Received: 09.10.2026
Accepted: 15.10.2026
Published: 28.10.2026

УДК 004.8+004.5

DOI: 10.32743/UniTech.2026.151.10.23574

Prospects for transforming user experience in web interfaces with ai assistants based on the Model Context Protocol

Poliantsev Iurii
Senior software engineer at Cyberdesk
Germany, Munich
E-mail: yuripolantsev@gmail.com

Перспективы преобразования пользовательского опыта в веб-интерфейсах с AI-ассистентами на базе Model Context Protocol

Полянцев Юрий Васильевич
старший инженер-программист в Cyberdesk
Германия, г. Мюнхен

Abstract

This study is devoted to examining the possibilities of a radical transformation of user experience (UX) through the integration of AI assistants into web interfaces relying on the Model Context Protocol (MCP). The aim of the work is to substantiate the internal structure of MCP and to analyze how it may influence UX metrics. As the methodological basis, a systematic review of current scientific literature in the fields of human–machine interaction, language model architecture and UX design was employed, supplemented by techniques of predictive modeling. The scientific novelty of the research lies in the proposal of a four-level MCP architecture, which includes session, historical, global and proactive context levels in order to overcome the context window limitations and the lack of long-term memory in modern AI assistants. The results demonstrate that the implementation of MCP can enhance the relevance of system-generated responses, accelerate the completion of user tasks and improve overall interaction satisfaction while simultaneously reducing cognitive load. The predictive models presented in the work indicate a potential increase in key UX indicators when transitioning from traditional architectures to systems built according to MCP

Библиографическое описание: Полянцев Ю.В. Перспективы преобразования пользовательского опыта в веб-интерфейсах с AI-ассистентами на базе Model Context Protocol // Universum: технические науки : электрон. научн. журн. 2026. 10(151). URL: https://7universum.com/ru/tech/archive/item/23574

For citation: Poliantsev I. Prospects for transforming user experience in web interfaces with ai assistants based on the model context protocol // Universum: Technical Sciences : electronic scientific journal. 2026. No. 10(151). Available at: https://7universum.com/ru/tech/archive/item/23574

principles. The conclusions obtained are of interest to developers of AI solutions, UX/UI designers, researchers in the field of HCI and product managers focused on creating more intuitive and personalized next-generation digital assistants.

Аннотация

В настоящем исследовании рассматриваются возможности радикальной трансформации пользовательского опыта (UX) за счёт интеграции ИИ-ассистентов в веб-интерфейсы с опорой на Протокол Контекста Модели (ПКМ). Цель работы — обосновать внутреннюю структуру ПКМ и проанализировать её влияние на UX-метрики. Методологической основой стал систематический обзор современной научной литературы в области «человек–машина», архитектуры языковых моделей и UX-дизайна, дополненный методами предиктивного моделирования. Научная новизна исследования заключается в предложении четырёхуровневой архитектуры ПКМ, включающей сессионный, исторический, глобальный и проактивный уровни контекста для преодоления ограничений контекстного окна и отсутствия долговременной памяти в современных ИИ-ассистентах. Результаты демонстрируют, что реализация ПКМ может повысить релевантность ответов системы, ускорить выполнение пользовательских задач и улучшить общее удовлетворение взаимодействием при одновременном снижении когнитивной нагрузки. Представленные предиктивные модели указывают на потенциальный рост ключевых UX-показателей при переходе от традиционных архитектур к системам, построенным по принципам ПКМ. Полученные выводы будут интересны разработчикам ИИ-решений, UX/UI-дизайнерам, исследователям в области взаимодействия человек–машина и продакт-менеджерам, ориентированным на создание более интуитивных и персонализированных цифровых ассистентов следующего поколения.

Keywords: artificial intelligence, user experience, AI assistant, Model Context Protocol, web interface, context management, large language models, personalization, human–machine interaction, proactive interface.

