چکیده مبسوط

(چکیده توسعه یافته Extended Abstract)

عنوان: ارائه چارچوبی جهت طراحی و ارزیابی چت‌بات هوش مصنوعی برای خدمات دولتی

نام نویسندگان: ناهید انتظاریان، دکتری مدیریت فناوری اطلاعات (گرایش کسب و کارهای هوشمند)، دانشکده علوم اداری و اقتصاد، دانشگاه فردوسی مشهد، na.entezarian@mail.um.ac.ir

محمد مهرآیین، استاد سیستم های اطلاعاتی، دانشکده علوم اداری و اقتصاد، دانشگاه فردوسی مشهد، Mehraeen@um.ac.ir

مقدمه: گسترش هوش مصنوعی و پردازش زبان طبیعی فرصت‌های جدیدی برای بهبود دسترسی، کیفیت و کارآمدی خدمات دولتی فراهم می‌کند. بااین‌حال، طراحی چت‌بات‌های هوش مصنوعی در بخش دولتی، علاوه بر الزامات فنی و کارکردی، مستلزم توجه به ارزش‌های عمومی و الزامات نهادی، از جمله قابلیت اعتماد، شفافیت، امنیت، حریم خصوصی و پاسخ‌گویی است. تمرکز مطالعات پیشین بر قابلیت‌های فنی یا رضایت کاربران، ضرورت تدوین چارچوبی یکپارچه برای طراحی چت‌بات‌های دولتی را برجسته می‌کند. پژوهش حاضر با هدف شناسایی الزامات و عناصر کلیدی طراحی، تعیین معماری مناسب و تدوین معیارهای ارزیابی، چارچوبی مفهومی برای طراحی چت‌بات‌های هوش مصنوعی دولتی ارائه می‌دهد و تناسب آن را برای ارائه خدمات حقوقی در بخش تعاون بررسی می‌کند.

روش‌شناسی: این پژوهش با رویکرد علم طراحی و با بهره‌گیری از روش‌شناسی کیفی چندروشی انجام شد. فرایند پژوهش شامل مرور نظام‌مند ادبیات، تحلیل استنتاجی، مصاحبه‌های اکتشافی با خبرگان و ارزیابی کیفی معماری مفهومی پیشنهادی بود. در مرحله نخست، ۱۱۲ مقاله با استفاده از پروتکل گزارش‌دهی ترجیحی برای مرورهای نظام‌مند و فراتحلیل (پریزما) بررسی و تحلیل شد. یافته‌های مرور ادبیات با استفاده از تحلیل استنتاجی ترکیب شد تا الزامات طراحی چت‌بات‌های هوش مصنوعی دولتی در دو دسته کارکردی و غیرکارکردی طبقه‌بندی شود. در مرحله بعد، مصاحبه‌های اکتشافی با ۱۰ نفر از خبرگان هوش مصنوعی و طراحان چت‌بات‌های سازمانی برای شناسایی عناصر کلیدی طراحی انجام گرفت. همچنین، با مصاحبه با ۱۵ نفر از خبرگان چت‌بات‌های هوش مصنوعی و ذی‌نفعان بخش دولتی، معیارهای ارزیابی در سه حوزه کیفیت سیستم، کیفیت اطلاعات و کیفیت خدمات تعیین شد. در ادامه، معماری مفهومی پیشنهادی در یک گروه کانونی متشکل از ۱۷ نفر از خبرگان هوش مصنوعی و ذی‌نفعان بخش دولتی، از نظر تناسب با نیازهای ارائه خدمات حقوقی در بخش تعاون و ابعاد ارزیابی تعیین‌شده، به‌صورت کیفی بررسی شد. تمرکز

ارزیابی بر اعتبار مفهومی و تناسب معماری پیشنهادی بود و نه سنجش فنی عملکرد یک چت‌بات عملیاتی.

