



Eixo 6 - O mundo digital: apropriações e desafios

## O chatbot começa na biblioteca: bases de conhecimento para inteligência artificial generativa em unidades de informação

*The chatbot starts in the library:  
knowledge bases for generative artificial intelligence in information units*

**Leonardo Assis** – Universidade de São Paulo (USP) –  
leonardodasilvadeassis@gmail.com

**Resumo:** O trabalho apresenta um método de criação de bases de conhecimento para chatbots com inteligência artificial generativa em bibliotecas e unidades de informação. A partir de bibliografia recente sobre chatbots, serviços de referência e mediação da informação, argumenta-se que a qualidade da conversação depende menos da ferramenta escolhida e mais da organização documental que a sustenta. O método propõe coleta de conteúdos institucionais, anonimização de dados sensíveis, categorização temática, definição de escopo, integração gradual com sistemas existentes e avaliação contínua. Conclui-se que o chatbot deve ampliar, e não substituir, o trabalho profissional de curadoria e mediação.

**Palavras-chave:** Bibliotecas. Unidades de informação. Inteligência artificial generativa. Chatbots. Mediação da informação.

**Abstract:** This paper presents a method for creating knowledge bases for generative artificial intelligence chatbots in libraries and information units. Based on recent literature on chatbots, reference services and information mediation, it argues that conversational quality depends less on the selected tool than on the documentary organization that supports it. The method proposes the collection of institutional content, anonymization of sensitive data, thematic categorization, scope definition, gradual integration with existing systems and continuous evaluation. It concludes that chatbots should expand, rather than replace, professional work in information curation and mediation.

**Keywords:** Libraries. Information units. Generative artificial intelligence. Chatbots. Information mediation.



## 1 INTRODUÇÃO

A expansão da inteligência artificial generativa recolocou em novos termos uma discussão antiga no campo da Biblioteconomia, dos Estudos da Informação e das bibliotecas e unidades de informação: como construir formas de atendimento mais ágeis, interativas e capazes de lidar com um volume crescente de demandas sem reduzir a complexidade da mediação da informação. O debate sobre chatbots em bibliotecas e serviços de referência não é novo, mas ganha destaque com a inteligência artificial generativa, quando sistemas baseados em grandes modelos de linguagem passaram a produzir respostas mais fluídas, contextualizadas e próximas da linguagem natural. Esse movimento ampliou expectativas em torno de agentes conversacionais em bibliotecas e outras unidades de informação, mas também tornou mais visíveis problemas já conhecidos, como respostas imprecisas, riscos de alucinação, uso indevido de dados e dependência de plataformas externas.

A literatura permite reconhecer alguns eixos recorrentes. Um primeiro descreve um campo ainda em consolidação, concentrado em bibliotecas acadêmicas, páginas institucionais e serviços de referência textuais (Yan; Zhao; Mazumdar, 2023). Um segundo enfatiza experiência do usuário, desenho conversacional e participação dos profissionais da informação na elaboração das respostas (Mckie; Narayan, 2019). Um terceiro reúne relatos de implementação que apontam ganhos de eficiência, mas também obstáculos técnicos, organizacionais e éticos (Lee, 2025; Wagwu et al., 2024). Há ainda um eixo mais recente que discute a integração entre inteligência artificial generativa, catálogos, sistemas de descoberta e bases especializadas (Hazarika; Konch, 2025; Yun; Choi, 2024).

É nesse ponto que se situa a principal contribuição deste artigo. Em vez de partir da ferramenta, da plataforma ou do entusiasmo em torno de um modelo específico, o debate é recolocado no lugar de onde deveria começar: a instituição e seus próprios documentos, fluxos e registros de atendimento. Sustenta-se que a qualidade de um chatbot institucional depende, antes de tudo, da qualidade da base de conhecimento que o sustenta. Sem esse trabalho prévio, o agente conversacional tende a reproduzir respostas genéricas, pouco aderentes ao funcionamento da instituição e vulneráveis a erros que comprometem a confiança do público.



O objetivo é apresentar um método de criação de bases de conhecimento para chatbots de conversação, com ênfase em bibliotecas, mas com potencial de adaptação para outras unidades de informação. Esse método organiza-se a partir de um percurso que vai da documentação institucional às respostas conversacionais, passando pela coleta de conteúdos, retirada de dados pessoais e sensíveis, categorização por recorrências temáticas, definição de escopo e limites, testes e aprimoramento contínuo. Busca-se demonstrar, assim, que a criação de um chatbot não é, em primeiro lugar, um problema de interface, mas de organização, curadoria e governança da informação.