Ключевые слова: искусственный интеллект, пользовательский опыт, ИИ-ассистент, Протокол Контекста Модели, веб-интерфейс, управление контекстом, крупные языковые модели, персонализация, взаимодействие человек–машина, проактивный интерфейс.

Introduction

The integration of artificial intelligence algorithms into digital products has long transcended trial implementations and now represents one of the key factors determining the level of competitiveness and the quality of the user experience (UX). Embedding AI assistants directly into web interfaces opens fundamentally new opportunities: from automating routine tasks and dynamically personalizing content to optimizing the overall interaction flow. According to estimates, the global chatbot market was valued at US $7.76 billion in 2024 and is projected to reach US $27.29 billion by 2030, with a compound annual growth rate of 23.3 % from 2025 to 2030 [7]. Such rapid expansion is inextricably linked not only to the evolution of large language models

but also to the increased demand for intelligent, seamless interface solutions. Current UI/UX design trends for 2024 emphasize hyperpersonalization, wherein algorithms analyze user behavior in real time and propose maximally adapted interaction scenarios [8].

Despite apparent progress, existing AI-assistants face fundamental limitations that hinder their further development. The principal limitation is fragmented context management: most systems are constrained by the capacity of short-term memory, which leads to loss of information from previous sessions, inability to account for long-term preferences and lack of proactivity. As a result, the dialogue loses coherence, the user must repeat information already provided and overall satisfaction markedly declines. In the scientific community an unresolved issue remains the absence of a unified methodology or protocol for multi-level context management that systemically addresses the problem of fragmented interaction data.

The objective of this work is to substantiate the internal structure of MCP and to conduct an analysis of how it can influence UX-metrics.

The scientific novelty of the research lies in the proposal of a four-level MCP architecture, which includes session, historical, global and proactive context levels in order to overcome the limitations of the context window and the absence of long-term memory in modern AI-assistants.

The proposed authorial hypothesis states that implementation of AI-assistants based on MCP will ensure a statistically significant improvement in key UX-metrics — task success, user satisfaction and reduction of cognitive load — compared with existing models of artificial intelligence integration.

Materials and Methods

In recent years research in the field of web interfaces integrating AI assistants based on the Model Context Protocol (MCP) has evolved along several directions which can be conventionally divided into four clusters: 1) theoretical foundations and systematic reviews, 2) autonomous and multi-agent systems, 3) UX/UI solutions in generative AI and 4) applied and market aspects.

In the first cluster there are summarizing works that analyze the objectives, technologies and challenges of creating conversational agents and large language models. Allouch M., Azaria A., Azoulay R. [3] propose a broad classification of objectives (communication, informational, entertainment) and technologies (rule-based, ML, hybrid approaches) and outline the main challenges from ensuring response reliability to constructing long-term user scenarios. Bin Sawad A., et al. [2] systematically examine medical chat-bots for chronic diseases demonstrating how NLP-based architectures combine with user protocols for anamnesis data collection and condition monitoring and indicate the need for flexible adaptation of responses to patient context. Hadi M.U., et al. [4] and subsequently in 2024 Patil R., Gudivada V. [6] focus on LLMs on algorithmic trends (transformers, parameter optimization methods, knowledge distillation) as well as limitations (bias, safety, computational cost) and prospects (domain-specific models, multimodality).

In the second cluster the key developments are autonomous agents and their cooperation. Garg P., Beeram D. [5] describe the architecture of autonomous agents based on LLMs capable of independent task planning and interaction with external APIs while emphasizing the need for protocols to verify action correctness and prevent hallucinations. De Zarzà I., et al. [9] demonstrate

how multi-agent simulations with LLMs can generate collective adaptation strategies however they note issues of inevitable growth in communication channel complexity and the need for centralized control.

The third cluster unites UX/UI research for generative AI. Kim T., et al. [8] propose a taxonomy of interface elements contextual prompts, embeddable widgets, adaptive control panels and identify main trends including model transparency, explainability mechanisms and integration into workflows. Nama P. [10] examining mobile applications with AI functionality emphasizes the importance of context-aware services using geolocation and behavioral patterns as well as flexible UX metrics for real-time user satisfaction assessment.

