

Um Agente Conversacional de Apoio à Linguagem Simples para Equipes de Desenvolvimento: Um Estudo com Mágico de Oz

Rodrigo Oliveira¹, Roberta Bordalo¹, Luciana Salgado¹, Claudia Cappelli²

¹Programa de Pós-Graduação em Computação (PPGC) da Universidade Federal Fluminense (UFF), Niterói, RJ, Brasil

²Departamento de Informática e Ciência da Computação da Universidade Estadual do Rio de Janeiro (UERJ), Rio de Janeiro, RJ, Brasil

rodrigoso@id.uff.br, rbordalo@id.uff.br, luciana@ic.uff.br,
claudia.cappelli@gmail.com

Abstract. *This article presents an evaluation of LSBot, a conversational agent designed to support development teams in applying Plain Language (PL) principles. The study used the Wizard of Oz technique to simulate the chatbot's operation in collaborative interaction sessions, allowing for analysis of its usefulness, ease of use, and perceived support. The results present empirical evidence demonstrating good acceptance of the tool and its potential to support reflection and new ideas about PL techniques for software teams.*

Resumo. *Este artigo apresenta a avaliação do LSBot, um agente conversacional concebido para apoiar equipes de desenvolvimento na aplicação de princípios de Linguagem Simples (LS). O estudo utilizou a técnica do Mágico de Oz para simular o funcionamento do chatbot em sessões de interação colaborativas, permitindo analisar sua utilidade, facilidade de uso e percepção de apoio. Os resultados apresentam evidências empíricas que demonstram boa aceitação da ferramenta e potencial de apoio à reflexão e novas ideias sobre a técnica de LS às equipes de software.*

1. Introdução

No Brasil, a recente aprovação de uma política nacional de Linguagem Simples (LS) [Gov.br 2025] consolida um marco institucional importante ao reconhecer a clareza comunicacional como um direito do cidadão e uma responsabilidade do Estado. Essa política reforça a necessidade de que informações públicas, cada vez mais mediadas por sistemas digitais, sejam compreensíveis, acessíveis e utilizáveis por todos. A LS da qual trata a política é, simultaneamente, uma técnica e um movimento internacional de cidadania que visa uma comunicação tão clara que seu público-alvo consiga encontrar facilmente o que precisa, compreender o que encontra e usar essa informação para atingir seus objetivos [IPLF *s.d.*]. Essa técnica se apresenta como um contraponto à comunicação rebuscada, obscura, burocrática e sem transparência que pode excluir grupos [Ranchordás 2022], aumentar os custos de um produto ou serviço [FDA *s.d.*], limitar a participação popular na tomada de decisões de interesse público e até mesmo contribuir para a desconfiança nas organizações [Schrivver 2017].

Nesse sentido, a LS tem sido considerada um requisito importante não somente em áreas como o governo digital, fomentando a participação ativa de todos os setores da sociedade [Nogueira et al. 2025], mas também na promoção e alfabetização em saúde, apoiando a entrega de informações compreensíveis para populações vulneráveis ou marginalizadas [Greene et al. 2017], no campo jurídico, permitindo que os cidadãos acessem melhor a legislação e compreendam as decisões judiciais [Arenas Arias 2023], entre outras áreas. O foco inicial do movimento de LS era a clareza e legibilidade do conteúdo textual. Entretanto, a área tem ampliado seu escopo com a aplicação em outros meios, como sistemas e interfaces [Nunes et al. 2023; Oliveira et al. 2026b]. Isso se deve principalmente à acelerada digitalização, em que a Computação enfrenta desafios como a exclusão, desigualdades e promoção do letramento digital [Wagner e de Medeiros Santos 2026 p.32], o que aproxima a LS dessas demandas e reforça seu potencial para reduzir barreiras de compreensão e uso em sistemas digitais, favorecendo o empoderamento e a participação em serviços mediados por tecnologia.

Do ponto de vista da colaboração, esse cenário cria oportunidades e desafios para equipes de desenvolvimento, uma vez que a atividade é inerentemente colaborativa [Diniz et al., 2025] e projetos orientados à LS demandam mudanças culturais e organizacionais, com maior integração entre equipes e adaptação de processos [Coutinho et al., 2025; Card e Bacigalupa 2025]. A baixa familiaridade com LS e sua percepção de valor, muitas vezes associada apenas ao usuário final, dificultam sua adoção [Johannessen et al. 2017; Coutinho et al. 2025]. Assim, a adoção da LS em sistemas configura um desafio coletivo, que envolve decisões compartilhadas sobre linguagem, organização da informação e adequação ao público, exigindo negociação, coordenação e construção conjunta de entendimento, aspectos centrais dos Sistemas Colaborativos.

Para explorar essa questão, esta pesquisa propôs um *chatbot* de suporte a equipes de desenvolvimento, concebido como artefato para apoiar a construção colaborativa de soluções orientadas à LS. Diante da complexidade envolvida, como a orientação de práticas e a tradução para LS, ainda carentes de alinhamento com diretrizes formais [Sastre et al. 2024; Lyu et al. 2023], optou-se por uma implementação simulada. O *chatbot* foi desenvolvido com a técnica do Mágico de Oz (*Wizard of Oz* - WOZ) [Menten et al. 2024], sendo operado por um humano sem o conhecimento dos participantes. O estudo permitiu observar, em ambiente controlado, a recepção da ferramenta, sua utilidade, facilidade de uso e a percepção de apoio às equipes na aplicação de LS. Inserido em uma pesquisa mais ampla sobre o projeto de sistemas orientados à LS [Oliveira et al. 2026a], este artigo investiga como agentes conversacionais podem atuar como tecnologias de apoio no desenvolvimento de software, especialmente na construção coletiva de decisões de design, ao aproximar a LS da área de Sistemas Colaborativos. Como contribuição, apresentamos evidências empíricas sobre o uso de artefatos conversacionais e uma aplicação do WOZ nesse contexto, relevante para pesquisas futuras em ambas as áreas.

A partir desta introdução, o artigo apresenta sua fundamentação teórica na Seção 2. Na Seção 3, é descrita a metodologia utilizada, juntamente com as atividades realizadas. Em seguida, na Seção 4, apresentamos os resultados, seguidos por uma breve discussão. Por fim, as conclusões são apresentadas na Seção 5.