## 2 METODOLOGIA

O trabalho tem natureza teórico-metodológica e aplicada. Parte de uma revisão narrativa da literatura recente sobre chatbots em bibliotecas, serviços de referência, experiência do usuário, inteligência artificial generativa e mediação da informação, articulando esse conjunto bibliográfico à formulação de um método para criação de bases de conhecimento institucionais. A revisão não teve como finalidade realizar um mapeamento exaustivo do campo, mas identificar problemas recorrentes que atravessam a implementação desses sistemas: qualidade dos dados, desenho conversacional, delimitação de escopo, privacidade, governança e participação profissional.

A proposta metodológica também se apoia em experiência formativa desenvolvida em curso sobre criação de chatbots com inteligência artificial para bibliotecas e unidades de informação, no qual foram trabalhadas etapas de modelagem, definição de escopo, organização de bases de conhecimento, anonimização de registros e experimentação com ferramentas de conversação. Essa experiência permitiu observar que a construção do chatbot exige menos uma adoção imediata de plataformas e mais um processo prévio de leitura da instituição, de seus documentos, de seus atendimentos e de suas responsabilidades informacionais.

A análise considerou estudos que ajudam a visualizar tanto o estado de consolidação do tema quanto experiências de implementação. Yan, Zhao e Mazumdar (2023) contribuem ao mostrar que parte significativa das pesquisas sobre chatbots em bibliotecas ainda se concentra em aspectos de desenho, manutenção e avaliação



tecnológica. Mckie e Narayan (2019) oferecem uma chave importante para compreender a experiência do usuário e a necessidade de profissionais da informação participarem da construção da biblioteca de respostas. Schulz (2024) e Souza et al. (2025) permitem aproximar a discussão da entrevista de referência e da atuação bibliotecária supervisionada. Lee (2025), Wagwu et al. (2024), Hazarika e Konch (2025) e Yun e Choi (2024) ajudam a pensar implementação, integração com sistemas e qualidade dos conjuntos de dados.

A partir desse diálogo, o método proposto foi organizado em quatro movimentos. O primeiro consiste em deslocar o ponto de partida da ferramenta para a instituição. O segundo identifica o corpus informacional que pode alimentar a base de conhecimento, combinando documentos institucionais e registros de interação com o público. O terceiro define camadas de tratamento desse corpus: coleta, anonimização e categorização. O quarto articula a base resultante a escopo de atendimento, limites de atuação, padrões de resposta, integração gradual com sistemas existentes e avaliação contínua.

Esse percurso metodológico assume que o chatbot não cria conhecimento do zero. Ele reorganiza, recupera e apresenta, em linguagem conversacional, conhecimentos previamente documentados ou validados. Por isso, o foco da análise não está na escolha de uma plataforma específica, mas nas condições informacionais que tornam um agente conversacional mais confiável, situado e compatível com a realidade da instituição. A proposta pode ser aplicada a bibliotecas e outras unidades de informação, desde que respeite a documentação existente, as rotinas de atendimento e os direitos dos usuários.

### **3 RESULTADOS E DISCUSSÕES**

#### **3.1 O ponto de partida não é a ferramenta, é a instituição**

A principal defesa resultante da análise é simples, mas tem consequências decisivas: um chatbot institucional deve começar pela instituição, não pela plataforma. Isso significa inverter uma lógica comum no mercado e em parte da literatura aplicada, que inicia o processo pela escolha da ferramenta, pela contratação de um fornecedor ou pelo entusiasmo com determinado modelo de inteligência artificial. Quando se parte



desse lugar, a tendência é adaptar a instituição à tecnologia. O movimento inverso consiste em adaptar a tecnologia às condições, às necessidades, aos documentos e às formas de mediação da instituição.

Nas bibliotecas e outras unidades de informação, o atendimento cotidiano já produz um repertório valioso de conhecimento institucional. Esse repertório está disperso em normas, regulamentos, manuais, políticas de empréstimo, documentos sobre acessibilidade, perguntas frequentes, e-mails, formulários, registros de mediação, descrições de serviços, programação cultural e informações sobre acervo. Em muitos casos, também está presente na memória profissional das equipes. O problema é que esse conhecimento frequentemente permanece fragmentado, pouco sistematizado e, portanto, difícil de ser convertido em resposta conversacional consistente.