یافته های اصلی: یافته‌های مرور نظام‌مند ادبیات و تحلیل استنتاجی، مجموعه‌ای از الزامات طراحی را در دو دسته شناسایی کرد. الزامات کارکردی شامل درک زبان طبیعی، پاسخ‌گویی دقیق، مدیریت و شخصی‌سازی گفتگو، قابلیت همکاری با سامانه‌های دیگر، یادگیری و دریافت بازخورد، ارجاع به اپراتور انسانی و احراز هویت و مدیریت مجوزهای دسترسی بود. الزامات غیرکارکردی نیز قابلیت اطمینان، دسترس‌پذیری، امنیت، حفاظت از حریم خصوصی، شفافیت، قابلیت نگهداری، انعطاف‌پذیری، کارایی و قابلیت استفاده را دربرگرفت. مصاحبه‌های خبرگان به شناسایی عناصر کلیدی معماری، از جمله اجزای سمت کاربر و سمت سرور، رابط کاربری، مؤلفه‌های پردازش زبان طبیعی، جست‌وجوی معنایی، مدیریت گفتگو، پایگاه دانش، رابط‌های برنامه‌نویسی کاربردی، نظارت بلادرنگ، ثبت و تحلیل رویدادها، تولید پاسخ و سازوکارهای امنیت و حفاظت از حریم خصوصی انجامید. بر اساس ترکیب این یافته‌ها، معماری تولید تقویت‌شده با بازیابی به‌عنوان رویکرد معماری پیشنهادی برای پشتیبانی از پاسخ‌گویی مبتنی بر اسناد معتبر دولتی انتخاب شد. همچنین، معیارهای ارزیابی در سه حوزه کیفیت سیستم، کیفیت اطلاعات و کیفیت خدمات سازمان‌دهی شد. نتایج ارزیابی کیفی گروه کانونی نیز نشان داد که معماری مفهومی پیشنهادی، از دیدگاه مشارکت‌کنندگان، با نیازهای ارائه خدمات حقوقی در بخش تعاون تناسب دارد.

بحث و نتیجه گیری: پژوهش حاضر با تلفیق الزامات فنی، نیازهای کاربران و ملاحظات اختصاصی بخش دولتی، چارچوبی مفهومی برای طراحی چت‌بات‌های هوش مصنوعی دولتی ارائه می‌کند. یافته‌ها نشان می‌دهد که طراحی چنین سامانه‌هایی صرفاً به انتخاب مدل هوش مصنوعی یا بهبود قابلیت‌های پردازش زبان طبیعی محدود نمی‌شود، بلکه مستلزم توجه هم‌زمان به کیفیت سیستم، اعتبار و تناسب اطلاعات، کیفیت خدمات، امنیت، شفافیت و پاسخ‌گویی است. پیشنهاد معماری تولید تقویت‌شده با بازیابی نیز بر اهمیت اتکا به اسناد معتبر در پاسخ‌گویی به پرسش‌های مرتبط با خدمات حقوقی دولتی تأکید می‌کند. ارزیابی کیفی انجام‌شده، شواهدی درباره تناسب مفهومی معماری پیشنهادی از دیدگاه خبرگان و ذی‌نفعان فراهم می‌آورد؛ بااین‌حال، این نتیجه به معنای اثبات عملکرد فنی یا اثربخشی عملیاتی سامانه نیست. نوآوری پژوهش در ارائه چارچوبی یکپارچه است که الزامات کارکردی و غیرکارکردی، عناصر معماری و معیارهای ارزیابی را در بستر خدمات دولتی به یکدیگر پیوند می‌دهد. این چارچوب می‌تواند راهنمای طراحی و توسعه چت‌بات‌های دولتی، به‌ویژه در حوزه خدمات حقوقی، باشد. در پژوهش‌های آینده، پیاده‌سازی نمونه عملیاتی،

ارزیابی کمی دقت و بازیابی اطلاعات، بررسی نرخ پاسخ‌های نادرست، مقایسه روش‌های معماری و سنجش تجربه کاربران می‌تواند اعتبار و کارآمدی چارچوب را در محیط واقعی بررسی کند.