2. Fundamentação Teórica

2.1. Linguagem Simples

Desde 2014, a Federação Internacional de Linguagem Simples (*International Plain Language Federation*) adotou uma definição de LS baseada em três princípios fundamentais: a facilidade com que o público-alvo pode encontrar, compreender e utilizar as informações determina se o conteúdo está em LS [IPLF *s.d.*]. Em sua forma atual, a LS ganhou impulso na década de 1970 por meio de iniciativas que buscavam ajudar os cidadãos a compreender seus direitos e deveres [Adler 2012; Schriver 2017]. Atualmente, é reconhecida como uma técnica que reduz a obscuridade das informações apresentadas por governos e organizações, que muitas vezes limita a compreensão e o acesso a informações essenciais.

No desenvolvimento de software, a LS é vista como uma aliada que apoia a comunicação entre as partes interessadas (*stakeholders*) [Parvathi et al. 2017]. Mais recentemente, ela também tem sido considerada relevante para a implementação dentro dos próprios sistemas [Oliveira et al. 2023; 2026b]. A era digital acelerou essa mudança, fazendo com que as organizações percebessem que a LS vai além da linguagem textual. Em contextos digitais, seu escopo inclui interfaces de usuário, *design* visual e recursos interativos [Nunes et al. 2023; Williams 2022].

No entanto, mesmo com esses avanços, os processos orientados à LS em sistemas ainda enfrentam desafios como baixa escalabilidade e automação limitada [Backowski et al. 2022], bem como a necessidade de abordar a diversidade cultural e linguística ao projetar interfaces inclusivas [Nunes et al. 2023]. Mesmo com os avanços em Inteligência Artificial (IA) e em Grandes Modelos de Linguagem (*Large Language Models* - LLMs), o uso dessas tecnologias para tradução de LS, embora promissor, tem sido apontado como de baixo alinhamento às diretrizes formais da técnica de LS [Sastre et al. 2024; Lyu et al. 2023].

2.2. Mágico de Oz

O Mágico de Oz (*Wizard of Oz* - WOz) é uma técnica bem estabelecida na pesquisa, utilizada para prototipagem e simulação do comportamento de sistemas que ainda não foram implementados [Joshi et al. 2024; Menten et al. 2024]. Nessa abordagem, os usuários interagem com o que acreditam ser um sistema real, neste caso um *chatbot* atuando como assistente de apoio, enquanto as respostas são, na verdade, fornecidas por um operador humano, denominado “mágico”, que controla o sistema nos bastidores. O WOz foi escolhido neste estudo especificamente porque as incertezas relacionadas à LS emergem de forma mais clara quando os desenvolvedores precisam tomar decisões de design em tempo real de projeto, revelando lacunas de conhecimento que podem não ser capturadas por entrevistas ou questionários aplicados em momentos posteriores.

Esta pesquisa se concentra principalmente na identificação de percepções (*insights*) que ajudem a refinar a técnica ou uso da tecnologia na qual o WOz se baseia [Hu et al. 2023]. Assim, o principal objetivo não é a tecnologia em si, mas a compreensão de como as pessoas interagem e se relacionam com ela. Nesse contexto, a simulação envolveu um assistente humano que emulou, de forma estruturada, o

comportamento de um sistema de IA. Essa abordagem apresenta desafios, como a carga cognitiva de gerenciar múltiplas tarefas e a necessidade de lidar com respostas abertas em tempo real [Hu et al. 2023]. Além do mágico, o estudo contou com um facilitador, ator responsável por conduzir a sessão e interagir diretamente com os participantes, cuja atuação é fundamental para a coerência da experiência. Como discutido por Porcheron et al. (2020), estudos WOz demandam um trabalho colaborativo e altamente coordenado entre os pesquisadores “em cena” e “nos bastidores”. Essa abordagem também inclui desafios relacionados à coordenação em tempo real e à manutenção da aparência de autonomia do sistema durante a interação. Para lidar com tais desafios, os autores destacam o uso de artefatos metodológicos compartilhados, como protocolo, roteiro e respostas pré-definidas, que orientaram a ação dos pesquisadores e viabilizam a condução consistente do estudo, cujas estratégias estão detalhadas na Seção 3.

2.3. Trabalhos Relacionados

Há um crescente avanço no uso de *chatbots* e assistentes de IA no apoio a tarefas de desenvolvimento de software, em sua maioria atuantes na gestão de projetos [Moguel-Sánchez et al. 2023]. Além disso, essas ferramentas também são utilizadas para fornecer informações sob demanda, auxiliando na tomada de decisões dos desenvolvedores [Süße et al. 2023], similar à proposta do LSBot neste trabalho. Sob a perspectiva colaborativa, tais tecnologias podem contribuir para a resolução de problemas em equipe, reduzir lacunas de comunicação e diminuir o atrito social [Johnson et al. 2025]. Em uma revisão de mais de 80 estudos, Moguel-Sánchez et al. (2023) apontam como principal desafio o desempenho ainda insatisfatório dos *chatbots* em tarefas específicas, levantando preocupações quanto à sua adaptabilidade, raciocínio e autonomia. Embora esses trabalhos evidenciem o potencial dos agentes conversacionais no desenvolvimento de software, eles se concentram majoritariamente em tarefas gerais ou técnicas, com pouca atenção ao suporte a decisões de design relacionadas à aplicação de princípios práticos como a LS. Nesse sentido, este trabalho se diferencia ao investigar o uso do agente conversacional como artefato de apoio à construção coletiva de soluções orientadas à LS, enfatizando seu papel na mediação de decisões colaborativas de design.

No que se refere ao uso da técnica do WOz nesse contexto, trata-se de uma abordagem de menor custo, frequentemente empregada para avaliar aspectos complexos de sistemas ainda não implementados [Hu et al. 2023]. A colaboração é um desses aspectos, sendo comum o uso da técnica para testar plataformas multiusuário em fases iniciais ou compreender dinâmicas colaborativas. No trabalho de Hu et al. (2023), investiga-se a colaboração entre múltiplos usuários no papel de mágico, analisando seus efeitos na experiência do usuário, na divisão de tarefas e na auto-organização. Já Simpson et al. (2022) exploram o desempenho e a colaboração de equipes no uso de interfaces conversacionais em ambientes virtuais de aprendizagem colaborativa. Por sua vez, Pereira et al. (2022) analisam o incentivo à interação entre estudantes em uma rede social educativa por meio de recomendações baseadas em WOz. Diferentemente dessas abordagens, este trabalho emprega a técnica para investigar como um agente conversacional pode apoiar diretamente as equipes de desenvolvimento, com foco na percepção de utilidade, facilidade de uso e apoio à colaboração.

3. Metodologia

Este estudo envolveu as interações com um protótipo simulado de *chatbot* projetado para fornecer suporte à LS. Desta forma, as percepções, desenvolvidas através da coleta de dados, fornecem a base para este trabalho.