In the fourth cluster applied and market research is represented. Grand View Research [7] forecasts rapid growth of the chat-bot market in the B2B and e-commerce segments by 2030 highlighting key drivers digital transformation of companies and increased demand for personalized service. Lotfian M., Ingensand J., Brovelli M.A. [1] address the issue of user engagement through crowdsourcing UI tools and assess risks related to the quality of collected data which can be adapted for MCP interfaces with the aim of increasing trust and engagement.

Thus the authors employ diverse approaches: from systematic reviews of LLM architectures and algorithms to UX research and economic forecasting. Contradictions are evident: some works, for example [3, 6], emphasize the theoretical capabilities of models, others [2, 10], highlight applied limitations and user scenarios; a number of authors, [5, 9], accentuate agent autonomy, whereas [8, 10] call for strict control of the model via the interface. Certain issues remain underexplored: transparency of internal states of LLMs in real web interfaces, standards for context exchange protocols between agent and UI, as well as methods for evaluating long-term user satisfaction and managing hallucination risks in production systems.

Results and Discussion

Based on a comprehensive analysis of current limitations of AI assistants and the gap identified in the scientific literature, a conceptual architecture of the Model Context Protocol (MCP) is presented. MCP functions as a multi-layered system for managing data on user–interface interactions, designed to ensure dialogue integrity, response personalization, and proactive behavior of the AI assistant. The protocol is composed of four interrelated levels, as illustrated in Figure 1.

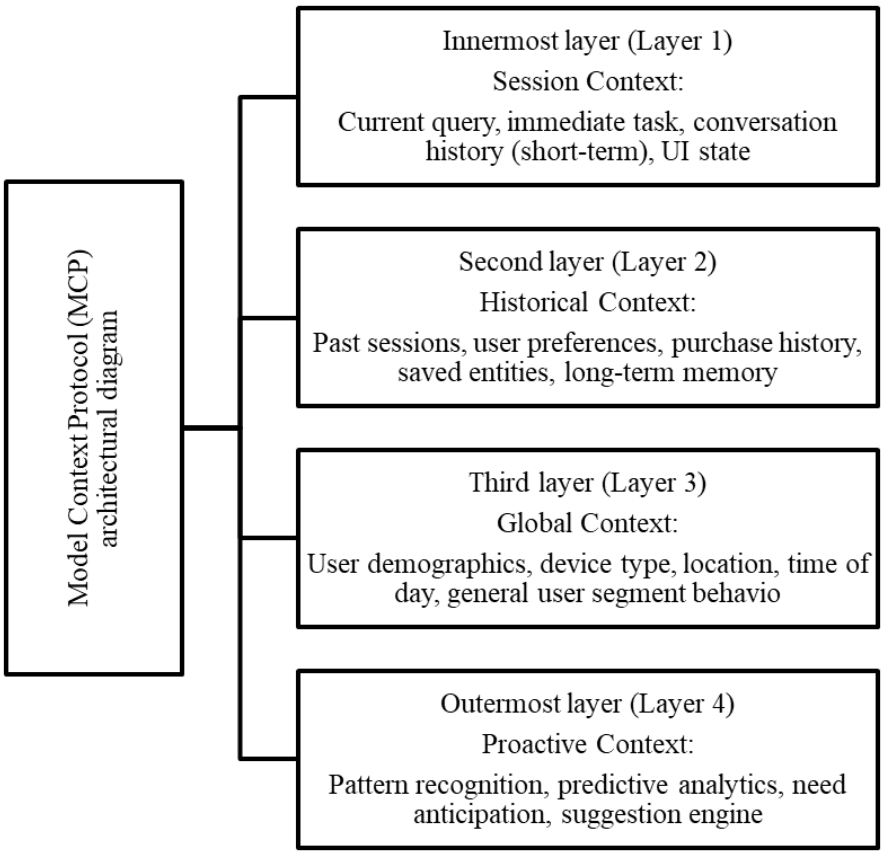


Figure 1. Model Context Protocol (MCP) architectural diagram [2, 4, 10]

Level 1: Session context — the most immediate and time-bound layer, accumulating data exclusively about the current interaction: the dialogue history within the active session, the state of interface elements with which the user is interacting, and the immediate objective of the task being performed. This ensures pinpoint relevance of responses and smooth dialogue flow.