Partir da instituição implica reconhecer que o modelo pode ser avançado, mas, se a base estiver desorganizada, a resposta continuará ruim. Essa distinção desloca o debate do milagre tecnológico para o trabalho de curadoria informacional. Um chatbot só será institucionalmente confiável se suas respostas puderem ser ancoradas em documentos, fluxos e decisões reais da organização.

### **3.2 Corpus, coleta e anonimização**

Uma vez definido que o ponto de partida é a instituição, a etapa seguinte consiste em identificar o corpus que poderá alimentar a base de conhecimento. Esse corpus deve ser composto por duas grandes famílias documentais. A primeira reúne documentos institucionais estáveis ou semiestáveis: regimentos, regulamentos, normas de funcionamento, guias de serviços, protocolos de atendimento, políticas de uso de espaço, textos sobre acessibilidade, políticas de privacidade, orientações sobre cadastro, descrições de coleções, calendários, programação cultural e páginas oficiais. A segunda família é formada pelos registros de interação: perguntas frequentes, mensagens de atendimento, formulários recebidos, demandas recorrentes, dúvidas sobre serviços e respostas já validadas pelas equipes.

Essas duas famílias cumprem funções diferentes. Os documentos institucionais oferecem autoridade e estabilidade; os registros de interação mostram como o público pergunta, quais termos usa, quais dúvidas reaparecem e onde a documentação formal não responde bem. Por isso, a base de conhecimento não deve ser apenas normativa,



nem apenas empírica. Ela precisa combinar a linguagem institucional com a linguagem do usuário. É nesse encontro que a conversação ganha aderência.

A coleta, portanto, não é acumulação indiscriminada. É uma seleção orientada por finalidade: responder melhor ao público sobre serviços, acervos, regras, eventos, acessibilidade, formas de contato e encaminhamentos. A instituição identifica materiais, separa conteúdos oficiais, verifica versões atualizadas, retira documentos obsoletos e registra a origem de cada informação. Um documento entra na base quando ajuda a responder a uma demanda real ou previsível.

A anonimização é condição inegociável. Históricos individuais de usuários, dados de empréstimo nominados, informações sobre saúde, orientação sexual, vínculos familiares, perfis ideológicos e outros dados sensíveis não devem ser levados diretamente para a base. O que interessa ao projeto não é a identidade da pessoa atendida, mas o padrão informacional da dúvida, do problema ou da resposta validada. Isso significa retirar nomes, números de identificação, endereços e qualquer elemento que permita reidentificação. A declaração da IFLA sobre bibliotecas e inteligência artificial enfatiza que privacidade, equidade e liberdade intelectual devem orientar o uso de IA em ambientes informacionais (IFLA, 2020).

### **3.3 Categorização, escopo e limites**

Depois de coletado e anonimizado, o material precisa ser organizado em temas, intenções e situações de atendimento. Em uma biblioteca, por exemplo, as categorias podem envolver cadastro, empréstimo, renovação, reserva, localização de materiais, bases de dados, acesso remoto, acessibilidade, uso de espaços, programação cultural e encaminhamentos. Em outras unidades de informação, podem envolver acesso a serviços, consulta a coleções, instrumentos de pesquisa, recursos de acessibilidade e encaminhamentos especializados. A categorização transforma documentos dispersos em blocos de conhecimento que podem ser recuperados em linguagem conversacional.

Essa etapa também define padrões de resposta. O chatbot precisa saber quando responder diretamente, quando oferecer links ou documentos de apoio, quando fazer uma pergunta de esclarecimento e quando encaminhar para atendimento humano. Os limites de atuação são a contrapartida do escopo: o agente não deve opinar sobre temas que escapam ao conjunto documental que o alimenta, tomar decisões administrativas,



emitir pareceres ou substituir mediação humana em casos sensíveis. Ele pode orientar, localizar, encaminhar, resumir regras e apresentar possibilidades de busca, mas precisa reconhecer quando a demanda exige escuta especializada ou checagem adicional.

Ao cruzar essa etapa com a literatura, percebe-se que o desenho conversacional não pode ser separado da governança da informação. Mckie e Narayan (2019) já apontavam para a necessidade de envolver profissionais da informação na elaboração dos fluxos conversacionais. Radicaliza-se, aqui, essa percepção: o padrão de resposta é produto da própria base de conhecimento e do modo como ela foi classificada, hierarquizada e validada. Quando a base é frágil, o estilo também se fragiliza. Quando a base é sólida, o tom pode ser ajustado com mais segurança e consistência.