A pesquisa contou com um protocolo, que descrevia a estrutura geral do estudo, contendo um roteiro da sessão, dividido nas etapas: introdução, preparação para a interação, interação com o *chatbot* de apoio à Linguagem Simples (LSBot), intervenção para fluidez da sessão (recomendada, caso necessário) e resposta a um questionário final da avaliação do *chatbot*. Adicionalmente, foi elaborado um *slide* de apoio com o objetivo de padronizar as ações e fomentar o diálogo entre o facilitador e os desenvolvedores. Esses e outros dados relacionados à pesquisa estão disponíveis em um repositório público em: <https://zenodo.org/records/18954730>.

Para validação do protocolo e do *slide* de apoio, foi realizada uma sessão piloto, com a participação de três pesquisadores, a saber: o mágico, o facilitador e um voluntário desempenhando o papel de desenvolvedor. A partir dessa sessão, foi possível refinar tanto o protocolo quanto o material de apoio, bem como aprimorar a dinâmica de interação entre o *chatbot* e os participantes. Como exemplo dessas melhorias, destaca-se a criação de *stickers*¹, como forma de *feedback* aos participantes em tempo de interação para as ações do *chatbot*, como apresentado na Figura 1.

3.1. Contexto e participantes

O estudo envolveu oito equipes de estudantes de graduação dos cursos de Ciência da Computação e Sistemas de Informação. Cada grupo aplicou a LS em um projeto da disciplina de Interação Humano-Computador (IHC), abordando diferentes temas relacionados à saúde e ao bem-estar, abrangendo aplicações para o acesso à informação de saúde, gerenciamento de medicamentos e agendamento de cuidados médicos (G1–G8). As equipes foram selecionadas por amostragem de conveniência com base na disponibilidade durante o período da disciplina. Embora não seja probabilística, esta seleção reflete desenvolvedores iniciantes típicos em projetos de software.

A participação foi voluntária, desvinculada de qualquer avaliação acadêmica da disciplina e executada com autorização do docente responsável. Os participantes poderiam desistir a qualquer momento, sem qualquer prejuízo. Todos os estudantes foram informados sobre os objetivos e procedimentos do estudo, assinaram o Termo de Consentimento Livre e Esclarecido e autorizaram as gravações para fins de coleta de dados. Não foram utilizados dados pessoais e tanto os participantes quanto as equipes foram identificados apenas por códigos numéricos.

Antes da pesquisa em si, realizamos atividades com as equipes para introduzir o tema LS, considerando que os participantes não tinham conhecimento prévio sobre o assunto. Os alunos eram iniciantes em IHC, mas possuíam conhecimento em Engenharia de Software, incluindo levantamento de requisitos. As atividades incluíram: exemplos de simplificação de interface, exercícios de reescrita de microtextos

¹ *Stickers* são adesivos digitais usados para personalização ou expressão em conversas virtuais. No WhatsApp, estes adesivos são conhecidos como figurinhas.

relacionados à saúde e conversas dos grupos para identificar dúvidas sobre LS cabíveis ao projeto em desenvolvimento. No total, participaram oito equipes, cada uma com cinco integrantes. No entanto, a interação direta com o *chatbot* envolveu 15 desenvolvedores no total, com cada equipe contando de 1 a 4 membros por sessão. A ausência de alguns membros se deu devido a conflitos de agenda. Ainda assim, ao menos um desenvolvedor de cada grupo atuou como representante dos demais, apresentando dúvidas dos outros participantes coletadas nas atividades prévias.

3.2. Procedimento e Estrutura do WOz

O estudo envolveu três papéis: desenvolvedores (usuários do *chatbot*), um facilitador (pesquisador de IHC responsável pela mediação) e um mágico (pesquisador experiente em LS que respondia secretamente como o *chatbot*). Cada equipe participou de uma sessão *online* com duração aproximada de 50 minutos. As sessões aconteceram em dois dias, conforme a disponibilidade dos participantes, contando com o mesmo facilitador e mágico. Durante as sessões, um ou mais desenvolvedores de cada equipe puderam fazer perguntas específicas sobre como aplicar a LS em seus próprios projetos, seja em textos de interface, fluxos de navegação ou protótipos de telas, como mostra a Figura 1. A comunicação ocorreu por meio de um aplicativo de mensagens (Figura 1A) em um grupo que incluía o facilitador (Figura 1B), os desenvolvedores (Figura 1C) e o *chatbot* (Figura 1D). Na área de interação (Figura 1D), temos um exemplo de mensagem enviada pelo mágico ao desenvolvedor. Todas as sessões seguiram a mesma estrutura e conjunto de instruções para padronizar as interações entre as equipes.

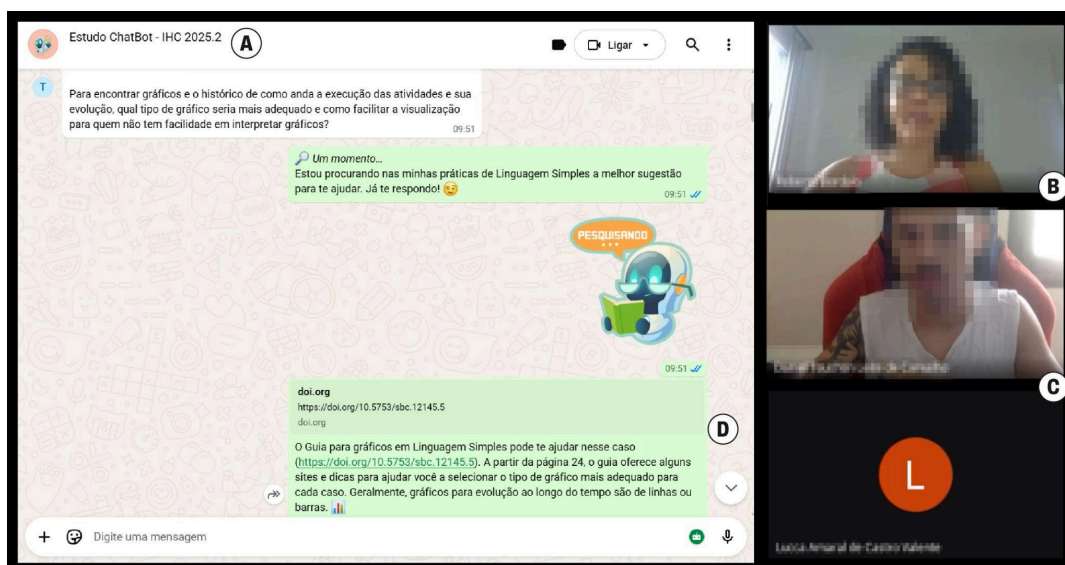


Figura 1 - Exemplo de sessão de interação *online* com o *chatbot*.

