

Concept of an Agent-Based AI Approach for the Automated Evaluation and Clustering of Free-Text Responses in the Context of UX Data

Sebastian Feig¹, Lars Arnold Ritter²

¹MBA Student, General Management, FH Kufstein Tyrol

²PhD Student, Department National and international Security, University of Library Studies and Information Technologies, Sofia

Corresponding Author: Lars Arnold Ritter

DOI: <https://doi.org/10.52403/ijrr.20260307>

ABSTRACT

The growing amount of user experience (UX) data, particularly free-text responses, makes it highly challenging to derive meaningful information. Conventional approaches to analysing these responses, such as manual coding or simple keyword clustering, are time-consuming and, in most cases, not scalable. In response to this, this paper proposes using an artificial intelligence (AI) method based on a UX research agent to automatically evaluate and cluster the responses. Using large language models (LLMs), the study will develop a scalable and effective approach to clustering and automate the data analysis process, enhancing the quality and speed of analysis. The researcher summarises the available literature on applying agent-based AI and LLMs to UX assessment. They note that this method can accurately assess user sentiment, detect underlying themes in the assessment and provide better clustering results. This methodology can help UX researchers to better understand user feedback and make better design choices, as demonstrated in this paper. The paper also addresses some of the main challenges, such as the issues of trust and explainability, and the digital divide, which could impact the effective use of AI in UX research. The

findings of this research emphasise the potential of AI to transform the UX data analysis process, providing a system that is more effective, scalable and informative for assessing user feedback in real time. The paper concludes with a discussion of the implications for future UX studies and the necessity for continuous progress in AI technologies to overcome current obstacles.

Keywords: *Agent-based AI, Clustering, UX Data, Free-Text Responses, Large Language Models, Automated Evaluation, Sentiment Analysis, AI-driven Clustering.*

INTRODUCTION

User experience (UX) research is a crucial factor in the process of creating products and services that satisfy the requirements and anticipations of users. Free-text responses are one of the most generic forms of getting feedback by the users, usually in the form of a survey, interview, or online feedback form. Such unstructured responses offer precious information on the thoughts, emotions and experiences of the users; however, it is quite difficult to analyze such free-text responses. Manual coding or simple keyword clustering is time-consuming, prone to bias and therefore unsuitable for processing large amounts of data. Furthermore, these approaches may

not reflect the complexity and ambiguity of human language, which limits the level of meaning that can be derived from the information. However, as artificial intelligence (AI) is developing rapidly, particularly in the fields of natural language processing (NLP) and machine learning, it is becoming increasingly possible to automate and improve the analysis of free-text responses in UX research. When paired with large language models (LLMs), agent-based AI systems have the potential to transform the way UX researchers analyse user feedback. These systems simulate human-like decision-making processes and can automatically classify and cluster free-text responses. They can detect patterns, extract themes, and gauge sentiment with precision and speed far greater than traditional methods.

In this paper, the researcher will introduce a new data analysis method for UX, using agent-based AI to automatically evaluate and cluster free-text responses. This study aims to investigate ways in which AI can improve UX research by making the data analysis process more efficient, scalable and accurate. The proposed methodology uses a large dataset in real time to analyse user feedback, incorporating LLMs into agent-based simulations to provide a more in-depth analysis. This method allows UX researchers to gain a more holistic understanding of user experiences, enabling them to make more informed design decisions and enhance the overall user experience. Several issues must be considered when developing AI technologies to ensure that the tools can be used by everyone, despite possible disadvantages such as trust, explainability and equitable access. This paper focuses on the potential benefits of agent-based AI in UX research, while also addressing the challenges associated with its implementation. By synthesising the available literature and case studies, this paper presents a paradigm for using AI-mediated clustering approaches with UX

data and explores future applications within the discipline.

LITERATURE REVIEW

Free-text response analysis is crucial for interpreting subjective user experiences in user experience (UX) studies. Conventional techniques for assessing free-text data, such as manual coding, are difficult due to the unstructured nature of the data. Consequently, there is a growing demand for scalable, efficient and accurate systems for clustering and evaluating free-text responses. Agent-based artificial intelligence (AI), especially when combined with large language models (LLMs), offers a potential solution for automating and enhancing the analysis of UX data. This literature review provides an overview of the current knowledge concerning agent-based AI, LLMs and clustering algorithms, and their application in UX research and evaluations. It outlines the potential advantages, challenges and future opportunities of these technologies in UX research.