### **3.4 Integração, avaliação e sustentabilidade**

Outro resultado da análise é a necessidade de compreender o chatbot como interface de acesso a um ecossistema informacional, e não como camada isolada. O maior uso da inteligência artificial, nesse contexto, não é substituir a biblioteca, mas potencializar aquilo que ela já produziu: seus regulamentos, orientações, perguntas frequentes, registros de atendimento, políticas, serviços e formas de relação com a comunidade. No caso das bibliotecas e unidades de informação, isso abre possibilidades concretas: intermediar acesso ao catálogo, explicar regras de uso, orientar sobre acesso remoto, apresentar programações, indicar recursos de acessibilidade e encaminhar o usuário para páginas específicas.

Ao mesmo tempo, essa integração exige cuidado para que a promessa do serviço não ultrapasse o que a infraestrutura permite. Nem toda biblioteca ou unidade de informação terá equipe técnica, orçamento ou tempo para integrações profundas. Por isso, mostra-se pertinente um modelo gradual. Em uma fase inicial, a base pode ser alimentada por documentos institucionais e FAQs validadas. Em uma segunda fase, pode incorporar históricos de atendimento já anonimizados. Só depois disso, e se houver condições, a instituição avança para integrações mais robustas com catálogos, bases de dados e sistemas internos.

Nenhum chatbot institucional fica pronto quando a primeira versão é publicada. A avaliação contínua é condição de existência do serviço. Há testes de cobertura, precisão, linguagem, encaminhamento e atualização. Ehrenpreis e DeLooper (2025), ao



avaliarem um chatbot híbrido em biblioteca acadêmica, mostram que esse processo pode levar não apenas ao ajuste do agente, mas também à reorganização de páginas, guias e conteúdos institucionais. Essa revisão contínua desloca o debate de uma lógica de produto fechado para uma lógica de serviço em evolução.

A longevidade desse tipo de projeto depende de governança. Uma base de conhecimento não pode depender exclusivamente de uma pessoa, de um entusiasmo momentâneo ou de uma solução comprada sem apropriação interna. É preciso definir quem atualiza o corpus, quem valida novas respostas, quem acompanha alterações em regulamentos, quem observa falhas recorrentes e quem decide quando um conteúdo deve ser incluído, revisto ou removido. Sem esse arranjo mínimo, o chatbot envelhece rapidamente, perde alinhamento institucional e passa a comprometer o atendimento em vez de qualificá-lo.

A formação profissional aparece, aqui, como eixo estruturante. Parte das dificuldades de implementação decorre de insegurança, falta de capacitação e percepção de ameaça por parte das equipes. Em vez de tratar isso como resistência irracional, é mais produtivo interpretar tais reações como sinal de que o projeto precisa ser construído com participação e clareza sobre objetivos, limites e responsabilidades. A instituição que apresenta o chatbot como substituição inevitável do trabalho humano tende a produzir recusa. A instituição que o apresenta como ferramenta de apoio, organizada a partir de documentação e supervisionada por profissionais, tende a construir adesão mais consistente.

#### **4 CONSIDERAÇÕES FINAIS**

Ao longo deste artigo, buscou-se demonstrar que a criação de chatbots com inteligência artificial para bibliotecas e outras unidades de informação não deve ser pensada como simples adoção de ferramenta. O verdadeiro problema é informacional, documental e institucional. A literatura analisada mostra que os chatbots vêm sendo testados em serviços de referência, orientação, recomendação e integração com sistemas nas unidades de informação, mas também evidencia que os desafios de implementação continuam significativos. Desenho conversacional, qualidade dos dados,



formação de equipes, privacidade, infraestrutura e confiança institucional seguem no centro do debate.

A contribuição central foi sistematizar um método que recoloca a base de conhecimento no centro da arquitetura conversacional. Sustentou-se que um chatbot institucional deve ser construído a partir de documentos internos, registros de atendimento e repertórios já produzidos pela própria organização. Para que isso ocorra de forma ética e consistente, foram propostas três camadas fundamentais de tratamento informacional: coleta, anonimização e categorização. Assinalou-se também que a base precisa ser articulada a definições claras de escopo, limites de atuação, padrões de resposta, integração gradual com sistemas existentes e avaliação contínua. Sem esse percurso, o chatbot tende a se tornar um agente genérico, desconectado da realidade da instituição e mais vulnerável a erros.