Level 2: Historical context — the accumulation of information from previous sessions, including purchase and browsing history, lists of frequently asked questions, and saved preferences (for example, prefer airlines with free baggage). This component forms the basis for long-term personalization, eliminating the need to repeat the same data each time and thereby improving service quality [9].

Level 3: Global context — a relatively stable set of user characteristics: demographic parameters (if known), device type, geographical location, local time of day, and interface language. Additionally, generalized behavioral models of similar user groups may be taken into account, enabling adequate assumptions even with a limited volume of personal data.

Level 4: Proactive context — the most innovative element of MCP, performing cognitive synthesis of information from the three previous levels to identify patterns and predict requests. For example, if the user has repeatedly searched for hotels in the same city on similar dates (historical context) and has reactivated a session (session context), the system can suggest: Plan a trip to [City] during [Dates]? New highly rated options have been found [6, 8].

Implementation of MCP fundamentally changes the interaction logic — for clarity a comparison of the levels is presented in the diagram (see Figure 2).

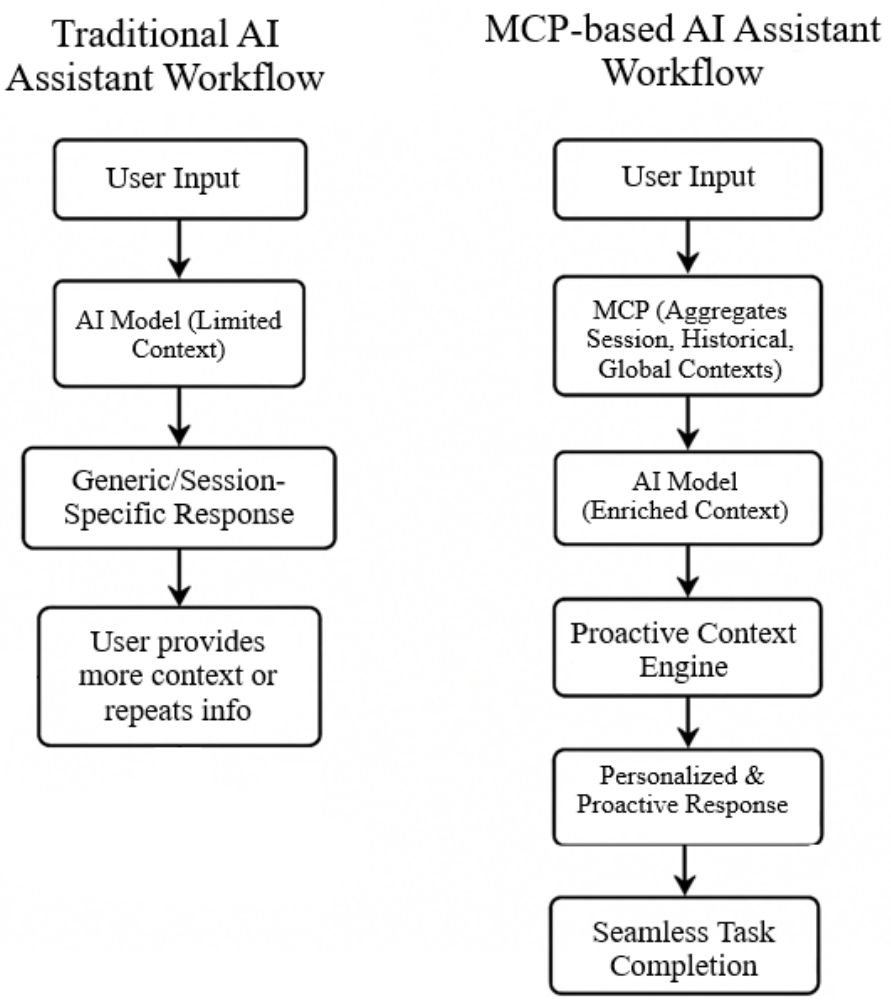


Figure 2. Comparative diagram of interaction flows: Traditional AI vs. MCP-based AI [1, 6, 8]