A Figura 2 apresenta um esquema da estrutura do estudo WOz. Durante as sessões *online*, os desenvolvedores interagem diretamente com o *chatbot* (Figura 2A), enquanto um pesquisador atuava como facilitador, gerenciando a sessão e controlando o tempo (Figura 2B). O *chatbot* respondia em tempo real às perguntas dos desenvolvedores (Figura 2C), sem que eles soubessem que as respostas eram geradas por um assistente humano remoto. Esse assistente, denominado “mágico”, gerou respostas coerentes e contextualizadas a partir de uma base de respostas pré-definida já

baseada em LS (Figura 2E). Quando necessário, o assistente utilizava *prompts* predefinidos com uma LLM, como ChatGPT², para gerar sugestões adicionais (Figura 2F-G). Em seguida, o assistente poderia então refinar ou adaptar a resposta, por exemplo, adicionando um *link* de referência, antes de enviá-la de volta por meio do *chatbot* (Figura 2D). Ao longo das sessões, os participantes foram incentivados a utilizar a técnica de “pensar em voz alta” (*think-aloud*), verbalizando seus pensamentos à medida que interagem. Para minimizar o desconforto potencial, os participantes foram lembrados de que todas as perguntas eram igualmente válidas e que poderia ocorrer algum atraso nas respostas do *chatbot* em razão da ferramenta ser um protótipo.

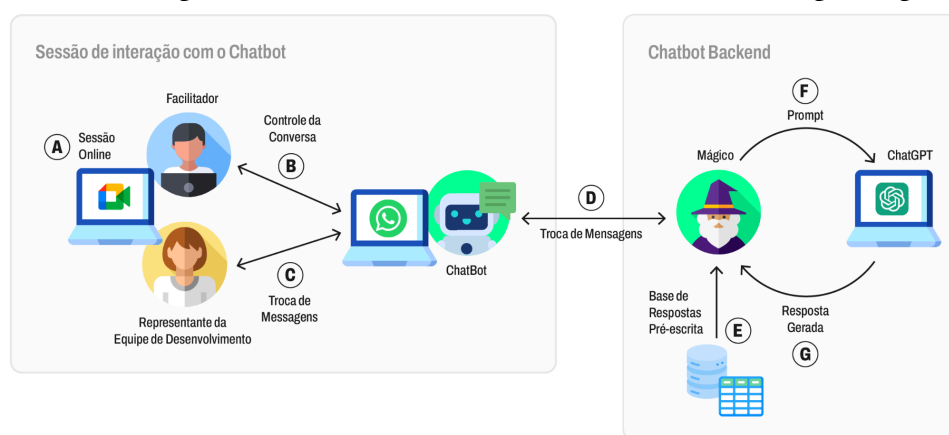


Figura 2 - Estrutura do Estudo do Mágico de Oz. Sessão de interação online com o chatbot (esquerda) e Backend do chatbot (à direita) com a presença do mágico.

3.3. Coleta e análise de dados

A coleta de dados se concentrou na percepção das equipes quanto ao uso do LSBot na atividade proposta. Embora este trabalho seja parte de uma pesquisa mais ampla sobre o projeto de sistemas com LS [Oliveira et al. 2026a], neste artigo o recorte analítico será exclusivamente sobre dados quantitativos obtidos por meio de dois questionários. O primeiro foi direcionado aos projetistas e estruturado com base no *Chatbot Usability Questionnaire* (CUQ) [Holmes et al. 2019], adaptado para contemplar aspectos específicos do contexto do estudo e da temática de LS, conforme prática recorrente em avaliações de *chatbots* [Carvalho et al. 2021]. O questionário incluiu itens relacionados à clareza e utilidade das respostas, facilidade de uso, percepção de apoio ao trabalho da equipe e dinâmica da atividade, organizados em escala Likert de cinco pontos.

O segundo questionário buscou captar a percepção do facilitador quanto à colaboração das equipes durante a atividade, considerando aspectos como compreensão das tarefas, envolvimento na atividade, facilidade na formulação de perguntas ao *chatbot*, iniciativa na geração de novas ideias a partir das respostas recebidas, demanda por apoio adicional e segurança em relação ao tema de LS. Os dados foram analisados por meio de estatística descritiva, considerando médias e distribuição das respostas, visando identificar tendências na percepção de usabilidade, utilidade e apoio oferecido pelo LSBot no contexto colaborativo. Os resultados são apresentados na seção seguinte.

² <https://chatgpt.com/>

4. Resultados

4.1. Avaliação do LSBot pelos Projetistas

A Figura 3 apresenta a distribuição das respostas dos participantes aos itens dos questionários. O gráfico resume as percepções quanto à clareza e utilidade das respostas do LSBot, facilidade de uso, percepção de apoio e avaliação da dinâmica da atividade. A maioria das afirmações foi formulada em sentido positivo (sinalizadas pelo ícone verde ao lado das perguntas), de modo que se espera, nesses casos, maior concentração de respostas em concordância. Por outro lado, os itens sinalizados com o ícone vermelho correspondem a afirmações formuladas em sentido negativo (por exemplo, relacionadas à demora nas respostas ou a possíveis confusões ao usar o *chatbot*). Nesses casos, espera-se mais respostas em discordância, indicando uma avaliação mais favorável.