کلیدواژه ها: چت‌بات هوش مصنوعی؛ خدمات دولتی؛ معماری تولید تقویت‌شده با بازیابی؛ الزامات طراحی؛ عناصر طراحی؛ معماری چت‌بات؛ ارزیابی کیفی

Extended Abstract

Title

Developing a Framework for the Design and Evaluation of Artificial Intelligence Chatbots for Government Services

Authors

Nahid Entezarian
Ph.D. In Information Technology Management (Intelligent Business), Faculty of Administrative Sciences and Economics, Ferdowsi University of Mashhad, Mashhad, Iran
Email: na.entezarian@mail.um.ac.ir

Mohammad Mehraeen
Professor of Information Systems, Faculty of Administrative Sciences and Economics, Ferdowsi University of Mashhad, Mashhad, Iran
Email: Mehraeen@um.ac.ir

Introduction

The growing application of artificial intelligence (AI) and natural language processing (NLP) creates new opportunities to improve the accessibility, quality, and efficiency of government services. However, designing AI chatbots for the public sector requires attention not only to technical and functional requirements but also to public values and institutional considerations, including trustworthiness, transparency, security, privacy, and accountability. The focus of previous studies on either technical capabilities or user satisfaction highlights the need for an integrated framework for designing government chatbots. Accordingly, this study aims to identify key design requirements and elements, determine an appropriate architecture, and establish evaluation criteria. It proposes a conceptual framework for designing AI chatbots for government services and examines its suitability for delivering legal services in the cooperative sector.

Methodology

This study adopted a Design Science Research (DSR) approach and employed a qualitative multi-method research design. The research process comprised a systematic literature review, abductive analysis, exploratory expert interviews, and qualitative evaluation of the proposed conceptual architecture. First, 112 articles were reviewed and analyzed using the Preferred

Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol. The literature review findings were combined with abductive analysis to classify the design requirements of government AI chatbots into functional and non-functional categories. Subsequently, exploratory interviews were conducted with 10 AI experts and designers of organizational chatbots to identify key design elements. In addition, interviews with 15 AI chatbot experts and public-sector stakeholders were conducted to establish evaluation criteria across three dimensions: system quality, information quality, and service quality. Finally, the proposed conceptual architecture was qualitatively examined through a focus group comprising 17 AI experts and public-sector stakeholders. The evaluation focused on the conceptual validity and suitability of the proposed architecture in relation to the requirements of legal service delivery in the cooperative sector and the identified evaluation dimensions, rather than on the technical performance of an operational chatbot.

Main Findings

The findings of the systematic literature review and abductive analysis identified a set of design requirements in two categories. Functional requirements included natural language understanding, accurate response generation, conversation management and personalization, interoperability with other systems, learning and feedback mechanisms, human operator handover, and authentication and access control. Non-functional requirements encompassed reliability, availability, security, privacy protection, transparency, maintainability, flexibility, efficiency, and usability. Expert interviews led to the identification of key architectural elements, including client-side and server-side components, user interface, natural language processing components, semantic search, conversation management, knowledge base, application programming interfaces (APIs), real-time monitoring, event logging and analysis, response generation, and security and privacy protection mechanisms. Based on the synthesis of these findings, Retrieval-Augmented Generation (RAG) was identified as the proposed architectural approach to support responses grounded in authoritative government documents. Furthermore, the evaluation criteria were organized into three dimensions: system quality, information quality, and service quality. The results of the qualitative focus group evaluation also indicated that, from the participants' perspectives, the proposed conceptual architecture was suitable for meeting the requirements of legal service delivery in the cooperative sector.

Discussion and Conclusions

By integrating technical requirements, user needs, and public-sector-specific considerations, this study proposes a conceptual framework for designing AI chatbots for government services. The findings indicate that designing such systems involves more than selecting an AI model or improving natural language processing capabilities; it requires simultaneous consideration of system quality, information credibility and relevance, service quality, security, transparency, and accountability. The proposed use of Retrieval-Augmented Generation also highlights the importance of relying on authoritative documents when responding to inquiries concerning government legal services. The qualitative evaluation provides evidence of the conceptual suitability of the proposed architecture from the perspectives of experts and public-sector stakeholders. However, this finding does not constitute evidence of the technical performance or operational effectiveness of the system. The main contribution of this study lies in developing an integrated framework that connects functional and non-functional requirements, architectural elements, and evaluation criteria within the context of government services. This framework can serve as a guideline for designing and developing government AI chatbots,

particularly in the field of legal services. Future research can examine the validity and effectiveness of the framework in real-world settings through the development of an operational prototype, quantitative evaluation of information retrieval accuracy, assessment of incorrect response rates, comparison of alternative architectural approaches, and evaluation of user experience.

Keywords

Artificial Intelligence Chatbots; Government Services; Retrieval-Augmented Generation (RAG); Design Requirements; Design Elements; Chatbot Architecture; Qualitative Evaluation.