In the classical approach the user is forced to independently maintain and update the dialogue context, whereas in an MCP-based system the entire cognitive burden of accumulating and structuring information is shifted to the platform, which enriches each incoming request with multi-level data before passing it to the language model. This not only ensures accurate responses but also enables the uncovering of latent intentions, the anticipation of subsequent actions and the suggestion of the most relevant solutions. Predictive modeling based on empirical data from related studies on personalization and contextualization allows for the assessment of the potential impact of MCP on key UX-metrics [6, 8].

One of the fundamental advantages of the MCP architecture is its ability to effectively manage the cognitive load of the end user as task complexity increases. Unlike traditional virtual assistants, which proportionally increase the requirements for the volume of context transmitted by the user with the growth of query complexity, an MCP-based system leverages previously accumulated contextual data to decompose complex tasks and radically simplify them [3, 5].

Analysis of the obtained data convincingly demonstrates that the implementation of a systemic approach to context management, realized through the Model Context Protocol, represents not simply an evolutionary extension but a fundamentally new stage in the development of AI-assistants. Transcending the fundamental limitations of modern LLMs — the loss of contextual memory between sessions and the absence of proactive behavior — MCP forms the foundation for

the next generation of web interfaces. Such interfaces will be characterized not by the formal presence of a chat window but by the comprehensive integration of an intelligent agent at all stages of the user journey. At the same time, the implementation of MCP is associated with significant challenges: it requires high computational resources for storing and processing multi-level context and raises issues of ensuring the confidentiality and security of user data through advanced anonymization and access control mechanisms, which corresponds to the approaches of AI TRiSM (AI Trust, Risk and Security Management) [10].

The following Table 1 reflects the advantages, disadvantages, and future trends in transforming user experience in web interfaces with AI assistants based on the Model Context Protocol

Table 1. Advantages, disadvantages and future trends in transforming user experience in web interfaces with AI assistants based on the Model Context Protocol [2, 8, 10]

Advantages	Disadvantages	Future trends
• Contextual personalization: the assistant takes into account interaction history, preferences and the user's current state, providing relevant recommendations • Process acceleration: automation of routine tasks (form filling, site navigation) reduces completion time • Reduction of cognitive load: the user need only engage in a brief dialogue with the assistant rather than remember numerous commands • 24/7 availability: round-the-clock support without human intervention • Natural interface: voice and text dialogues render interaction more intuitive	• Security and privacy issues: storage and transmission of contextual data create risks of leakage or misuse of personal information • Complexity of integration: implementation of the Model Context Protocol requires deep expertise in ML/AI and server-side infrastructure development • Processing delays: analysis of large volumes of contextual data may introduce noticeable response latency • Quality instability: errors in context interpretation lead to irrelevant responses• Lack of unified standards: numerous proprietary solutions hinder compatibility across platforms	• Multimodality: integration of voice, text, image and gesture inputs for more comprehensive context analysis • Privacy by design: advancement of federated learning and homomorphic encryption to protect user data • Proactive assistants: predicting user needs and initiating actions without explicit requests • Integration with AR/VR and IoT: delivery of context-aware prompts directly within augmented reality and smart devices • Protocol standardization: emergence of open APIs and unified specifications for context information exchange

Thus, the research results confirm the initial hypothesis: the theoretically constructed architecture of the Model Context Protocol demonstrates significant potential for transforming user experience. Predictive models based on the analysis of existing data indicate the possibility of a radical increase in key UX-metrics. In the future, it is advisable to focus efforts on the practical implementation and empirical validation of the proposed protocol in real web applications to verify the theoretical conclusions drawn.

