIS. **Renote**, v. 16, n. 1, p. 113-122, 2018.

MAKASI, T. *et al.* A Typology of Chatbots in Public Service Delivery. **IEEE Software**, v. 39, n. 3, p. 58-66, 2022.

NIELSEN, Jakob; MOLICH, Rolf. Heuristic evaluation of user interfaces. *In: SIGCHI Conference on Human Factors in Computing Systems*, 1990, Seattle. **Proceedings...** New York: ACM, 1990. p. 249–256.

KHAN, Rashid; DAS, Anik. **Build Better Chatbots**. Berkeley: Apress, 2018.

MACPHERSON, C. B. The Life and Times of Liberal Democracy. Oxford: **Oxford University Press**, 1977.

MANGIPUDI, Phanidra. AI powered chatbots vs human agents: a compatible study on customer satisfaction. **SSRN Electronic Journal**, 2025.

MCDANIEL, E.; GONG, G. The language of robotics: Use and abuse of personification. In **IEEE Transactions on Professional Communication**, v. PC-25, n. 4, p. 178-181, 1982.

MEHR, Hila. Artificial Intelligence for Citizen Services and Government. Cambridge: **Ash Center for Democratic Governance and Innovation**, 2017.

MUDOFI, Luthfi Nurul Hidayah; YUSPIN, Wardah. Evaluating quality of chatbots and intelligent conversational agents of bca (VIRA) line. **Interdisciplinary Social Studies**, v. 1, n. 5, p. 532-542, 2022.

PATEMAN, Carole. **Participação e teoria democrática**. Rio de Janeiro: Paz e Terra, 1992.

PERAS, Dijana. Chatbot evaluation metrics. **Economic and Social Development: Book of Proceedings**, Varazdin, 14-15 dez. 2018.

PETTERS, Lorreine Beatrice. Chatbots em campanhas de sensibilização, narrativa conversacional e possibilidades interativas: o caso do bot Fabi para Unicef Brasil e Facebook. Comunicação, **Mídia e Consumo**, v. 16, n. 46, p. 252-270, 2019.

RADZIWILL, Nicole; BENTON, Morgan. Evaluating Quality of Chatbots and Intelligent Conversational Agents. **Software Quality Professional**, v. 19, n. 3, p. 25-36, 2017.

REINA, Darío Reyes; CRUZ, Clarice. Curadoria de chatbots: conceptualização, estratégias e indicadores de desempenho. **Revista Ibérica de Sistemas y Tecnologías de Información**, n. 40, p. 1-14, 2020.

SANTOS, Gabriel; SILVA, Williamson; VALLE, Pedro Henrique Dias. O que sabemos sobre testes em chatbots? Uma revisão sistemática da literatura. In: Workshop sobre Aspectos Sociais, Humanos e Econômicos de Software (WASHES), 9, 2024, Brasília/DF. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2024. p. 106-117.

SANTOS, Luiza; POLIVANOV, Beatriz. O que têm em comum Alexa, Siri, Lu e Bia? Assistentes digitais, sexismo e rupturas de performances de gênero. **Galáxia** (São Paulo), 2021.

SHIRAISHI, Guilherme de Farias *et al.* Para o high-tech ser high touch – um estudo exploratório com chatbots. **Administração em Diálogo**, São Paulo, v. 22, n. 1, p. 19, 2020.

SIMONDON, Gilbert. **Du mode d’existence des objets techniques**. Paris: Aubier, 1958.

TURNER, Jacob. **Robot Rules: Regulating Artificial Intelligence**. Londres: Palgrave Macmillan, 2019.

WACH, Krzysztof *et al.* The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. **Eber**, v. 11, n. 2, 2023.

Dados de Autoria

Sivaldo Pereira da Silva

E-mail: sivaldop@unb.br

ORCID: <https://orcid.org/0000-0001-8767-7679>

Instituição: Universidade de Brasília, Brasília, Distrito Federal, Brasil

Minibiografia: Doutor pelo Programa de Pós-Graduação em Comunicação e Cultura Contemporâneas da Universidade Federal da Bahia (Póscom-UFBA), com estágio doutoral na *University of Washington*, Estados Unidos. Professor da Faculdade de Comunicação (FAC) e do Programa de Pós-Graduação em Comunicação da Universidade de Brasília (PPGCOM-UnB). Atualmente é coordenador do grupo de pesquisa Centro de Estudos em Comunicação, Tecnologia e Política (CTPol); pesquisador do Instituto Nacional de Ciência e Tecnologia em Democracia Digital (INCT-DD).

Felipe Soares de Moraes

E-mail: felipe.moraes2018@gmail.com

ORCID: <https://orcid.org/0000-0002-6585-9073>

Instituição: Universidade de Brasília, Brasília, Distrito Federal, Brasil

Minibiografia: Mestrando do Programa de Pós-Graduação em Comunicação da Universidade de Brasília (PPGCOM-UnB). Graduado em Jornalismo e Humanidades pela Universidade Federal do Sul da Bahia (UFSB). Pesquisador do Centro de Estudos em Comunicação, Tecnologia e Política (CTPol/UnB).