Esse método permite uma leitura consistente do papel profissional nas unidades de informação. Em vez de substituir bibliotecários, documentalistas e mediadores, o chatbot evidencia a centralidade do seu trabalho. São esses profissionais que conhecem a documentação institucional, identificam recorrências do atendimento, reconhecem dados sensíveis, definem categorias, validam respostas e interpretam quando o sistema deve ou não responder. A inteligência artificial, nesse modelo, aparece como extensão das capacidades de organização e mediação, e não como substituição do juízo profissional.

Por fim, sustenta-se que a passagem da documentação institucional à conversação é, antes de tudo, uma passagem de curadoria. O chatbot só será confiável se traduzir a instituição com precisão, responsabilidade e clareza. Em contextos de recursos limitados, essa perspectiva é especialmente importante, porque mostra que é possível começar pequeno, com poucos arquivos e escopo reduzido, desde que a base de conhecimento seja bem construída. Nesse sentido, o futuro dos chatbots em unidades de informação depende menos da ferramenta em si e mais da capacidade institucional de organizar, validar e transformar informação em diálogo público qualificado.



## REFERÊNCIAS

EHRENPREIS, M.; DELOOPER, J. Chatbot assessment: Best practices for artificial intelligence in the library. **portal: Libraries and the Academy**, v. 25, n. 4, p. 671-702, 2025. Disponível em: <https://doi.org/10.1353/pla.2025.a971028>. Acesso em: 15 jan. 2026.

HAZARIKA, H. J.; KONCH, P. K. LibraryGPT: Bridging AI and academic libraries with tawk and machine learning integration. **Library Hi Tech News**, 2025. Disponível em: <https://doi.org/10.1108/LHTN-05-2025-0085>. Acesso em: 15 jan. 2026.

IFLA - INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **IFLA statement on libraries and artificial intelligence**. 2020. Disponível em: <https://repository.ifla.org/handle/20.500.14598/1646>. Acesso em: 15 jan. 2026.

LEE, Y. A. Implementing AI chatbots in academic libraries. **Journal of the Korean Society for Library and Information Science**, v. 59, n. 4, p. 315-335, 2025. Disponível em: <https://doi.org/10.4275/KSLIS.2025.59.4.315>. Acesso em: 15 jan. 2026.

MCKIE, I. A. S.; NARAYAN, B. Enhancing the academic library experience with chatbots: An exploration of research and implications for practice. **Journal of the Australian Library and Information Association**, v. 68, n. 3, p. 268-277, 2019. Disponível em: <https://doi.org/10.1080/24750158.2019.1611694>. Acesso em: 15 jan. 2026.

SCHULZ, O. What does ChatGPT have to do with library instruction and reference interviews? **Theological Librarianship**, v. 17, n. 2, p. 20-22, 2024. Disponível em: <https://doi.org/10.31046/1wgwab76>. Acesso em: 15 jan. 2026.

SOUZA, N.; PASSOS, K. G. F. dos; DITTRICH, E.; SILVA, L. M.; PICALHO, A. C. O uso do ChatGPT no serviço de referência em Bibliotecas Universitárias. **BiblioCanto**, v. 11, n. 1, p. 47-72, 2025. Disponível em: <https://doi.org/10.21680/2447-7842.2025v11n1ID39058>. Acesso em: 15 jan. 2026.

WAGWU, V.; OKPALA, A. E.; MAXWELL, C. E.; OLADOKUN, B. D. AI-driven chatbots for library services towards transforming patrons' engagement in Nigerian university libraries. **Communicate: Journal of Library and Information Science**, v. 26, n. 1, p. 89-98, 2024. Disponível em: <https://www.cjolis.org/cjolis/article/view/78>. Acesso em: 15 jan. 2026.

YAN, R.; ZHAO, X.; MAZUMDAR, S. Chatbots in libraries: A systematic literature review. **Education for Information**, v. 39, n. 4, p. 431-449, 2023. Disponível em: <https://doi.org/10.3233/EFI-230045>. Acesso em: 15 jan. 2026.

YUN, J. I.; CHOI, S. ChatGPT-based book recommendation system for learning Korean in a university library. **Journal of the Korean Society for Information Management**, v. 41, n. 3, p. 145-169, 2024. Disponível em: <https://doi.org/10.3743/KOSIM.2024.41.3.145>. Acesso em: 15 jan. 2026.