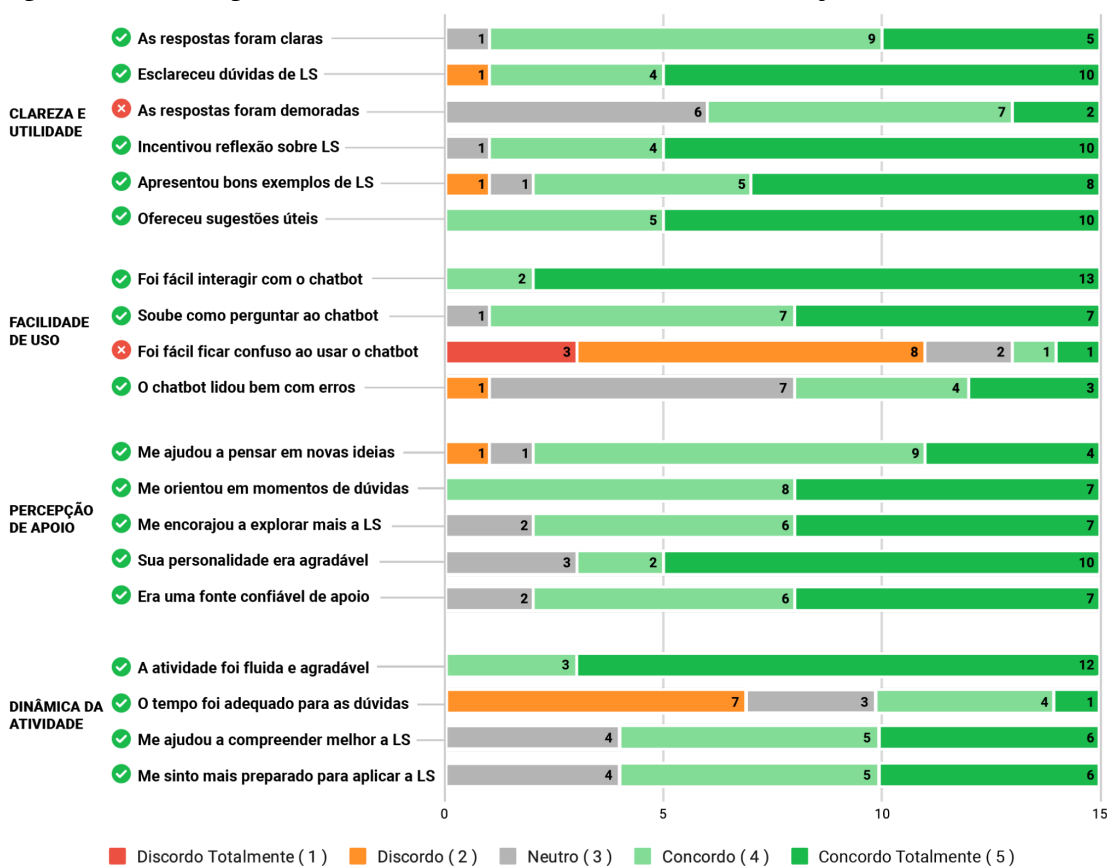


Figura 3 - Gráfico de Avaliação do LSBot pelos participantes.

Na avaliação da clareza e utilidade das respostas, conforme ilustrado na Figura 3, a maioria dos participantes indicou Concordo ou Concordo Totalmente, evidenciando que o conteúdo apresentado pelo LSBot foi claro, direto e útil. Da mesma forma, as respostas da ferramenta foram consideradas eficazes para esclarecer dúvidas sobre LS, sugerindo que o LSBot cumpre adequadamente sua capacidade informativa. Com relação ao tempo para respostas do LSBot, 9 dos participantes concordam que estas foram muito demoradas, enquanto os demais mantiveram-se neutros em relação à afirmativa. Esse resultado deve ser interpretado à luz do desenho do estudo: nos slides introdutórios da dinâmica, foi explicitado que o LSBot se tratava de um protótipo, sendo

esperado algum tempo de processamento nas respostas. Além disso, por se tratar de uma simulação via WOz, havia uma demora intrínseca associada à atuação do “mágico”, ainda que esta tenha sido minimizada por meio do uso de um banco estruturado de respostas e de recursos como stickers que sinalizavam que o sistema estava “processando” a solicitação. Mesmo sem perceberem que se tratava de um estudo WOz, alguns participantes pontuaram a demora como aspecto negativo, o que sugere que, ainda que contextualizada como característica de protótipo, a latência foi suficientemente perceptível para influenciar a experiência de uso. Por fim, quanto aos exemplos apresentados pelo LSBot, em apenas um dos casos o participante discordou de que eles ajudariam a compreender as práticas de LS.

Quanto à facilidade de uso, os resultados indicam que os desenvolvedores classificaram a interação com o LSBot como intuitiva e fluida, conforme evidenciado na Figura 3, com a maioria concordando nos itens de facilidade para interagir e formular perguntas para o *chatbot*. Esses dados indicam que a ferramenta apresenta baixa barreira de uso, uma vez que dificuldades operacionais não foram percebidas de forma recorrente durante a interação. Já no item sobre estar confuso ao usar a ferramenta, há maior discordância, dado que se trata de uma afirmação em sentido negativo (sinalizada com o ícone vermelho); nesse caso, a discordância deve ser interpretada como um indicativo favorável à ferramenta.

No item relacionado à capacidade do LSBot de lidar adequadamente com possíveis erros ou inconsistências, observa-se maior incidência de respostas neutras. Esse resultado pode estar associado ao perfil iniciante dos participantes, que, por ainda estarem em processo de aprendizagem sobre LS, podem ter assumido as respostas do *chatbot* como corretas, sem elementos suficientes para avaliá-las criticamente. Ainda assim, não foram registrados indícios de erro percebido ou situações que comprometessem a credibilidade da simulação, o que sugere que a condução do WOz manteve de forma consistente a aparência de autonomia do sistema.

No que se refere à percepção de apoio, essa dimensão recebeu uma avaliação amplamente favorável, visto que o LSBot foi considerado capaz de auxiliar no aprendizado, esclarecer dúvidas e apoiar atividades relacionadas à LS. Os participantes acreditam que o LSBot contribuiu para a compreensão da LS, auxiliando tanto na compreensão de conceitos quanto na aplicação no contexto do projeto de seus grupos. Apenas uma pequena parcela indicou discordância ou neutralidade, o que sugere que a percepção de apoio pode variar conforme o perfil e a experiência prévia do participante.

Por fim, com relação à dinâmica da atividade, todos os participantes a compreenderam como fluida e agradável. De forma igualitária, os participantes entenderam que a atividade os ajudou a compreender e os preparou para aplicar os conceitos de LS em projetos futuros. Sete participantes discordaram de que o tempo disponível foi adequado para suprir as principais dúvidas, aspecto apontado como ponto de melhoria nos relatos abertos de alguns participantes que serão apresentados a seguir.

Os projetistas avaliaram o *chatbot* de forma predominantemente positiva a partir dos relatos abertos ao final do questionário, destacando o uso de linguagem acessível e recursos visuais (*stickers*), que “ajudaram a manter um tom leve e aumentaram a vontade de interagir” (P1). Os participantes relataram que a interação com o *chatbot* contribuiu para a compreensão de aspectos do projeto antes não considerados,

afirmando que “a linguagem usada é acessível, e me ajudou a entender pontos do meu projeto que eu não considerava antes” (P1). As respostas foram percebidas como úteis, especialmente quando estimulavam a continuidade da interação, como nos casos em que o *chatbot* “respondia com uma pergunta final se eu precisava de mais ajuda em alguma outra etapa da LS” (P4). Como principais pontos de melhoria, além da latência para as respostas, houve menção ao desejo por exemplos mais proativos e variados (P13). Ainda assim, a experiência foi considerada positiva, mesmo diante das limitações de tempo, especialmente quando associada ao apoio oferecido na atividade.