Gabriella da Costa

E-mail: abrielladacosta@gmail.com

ORCID: <https://orcid.org/0000-0002-4234-4544>

Instituição: Universidade Federal Fluminense, Niterói, Rio de Janeiro, Brasil

Minibiografia: Doutoranda do Programa de Pós-Graduação em Comunicação da Universidade Federal Fluminense (PPGCOM-UFF). Mestre pelo Programa de Pós-Graduação em Comunicação e Cultura Contemporâneas da Universidade Federal da Bahia (Póscom-UFBA). Pesquisadora no Instituto de Ciência e Tecnologia em Democracia Digital (INCT-DD) e do Centro de Estudos em Comunicação, Tecnologia e Política (CTPol) da Universidade de Brasília (UnB).

Naiara Marques de Albuquerque

E-mail: naiaramalbuquerque@gmail.com

ORCID: <https://orcid.org/0009-0003-9266-0961>

Instituição: Universidade de Brasília, Brasília, Distrito Federal, Brasil

Minibiografia: Mestre em Comunicação Institucional e Política pela Universidade de Sevilla, graduada em Comunicação Social - Jornalismo pela Universidade de Brasília (UnB). Pesquisadora vinculada ao Centro de Estudos em Comunicação, Tecnologia e Política (CTPol) da Universidade de Brasília (UnB).

Amanda Xavier da Silva

E-mail: amanda.xaavier@gmail.com

ORCID: <https://orcid.org/0009-0006-8108-7110>

Instituição: Universidade de Brasília, Brasília, Distrito Federal, Brasil

Minibiografia: MBA em *Data Science e Analytics* pela Escola Superior de Agricultura Luiz de Queiroz da Universidade de São Paulo (USP/Esalq). Graduada em Relações Internacionais e Ciência Política pelo Centro Universitário do Distrito Federal (UDF). Pesquisadora vinculada ao Centro de Estudos em Comunicação, Tecnologia e Política (CTPol) da Universidade de Brasília (UnB).

Dados do artigo

Resultado de projeto de pesquisa, de dissertação, tese:

O artigo é parte da pesquisa intitulada “Direito à informação, transparência e responsividade no uso governamental de sistemas de Inteligência Artificial para interação com o público: parâmetros e indicadores qualitativos”.

Fontes de financiamento:

Pesquisa financiada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), através do edital Universal 2023. Código do processo: 421802/2023-5.

Apresentação anterior:

Uma versão preliminar deste artigo foi apresentada no Congresso do Instituto Nacional de Ciência e Tecnologia em Democracia Digital (INCT-DD), na cidade de Salvador, em novembro de 2023, sem publicação em anais.

Agradecimentos/Contribuições adicionais:

Pesquisa financiada com recursos do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Apenas para textos em coautoria

Concepção e desenho da pesquisa:

Sivaldo Pereira da Silva.

Coleta de dados:

Sivaldo Pereira da Silva, Gabriella da Costa, Naiara Marques de Albuquerque, Felipe Soares de Moraes e Amanda Xavier da Silva.

Análise e/ou interpretação dos dados:

Sivaldo Pereira da Silva, Gabriella da Costa e Naiara Marques de Albuquerque.

Escrita e redação do artigo:

Sivaldo Pereira da Silva, Gabriella da Costa, Naiara Marques de Albuquerque, Felipe Soares de Moraes e Amanda Xavier da Silva.

Revisão crítica do conteúdo intelectual:

Sivaldo Pereira da Silva, Gabriella da Costa e Naiara Marques de Albuquerque.

Formatação e adequação do texto ao template da E-Compós:

Sivaldo Pereira da Silva, Felipe Soares de Moraes e Amanda Xavier da Silva.

Dados sobre Cuidados Éticos e Integridade Científica**A pesquisa que resultou neste artigo teve financiamento?**

Sim.

Financiadores influenciaram em alguma etapa ou resultado da pesquisa?

Não.

Liste os financiadores da pesquisa:

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Autora, autor, autores têm algum tipo de vínculo ou proximidade com os financiadores da pesquisa?

Não.

Descreva o vínculo apontado na questão anterior:

Não se aplica.

Autora, autor, autores têm algum tipo de vínculo ou proximidade com alguma pessoa ou organização mencionada pelo artigo?

Não.

Descreva o vínculo apontado na questão anterior:

Não se aplica.

Autora, autor, autores têm algum vínculo ou proximidade com alguma pessoa ou organização que pode ser afetada direta ou indiretamente pelo artigo?

Não.

Descreva o vínculo apontado na questão anterior:

Não se aplica.

Interferências políticas ou econômicas produziram efeitos indesejados ou inesperados à pesquisa, alterando ou comprometendo os resultados do estudo?

Não.

Que interferências foram detectadas?

Não se aplica.

Mencione outros eventuais conflitos de interesse no desenvolvimento da pesquisa ou produção do artigo

Não se aplica.

A pesquisa que originou este artigo foi realizada com seres humanos?

Não.

Entrevistas, grupos focais, aplicação de questionários e experimentações envolvendo seres humanos tiveram o conhecimento e a concordância dos participantes da pesquisa?

Não se aplica, pois foram enviadas perguntas para órgãos públicos não direcionadas a seres humanos específicos. As respostas foram institucionais, não nominais.

Participantes da pesquisa assinaram Termo de Consentimento Livre e Esclarecido?

Não se aplica.

A pesquisa tramitou em Comitê de Ética em Pesquisa?

Não se aplica porque a pesquisa não envolveu a participação de seres humanos.

O Comitê de Ética em Pesquisa aprovou a coleta dos dados?

Não se aplica porque a pesquisa não envolveu a participação de seres humanos.

Mencione outros cuidados éticos adotados na realização da pesquisa e na produção do artigo:

Não se aplica.