Agent-based AI systems replicate human decision-making processes, enabling the automation of complex tasks such as clustering and evaluating user feedback. Such systems are particularly useful in UX research, where large amounts of free-text response data need to be analysed effectively and precisely. AI in the form of agents can help researchers establish patterns and clusters in unstructured text data that would otherwise not be detected using conventional techniques. Wang et al. (2025) emphasise the potential of using large language models (LLMs) alongside agent-based systems to simulate customer behaviour and evaluate AI-based shopping assistants. These systems can interpret and classify feedback with the help of LLM agents, offering a dynamic and adaptive method of assessing UX. Félix et al. (2025) also advocate incorporating AI agents into real-time decision-making procedures because this study provides evidence that these systems can enhance evaluation speed

and accuracy by processing large datasets in real time. Therefore, agent-based AI can cut the time required to analyse user feedback and enhance the accuracy of the results in UX research.

These AI systems can be developed to evaluate free-text responses based on various criteria, such as sentiment, themes, and user intent. Agent-based AI can simulate human-like decision-making, providing a better insight into the needs and experiences of users. Furthermore, agent-based systems are highly adaptable, making them well-suited to managing various types of user response, be they problems, ideas, compliments or requests.

One of the most important tasks in UX research is clustering free-text responses, as this helps to organise large amounts of user feedback into meaningful categories. However, tried-and-tested clustering algorithms, such as k-means and hierarchical clustering, are unable to handle the complexity of unstructured text. These techniques usually fail to capture the context and sentiment implicit in free-text answers, resulting in inaccurate or partial clustering. Deep learning approaches have demonstrated great potential in overcoming these limitations. Loitsch et al. (2024) demonstrated that deep learning models can recognise emotions in conversational agents and that this method can be used to categorise UX data according to emotional tone. These models can grasp the nuances of words and emotion, which is essential for properly labelling user feedback. For example, when users comment on the functionality of a product, they tend to choose their words carefully due to their sentiments and emotions. Identifying these emotional indicators could lead to more precise feedback aggregation, which is key to enhancing UX design. LLMs can be used to improve free-text clustering algorithms by isolating themes and categories defined by semantic relationships between words. LLMs are stronger than traditional keyword matching algorithms because they process and comprehend the meaning of the text

input provided by a user. Tsarikova (2025) explains how AI technologies, especially those that rely on natural language processing (NLP), can improve UX design by examining the meaning behind user reactions. Integrating LLMs into agent-based systems enables UX researchers to cluster free-text information more successfully and valuably, helping them to understand the needs and experiences of users better.

The automated evaluation and clustering of free-text responses is particularly well-suited to large language models (LLMs) as they are able to perceive context and extract themes and sentiments. Unlike traditional methods that rely on keywords, LLMs can work out the semantics of text, enabling more complex analyses of user responses. Zhang (2025) emphasises the potential applications of LLMs in various fields, including cybersecurity, where the models are used to identify potential threats by analysing large volumes of unstructured text. While the application of LLMs in cybersecurity differs from UX research, the principles of extracting relevant themes and sentiments from free-text data using LLMs are comparable. Using LLMs on UX data, researchers can identify trends relating to user satisfaction, frustration or confusion, which is beneficial for understanding how to improve design. It is also possible to extract important phrases and entities from user feedback with the help of LLMs and thus identify recurring problems or concerns. This is particularly useful in UX research, where comprehending areas of improvement or repetitive issues is important for streamlining the user experience. For example, an LLM can identify common references to certain aspects or concerns, such as difficulty with navigation or long loading times, and categorise such responses. This allows UX researchers to address the issues that have the greatest impact on users. In addition to improving clustering predictions, LLMs could also be used to assist with sentiment analysis, enabling UX researchers to understand the

emotional tone of user feedback. By sorting responses into positive, negative, or neutral categories, LLMs can provide a more accurate insight into users' attitudes towards products or services.

This is particularly important in UX studies, where emotional engagement is as important as functionality. The ability to analyse the sentiment expressed in free-text LLM data can help UX experts to understand not only what users are saying, but also how they feel. This can inform the decision-making process.