4.2. Avaliação do LSBot pelo Facilitador

O estudo contou também com um questionário a ser respondido pelo facilitador, que foi estruturado a partir de seis perguntas sobre: compreensão da tarefa (Q1), interesse e envolvimento na atividade (Q2), iniciativa em explorar novas ideias ou incorporar as soluções propostas pelo *chatbot* ao projeto (Q3), necessidade de apoio do facilitador (Q4), facilidade de formulação de perguntas ao LSBot (Q5) e familiaridade ou segurança em relação à LS (Q6). Cada dimensão buscou capturar aspectos comportamentais e processuais da dinâmica colaborativa durante a interação com o LSBot.

O facilitador respondeu ao questionário para cada um dos oito grupos participantes ao final de cada sessão de interação com o *chatbot*. Como as escalas das perguntas possuíam números distintos de categorias, as respostas foram convertidas em valores numéricos ordinais e normalizadas em um intervalo comum entre 0 e 1, utilizando a fórmula $(\text{valor} - \text{mínimo}) / (\text{máximo} - \text{mínimo})$. Essa padronização possibilitou a construção de um mapa de calor comparativo apresentado na Figura 4. Na visualização, tons em verde indicam avaliações mais favoráveis, enquanto tons mais próximos do vermelho representam menor desempenho na dimensão analisada.

No que se refere à compreensão para iniciar a interação com *chatbot*, 2 dos 8 grupos apresentaram dificuldades iniciais para compreender e iniciar a interação com o LSBot. Os demais (6 de 8 grupos), após uma orientação inicial do facilitador, conduziram a atividade com segurança, demonstrando familiaridade conceitual e confiança durante a interação. Com relação ao interesse e envolvimento na atividade, 6 dos grupos apresentaram desempenho alto, trabalhando com entusiasmo, enquanto um grupo demonstrou engajamento moderado (G7) e outro dificuldade para manter o foco durante a atividade (G4). Quanto aos sinais de exploração de novas ideias ou de incorporação das soluções ao trabalho, o facilitador avaliou que isso não ocorreu em apenas um grupo (G4), possivelmente em função da ausência de apoio dos demais integrantes e/ou da não realização da atividade prévia, que permitiu a elaboração de dúvidas sobre LS entre os grupos. Nos demais 7 grupos, o facilitador identificou indícios de reflexão, adaptação ou aplicação das sugestões discutidas com o LSBot de forma pontual em 2 grupos e de forma mais constante e autônoma nos outros 5 grupos.

Quanto às demandas de apoio para facilitar a interação com o LSBot e a execução da tarefa, em 6 dos casos, o suporte foi mínimo ou inexistente. Em um grupo (G7), o apoio foi moderado e, em outro (G4), foi necessária condução ativa por parte do facilitador, com sugestão de perguntas e explicações adicionais sobre a atividade. Em 6 dos grupos, o facilitador avaliou que os participantes trouxeram perguntas previamente

elaboradas ou souberam formulá-las com facilidade durante a atividade. Em um grupo (G8), houve dúvidas pontuais que foram superadas com pouco apoio, enquanto em outro observou-se maior esforço e hesitação na formulação das perguntas (G4). Quanto ao nível de familiaridade ou segurança dos participantes em relação à LS, em 2 grupos, observou-se insegurança, evidenciada por dificuldades em compreender ou aplicar o tema (G4 e G7). Em outros 2 grupos (G6 e G8), os participantes demonstraram compreensão geral, com dúvidas pontuais, indicando segurança relativa. Por fim, em 4 dos grupos, os participantes pareciam bem à vontade com o tema.

As observações qualitativas do facilitador, registradas em relatos abertos, indicaram comportamentos relevantes, como: engajamento elevado mesmo em atividades individuais; dificuldades pontuais em “pensar em voz alta”; subutilização do tempo disponível por um participante (G4), mesmo após alertas sobre o encerramento da sessão; e discussões espontâneas sobre LS entre participantes no momento de resposta ao questionário final, de aproximadamente 10 minutos, quando alguns projetistas discutiram ideias e melhorias nos projetos e suas percepções sobre a dinâmica da atividade.

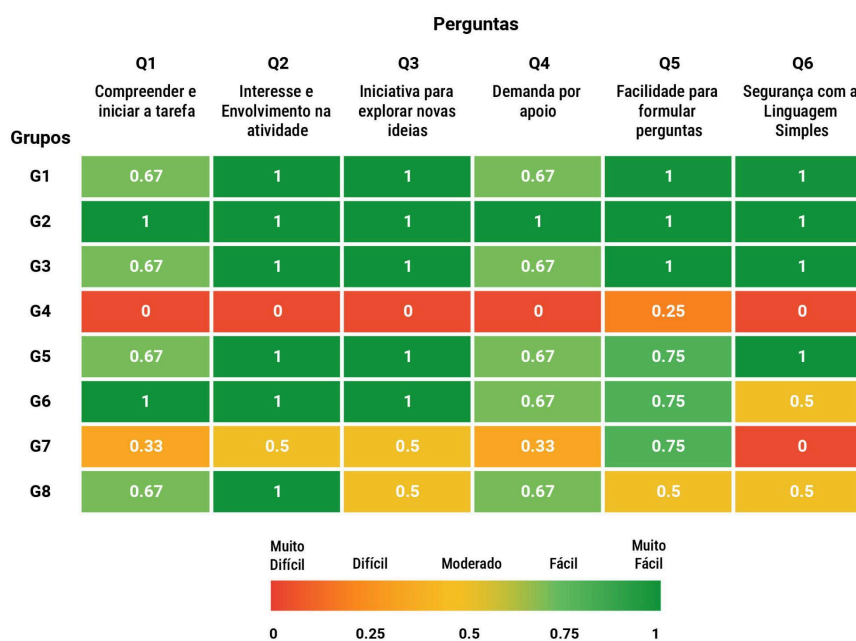


Figura 4 - Avaliação dos Grupos pelo Facilitador.

5. Conclusão

Esta pesquisa propôs um *chatbot* de suporte para equipes de desenvolvimento de sistemas (LSBot), criado por meio de uma abordagem inédita para o estudo de projetos orientados à Linguagem Simples (LS), utilizando a técnica do Mágico de Oz (*Wizard of Oz* - WOz). Ao articular LS e Sistemas Colaborativos, o estudo investigou o papel de agentes conversacionais como artefatos mediadores de reflexão, novas ideias e apoio nos processos de design nas etapas ainda conceituais do projeto de software e pode contribuir para o avanço de pesquisas sobre novas ferramentas automatizadas de LS, como assistentes baseados em LLM.