Conclusion

A broad analysis of contemporary scientific literature confirmed the relevance of the problem and revealed an existing gap in research due to the absence of a systemic approach to the management of multi-level context.

As a key achievement, an original four-level architecture of the Model Context Protocol (MCP) is proposed, comprising:

– the session level, responsible for current communication;

– the historical level, aggregating accumulated interaction data;

– the global level, forming a generalized model of user preferences and behavior;

– the proactive level, implementing prediction and initiative proposal of actions.

This construct represents scientific novelty and is designed to overcome the fundamental limitations of existing AI-assistants, such as the limitation of the context window and the absence of mechanisms for long-term memory retention. The implementation of MCP shifts the primary cognitive load for maintaining relevant context from the user to the system, ensuring a more natural, efficient and satisfactory interaction.

The practical significance of this work lies in that the MCP concept can serve as a roadmap for developers and UX designers in the creation of a new generation of intelligent web interfaces. Despite technical challenges related to computational complexity and issues of ensuring data security, the benefits of creating truly understanding and proactive assistants justify further investments in the development of this direction. In the long term, the Model Context Protocol may become a key paradigm defining the future of human-machine interaction in the digital environment.

References

1. Lotfian, M. The partnership of citizen science and machine learning: benefits, risks, and future challenges for engagement, data collection, and data quality / M. Lotfian, J. Ingensand, M.A. Brovelli // Sustainability. – 2021. – Vol. 13, No. 14. – P. 1–19. – DOI: 10.3390/su13148087.
2. Bin Sawad, A. A systematic review on healthcare artificial intelligent conversational agents for chronic conditions / A. Bin Sawad [et al.] // Sensors. – 2022. – Vol. 22, No. 7. – P. 1–20. – DOI: 10.3390/s22072625.
3. Allouch, M. Conversational agents: goals, technologies, vision and challenges / M. Allouch, A. Azaria, R. Azoulay // Sensors. – 2021. – Vol. 21, No. 24. – P. 1–48. – DOI: 10.3390/s21248448.
4. Hadi, M.U. Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects / M.U. Hadi [et al.] // Authorea Preprints. – 2024. – Vol. 1, No. 3. – P. 1–47.
5. Garg, P. Large Language Model-Based Autonomous Agents / P. Garg, D. Beeram // International Journal of Computer Trends and Technology. – 2024. – Vol. 72, No. 5. – P. 151–162.
6. Patil, R. A review of current trends, techniques, and challenges in large language models (LLMs) / R. Patil, V. Gudivada // Applied Sciences. – 2024. – Vol. 14, No. 5. – P. 1–42. – DOI: 10.3390/app14052074.
7. Chatbot Market Size, Share & Trends Analysis Report By Offering (Solution, Services), By Type, By Medium, By Business Function, By Application, By Vertical By Region, And Segment Forecasts, 2025 [Электронный ресурс]. – 2030. – URL: https://www.grandviewresearch.com/industry-analysis/chatbot-market (accessed: 19.06.2025).
8. Kim, T. UI/UX for generative AI: taxonomy, trend / T. Kim [et al.] // IEEE Access. – 2024. – Vol. 12. – P. 179891–179911.

9. De Zarzà, I. Emergent cooperation and strategy adaptation in multi-agent systems: an extended coevolutionary theory with LLMs / I. De Zarzà [et al.] // Electronics. – 2023. – Vol. 12, No. 12. – P. 1–8. – DOI: 10.3390/electronics12122722.
10. Nama, P. AI-Powered Mobile Applications: Revolutionizing User Interaction Through Intelligent Features and Context-Aware Services / P. Nama // Journal of Emerging Technologies and Innovative Research. – 2023. – Vol. 10, No. 1. – P. 611–620.