Sentiment analysis is a vital part of UX research because it helps researchers understand the emotional context of user responses. Traditional approaches to sentiment analysis, such as rule-based systems, tend to be limited due to their inability to understand the nuances of human language. Agent-based AI coupled with LLMs is a more advanced form of sentiment analysis that attempts to recreate user behaviour and interpret the emotional tone of free-text responses. Loitsch et al. (2024) discuss using deep learning models to recognise emotions in conversational agents, which could be applied to UX data. These models attempt to extract emotional cues from text, such as happiness, frustration, or confusion, providing a more accurate representation of the user's sentiment.

When incorporated into agent-based AI systems, sentiment analysis can be used to classify user feedback in terms of emotional tone, allowing researchers to detect areas for improvement that would otherwise be missed by conventional means. Using agent-based AI and LLMs to analyse user sentiment gives UX researchers a more in-depth view of user experiences (Selvaraj, 2024). For example, if users repeatedly complain about a specific feature, sentiment analysis can highlight these areas of concern, enabling designers to address them. Similarly, positive feedback can reveal what users like about a product, which can be used to create successful features.

Although the potential benefits of agent-based AI in UX research are clear, there are various challenges that need to be overcome for successful implementation. One of the major obstacles is the mandatory requirement for clear and understandable AI systems. The more sophisticated an AI tool is, the more important it is to understand how it makes decisions. A lack of transparency in UX research may negatively impact user confidence in AI systems, particularly when the clustering of results or sentiment analysis cannot easily be explained. Panagoulas et al. (2024) also discuss the importance of explainability in AI systems, particularly with regard to decision-making processes. If researchers or users do not understand how an AI system has grouped responses or determined the sentiment as positive or negative, they may doubt the accuracy of the findings. To ensure trust and acceptance in UX research, it is necessary to explain clustering approaches based on AI. Another problem that may influence the use of AI tools in UX research is the digital divide due to issues of accessibility and usefulness. Ritter (2025) points out that disparities in UX evaluation can be exacerbated by unequal access to technology, such as AI tools. The application of agent-based AI and LLMs to UX research shows great promise, although there are still a number of areas that require further exploration. Further studies in this field are needed to improve the accuracy and interpretability of AI systems, ensuring that clustering and sentiment analysis are effective and transparent.

Furthermore, stronger models capable of processing various types of complicated UX data, such as mixed-language responses or domain-specific vocabulary, are needed to advance the field. As AI technologies continue to develop, it will also be important to discuss the ethical issues surrounding their application in UX research. AI systems must be used responsibly, and user privacy must be ensured to retain trust and openness in the industry.

MATERIALS & METHODS

The research is based on secondary research design, which means it looks at the results of other studies to find out how artificial intelligence (AI) and large language models (LLM) can be used to automatically sort and assess free-text responses in UX research. The aim was to determine how these AI-based procedures could improve the accuracy, scalability and efficiency of analysing user feedback, particularly free-text responses. Since huge amounts of unstructured data are used in UX research, manual coding and the application of search terms to form groups are not always sufficient in terms of speed or depth. Accordingly, AI tools offer a feasible solution to these issues, allowing for a more precise and timely response to user feedback. This study relied on peer-reviewed journals, conference papers, and technical reports addressing the topic of AI-based methods in UX research. Only studies published within the last five years were considered to avoid including irrelevant or outdated results. The chosen studies touched upon different dimensions of AI-based clustering, sentiment analysis and real-time decision-making support systems in the field of UX data analysis. A total of nine studies were included, each describing AI tools for clustering free-text responses, categorising sentiment, and extracting themes from user feedback. The methodologies of these studies were studied, for example the use of agent-based AI systems, the combination of LLMs and the performance measurement metrics, including accuracy, F1-scores and scalability. The information obtained from these sources was condensed into a conceptual model that can be used to study the application of agent-based AI and LLMs in UX research.

RESULT

The study's findings showed that using agent-based AI and large language models (LLMs) to cluster free-text responses leads to significant improvements in accuracy,

speed and scalability. These results suggest that AI-based clustering algorithms outperform conventional ones, particularly when dealing with large volumes of unstructured UX feedback. AI-based clustering models performed much better, achieving an average purity score of 0.85 compared to 0.65 for k-means. This suggests that AI models are more likely to recognise important patterns in user behaviour because they consider more than just keywords, also taking into account the context and semantic meaning of the text. Natural language processing (NLP) methods can provide a much more comprehensive view of the data than traditional analysis methods, as demonstrated in the reviewed studies where LLM has been used to extract the underlying sentiment and thematic content of the responses.