No que se refere aos desenvolvedores, estes perceberam o LSBot como uma ferramenta de apoio ao aprendizado, auxiliando na compreensão dos princípios da LS. Embora a avaliação geral seja positiva, percepções minoritárias apontam para a necessidade de apoio adicional, especialmente por meio de exemplos mais contextualizados ou maior detalhamento de determinadas respostas. Os achados relacionados à facilidade de uso indicaram que o LSBot apresentou boa aceitação e usabilidade, mesmo em contextos nos quais os participantes contavam com menor participação do grupo. Nos grupos com maior atuação da equipe, a colaboração percebida e o estímulo a novas ideias sugerem que o uso da ferramenta tem potencial para estimular discussões e reflexões conjuntas, reforçando seu papel de apoio em contextos educacionais e de capacitação no tema.

No entanto, o estudo aponta também caminhos para melhorias, especialmente no sentido de ampliar a capacidade de adaptação ao contexto e necessidades específicas do desenvolvedor e seu projeto. Por exemplo, nos casos com menos participação ou interação do grupo, por exemplo, no Grupo 4, observou-se mais dificuldade em formular perguntas, menor exploração das respostas do *chatbot* e menor aproveitamento da atividade ao projeto. Isso indica que a ferramenta pode evoluir de um modelo apenas reativo, baseado em perguntas feitas pelos participantes, para um agente proativo, capaz de sugerir e estimular a participação e se adaptar melhor ao contexto da equipe. Além disso, a percepção de demora nas respostas, mesmo com a explicação prévia de que se tratava de um protótipo, mostra que o tempo de retorno influencia a experiência e pode afetar o ritmo da atividade colaborativa. Mesmo com as estratégias propostas no estudo, como o uso de base de respostas pré-definidas para mitigar esses efeitos, esse é um fator relevante para a atenção em pesquisas futuras.

Como contribuições para a área de Sistemas Colaborativos, este estudo apresenta evidências empíricas de que agentes conversacionais podem apoiar a tomada de decisão em equipe, estimular a exploração de ideias e mediar discussões sobre clareza comunicacional no desenvolvimento de sistemas interativos. O trabalho também demonstra o uso do WOz como estratégia metodológica para investigar dinâmicas colaborativas mediadas por IA em fases iniciais de concepção e contribui também para a avaliação de artefatos conversacionais voltados ao suporte de atividades colaborativas no desenvolvimento de software.

Em relação às limitações e ameaças à validade do estudo, a investigação foi conduzida com participantes de uma única instituição de ensino, todos do nível iniciante. O tamanho reduzido da amostra e a utilização de um perfil único de desenvolvedores podem limitar a generalização dos resultados. Uma possível ameaça diz respeito à participação parcial de algumas equipes durante as sessões, mitigada pela garantia de pelo menos um representante de cada grupo e por uma atividade prévia em sala de aula na qual as equipes reuniram suas perguntas coletivamente. Outra ameaça relaciona-se a possíveis variações nas orientações fornecidas pelo *chatbot* e seu mágico. Isso foi mitigado pelo uso de respostas predefinidas de LS, derivadas da mesma atividade em sala de aula citada antes, e por prompts padronizados de LLM, para garantir consistência entre as equipes. Pesquisas futuras podem incluir profissionais da indústria, acompanhar o uso do agente em projetos reais e investigar seus efeitos ao longo do tempo nas práticas colaborativas de desenvolvimento.

Agradecimentos

Este estudo foi financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Brasil, Código de Financiamento 001. Declaramos que a versão gratuita do ChatGPT foi utilizada para revisão textual de trechos específicos do artigo com a devida revisão dos autores.

Referências

- Adler, M. (2012). The Plain Language Movement. In *The Oxford Handbook of Language and Law*. Oxford Academic. <https://doi.org/10.1093/oxfordhb/9780199572120.013.0006>
- Arenas Arias, G. J. (2023). Publish, explain, understand, and comply: legislation in plain language. *The Theory and Practice of Legislation*, 11(2), 107–135. <https://doi.org/10.1080/20508840.2023.2182980>
- Backowski, R., Hinnant, K., & LaValle, L. (2022). Writing library database descriptions in plain language. *College & Undergraduate Libraries*, 29(3–4), 174–189. <https://doi.org/10.1080/10691316.2022.2149439>
- Card, D., & Bacigalupa, D. (2025). When Shall Must Go: A Case Study of Plain Language Implementation in State Government. In *Proceedings of the 43rd ACM International Conference on Design of Communication* (pp. 216-221). <https://doi.org/10.1145/3711670.3764645>
- Carvalho, L. V. D., Valle, P. H. D., Leifheit, B. R., Cabrejos, L. E. R., Nakamura, W., Guerino, G. C., ... & Silva, W. (2024). What do we know about usability evaluation for chatbots?: A systematic mapping study. In *Proceedings of the 20th Brazilian Symposium on Information Systems* (pp. 1-11). <https://doi.org/10.1145/3658271.3658324>
- Coutinho, V. D. O. G., Hazan, C., de Almeida Souza, M. R., & Freire, A. P. (2025). Practitioners' perceptions on human-centered design techniques in digital government development. In *Proceedings of the Conference on Digital Government Research*.
- Diniz, W., Gadelha, B., Steinmacher, I., & França, C. (2025). Habilidades Colaborativas no Mercado de TI: Uma Investigação sobre Requisitos de Soft Skills. In *Simpósio Brasileiro de Sistemas Colaborativos (SBSC)* (pp. 139-150). SBC.
- FDA. *United States Food and Drug Administration*. (s.d.). Plain Writing Act of 2010. Disponível em: <https://www.fda.gov/about-fda/plain-writing-its-law/plain-writing-act-2010>. Acesso em: 11 mar. 2025.
- Gov.br. (2025). Governo sanciona lei que garante comunicação pública mais clara e fácil de entender. Disponível em: <https://www.gov.br/servidor/pt-br/assuntos/noticias/2025/novembro/governo-sancion-a-lei-que-garante-comunicacao-publica-mais-clara-e-facil-de-entender>. Acesso em: 11 mar. 2025.