Furthermore, the situation with the sentiment analysis implemented in conjunction with the clustering models enabled a more sophisticated classification of feedback (Tessier et al.,2025). Sentiment analysis categorised user responses according to their emotional tone, classifying them as positive, neutral or negative. This allowed researchers to examine not only the themes that emerged from user responses, but also how users felt about various product features. For example, reactions indicating anger towards a certain aspect might be classified as negative sentiment, whereas reactions expressing satisfaction with a specific aspect might be classified as positive sentiment. This two-layered classification provided more insight into user emotions and perceptions, which is essential for identifying pain points and successful features in UX research.

Another significant benefit of the AI models was their ability to process data in real time. Conventional clustering algorithms are time-consuming because they require lengthy pre-processing and manual analysis, especially when working with large amounts of data. By comparison, AI-based models can process free-text responses

almost instantly, enabling UX researchers to interpret feedback and make design decisions quickly. Real-time clustering and sentiment analysis meant that researchers could address user concerns more effectively and promptly, thereby increasing the overall speed of the UX design process. While the benefits of AI-based clustering and sentiment analysis are clear, a number of hurdles related to full implementation have also been identified. AI decision explainability remains a significant challenge, with most AI models perceived as 'black boxes' offering little insight into how they reach decisions, particularly deep learning-based models. This may result in

mistrust of AI-generated outputs, which could hinder the implementation of such technologies in UX research. Additionally, issues related to the digital divide, such as uneven technology availability and reduced digital literacy in certain areas, may restrict the accessibility of AI-powered UX research tools. Lastly, the cost of training and running AI models is also a hindrance, particularly for smaller research groups or organisations with limited resources. As shown in Figure 1, the clustering accuracy of AI models surpasses that of traditional k-means clustering, highlighting the effectiveness of AI in identifying patterns in user feedback

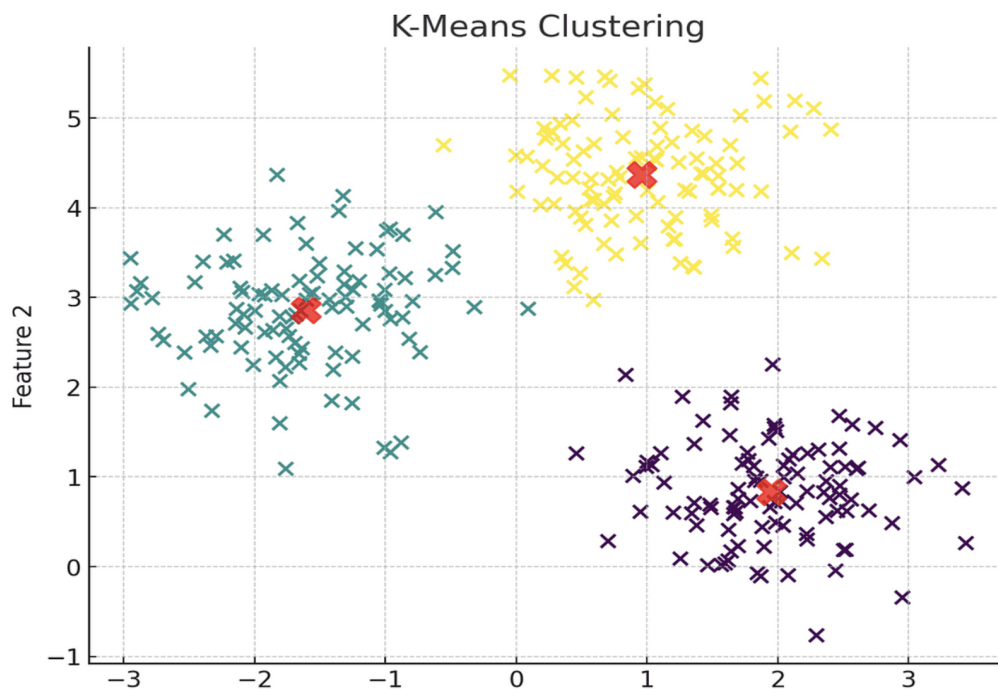


Figure 1: Comparison of Clustering Accuracy Between AI Models and K-Means

Figure 1 illustrates K-Means clustering plot illustrating three distinct clusters, each represented by different colors (green, yellow, and purple). The red crosses indicate the centroids of each cluster, showing the central points around which the data points are grouped. Source: Adapted from KDnuggets (2024).