- Greene, M., Cleary, Y., & Marcus-Quinn, A. (2017). Use of plain-language guidelines to promote health literacy. *IEEE Transactions on Professional Communication*, 60(4), 384–400. <https://doi.org/10.1109/TPC.2017.2761578>
- Holmes, S., Moorhead, A., Bond, R., Zheng, H., Coates, V., & McTear, M. (2019). Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces?. In *Proceedings of the 31st European conference on cognitive ergonomics* (pp. 207-214). <https://doi.org/10.1145/3335082.3335094>
- Hu, S., Yen, H. C., Yu, Z., Zhao, M., Seaborn, K., & Liu, C. (2023). Wizundry: a cooperative Wizard of Oz platform for simulating future speech-based interfaces with multiple wizards. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW1), 1–34. <https://doi.org/10.1145/3579591>
- IPLF. *International Plain Language Federation*. (s.d.). What is plain language? Disponível em: <https://www.iplfederation.org/plain-language/>. Acesso em: 11 mar. 2025.
- Johannessen, M. R., Berntzen, L., & Ødegård, A. (2017). A review of the Norwegian plain language policy. In *Proceedings of the International Conference on Electronic Government* (pp. 187–198). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-64677-0_16
- Johnson, J. G., Peralta, M., Kaur, M., Huang, R. S., Zhao, S., Guan, R., Rajaram, S. & Nebeling, M. (2025). Exploring collaborative GenAI agents in synchronous group settings: eliciting team perceptions and design considerations for the future of work. *Proceedings of the ACM on Human-Computer Interaction*, 9(7), 1-33. <https://doi.org/10.1145/3757595>
- Joshi, R., Graefe, J., Kraus, M., & Bengler, K. (2024). Exploring the impact of explainability on trust and acceptance of conversational agents: a Wizard of Oz study. In *Proceedings of the International Conference on Human-Computer Interaction* (pp. 199–218). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-60606-9_12
- Lyu, Q., Tan, J., Zapadka, M. E., Ponnatapura, J., Niu, C., Myers, K. J., Wang, G., & Whitlow, C. T. (2023). Translating radiology reports into plain language using ChatGPT and GPT-4 with prompt learning: results, limitations, and potential. *Visual Computing for Industry, Biomedicine, and Art*, 6(1), 9. <https://doi.org/10.1186/s42492-023-00136-5>
- Menten, R., Roveloruiz, G., & Vanacken, D. (2024). NexOz: a Wizard of Oz approach to facilitate the integration of AI in interactive systems. In *Proceedings of the International Symposium on Engineering Interactive Computer Systems* (pp. 110–125). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-91760-8_8
- Moguel-Sánchez, R., Martínez-Palacios, C. S., Ocharán-Hernández, J. O., Limón, X., & Sánchez-García, A. J. (2023). Bots in software development: A systematic literature review and thematic analysis. *Programming and Computer Software*, 49(8), 712-734. <https://doi.org/10.1134/S0361768823080145>

- Nogueira, SV, Botelho, LDLR, Bolter, SG, & Driemeier, GS (2025). Movimentos pela Linguagem Simples: Democratização da informação e cidadania ativa. *Pesquisa, Sociedade e Desenvolvimento*, 14 (8). <https://doi.org/10.33448/rsd-v14i8.49345>
- Nunes, V., Cappelli, C., Michel de Moraes, R., Oliveira, R., & Costa da Silva, J. L. (2023). Method for evaluating the use of plain language in interfaces of public service portals. In *Proceedings of the 24th Annual International Conference on Digital Government Research* (pp. 233–243). <https://doi.org/10.1145/3598469.3598560>
- Oliveira, R., Bordalo, R., Salgado, L., & Cappelli, C. (2026a). Bringing Plain Language Principles into Interaction Design Practices: Insights from a Wizard of Oz Approach. In: *WorldCIST 2026. Lecture Notes in Networks and Systems*. Springer.
- Oliveira, R., Salgado, L., Seixas, F., & Cappelli, C. (2026b). How is plain language applied in interactive systems design? A systematic mapping study. In *Human-Computer Interaction – INTERACT 2025. Lecture Notes in Computer Science*, v. 16111. Cham: Springer. https://doi.org/10.1007/978-3-032-05008-3_49
- Oliveira, R., Salgado, L., Seixas, F., & Cappelli, C. (2023). Perspectivas de uso da Linguagem Simples no Design da Colaboração e Interação Humano-Dados do Governo Digital. In *Simpósio Brasileiro de Sistemas Colaborativos (SBSC)* (pp. 105-108). SBC. https://doi.org/10.5753/sbsc_estendido.2023.228382
- Parvathi, A., Indira, C., & Pillai, A. S. (2017). Plain English in software project management. *Journal of English Language Teachers' Interaction Forum*, 8, 52. Disponível em: <http://www.eltif.in/journals/Journal%20VIII.3.%20.pdf#page=52>. Acesso em: 11 mar. 2025.
- Pereira, A. J., Gomes, A. S., & Primo, T. T. (2022). Design de Sistema de Recomendação Educacional: abordagens com Mágico de Oz. In *Simpósio Brasileiro de Informática na Educação (SBIE)* (pp. 1184-1195). SBC. <https://doi.org/10.5753/sbie.2022.225760>
- Porcheron, M., Fischer, J. E., & Reeves, S. (2021). Pulling back the curtain on the wizards of Oz. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3), 1-22. <https://doi.org/10.1145/3432942>
- Ranchordás, S. (2022). The digitization of government and digital exclusion: setting the scene. In *The Rule of Law in Cyberspace* (pp. 125–148). Cham: Springer International Publishing.
- Sastre, A., Iglesias, A., Morato, J., & Sanchez-Cuadrado, S. (2024). Is ChatGPT able to generate texts that are easy to understand and read? In *Proceedings of the World Conference on Information Systems and Technologies* (pp. 138–147). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-60221-4_14
- Schrifer, K. A. (2017). Plain language in the US gains momentum: 1940–2015. *IEEE Transactions on Professional Communication*, 60(4), 343–383. <https://doi.org/10.1109/TPC.2017.2765118>
- Simpson, J., Stening, H., Nalepka, P., Dras, M., Reichle, E. D., Hosking, S., ... & Richardson, M. J. (2022). DesertWoZ: A Wizard of Oz environment to support the

design of collaborative conversational agents. In *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing* (pp. 188-192). <https://doi.org/10.1145/3500868.3559711>

Süße, T., Kobert, M., Grapenthin, S., & Voigt, B. F. (2023). AI-powered chatbots and the transformation of work: Findings from a case study in software development and software engineering. In *Working Conference on Virtual Enterprises* (pp. 689-705). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42622-3_49

Wagner, F. R., & de Medeiros Santos, A. L. (2026). Grandes Desafios da Computação 2025–2035. *Computação Brasil*, (55), 8-12. <https://doi.org/10.5753/sbc.17434.6>

Williams, C. (2022). The impact of plain language on legal English in the United Kingdom. Routledge. <https://doi.org/10.4324/9781003025009>.