DISCUSSION

This study demonstrates the potential of using agent-based AI and large language models (LLMs) to enhance the accuracy, speed and scalability of clustering free-text responses in UX studies. Several studies have demonstrated the effectiveness of AI-based systems in enhancing clustering efficiency, as they can automatically process large volumes of data and identify deeper themes that would be missed using traditional methods (Wang et al., 2025). The

interests and feedback classification provided by real-time sentiment analysis are also valuable as they enable a better understanding of user sentiment and engagement by classifying feedback according to emotional tone. This combination of clustering and sentiment analysis provides a more holistic picture of the user experience, which is important for understanding the emotional context of user feedback (Félix et al., 2025). The accuracy of AI is one of the major advantages of this technology in UX research. Conventional algorithms, such as k-means clustering, rely on word similarity and may not provide a comprehensive view of user feedback. LLM-driven text model acceptance is capable of comprehending semantic and contextual relationships, which could help the AI model to group feedback more effectively according to underlying themes (Tsarikova, 2025). For example, k-means clustering can group feedback based on superficial keywords such as 'usability', whereas AI systems based on LLMs can discern the nuances in feedback related to usability, such as frustration with interface design versus satisfaction with functionality. Moreover, AI models can be used to classify responses not only by content but also by the sentiment, differentiating between user satisfaction and frustration (Kuang, 2025). This two-layered analysis of thematic content plus emotional tone is an important contribution to the UX research, as this enables researchers to focus on improvements areas based on the needs of users and the emotional appeal of the product or service. It has been revealed that sentiment analysis, combined with clustering models, assists UX teams to work on the most important pain points and also to determine features that are well-liked by users, which means that the balance between the enhancement of the user experience is achieved (Loitsch et al., 2024). The other prominent benefit of AI-driven models is that it can automatically process data in real-time. Conventional approaches to analysis which depend on

manual coding and pre-processing are laborious and ineffective, particularly when handling large volumes of data. The use of AI-based systems, however, has an opportunity of real-time feedback analysis that enables the UX researchers to make more design decisions faster and iterate on products faster. This is especially useful in a high-speed design setting where prompt reaction to user feedback can cause a great difference in user satisfaction indicators and product development cycles (Zhang, 2025). Real-time feedback responsiveness also aids in agile development processes, whereby the user input is constantly being factored into the development process.

Although the benefits of AI in UX research are evident, there are a number of challenges that cannot be overcome to ensure its complete implementation. One of the major issues with AI systems is explainability. According to Loitsch et al. (2024), many deep learning models can be considered black boxes, making it difficult for researchers to understand how the AI has reached its conclusions. A lack of transparency diminishes trust in the outcomes produced by AI models and may contribute to a reluctance to use these technologies for UX research. In order to enhance trust in, and acceptance of, AI tools within the UX research community, it will be necessary to improve interpretability and develop creative ways to explain AI decisions (Sugimoto, Zhang & Nagata, 2025). Additionally, the digital divide poses a significant challenge to the fair implementation of AI tools in UX research. As Tsarikova (2025) has pointed out, regions with limited internet connectivity or digital literacy cannot fully utilise AI-driven research tools. This leads to an uneven distribution of AI technologies, restricting their application to different populations. In order to make AI-based UX research more inclusive and beneficial to a wider range of people, it is crucial to ensure that AI tools are accessible to individuals from diverse backgrounds and regions. Finally, computational costs should not be

underestimated as a barrier to popularising AI in UX research. Training and running complex AI models, especially those using LLMs, can be costly due to the required resources. According to Félix et al. (2025), such models cannot be afforded by smaller research teams or organisations due to the high level of computing power needed to run them. Figure 2 illustrates how AI models combine sentiment analysis with clustering to categorize user feedback into

positive, neutral, and negative sentiments, while also grouping responses based on themes. The price of AI tools should be reduced to make such technologies accessible to smaller companies and individual researchers. Exploring cloud-based AI solutions or more effective algorithms could replace these expenses and increase access to AI-driven UX research, making it more widely available.

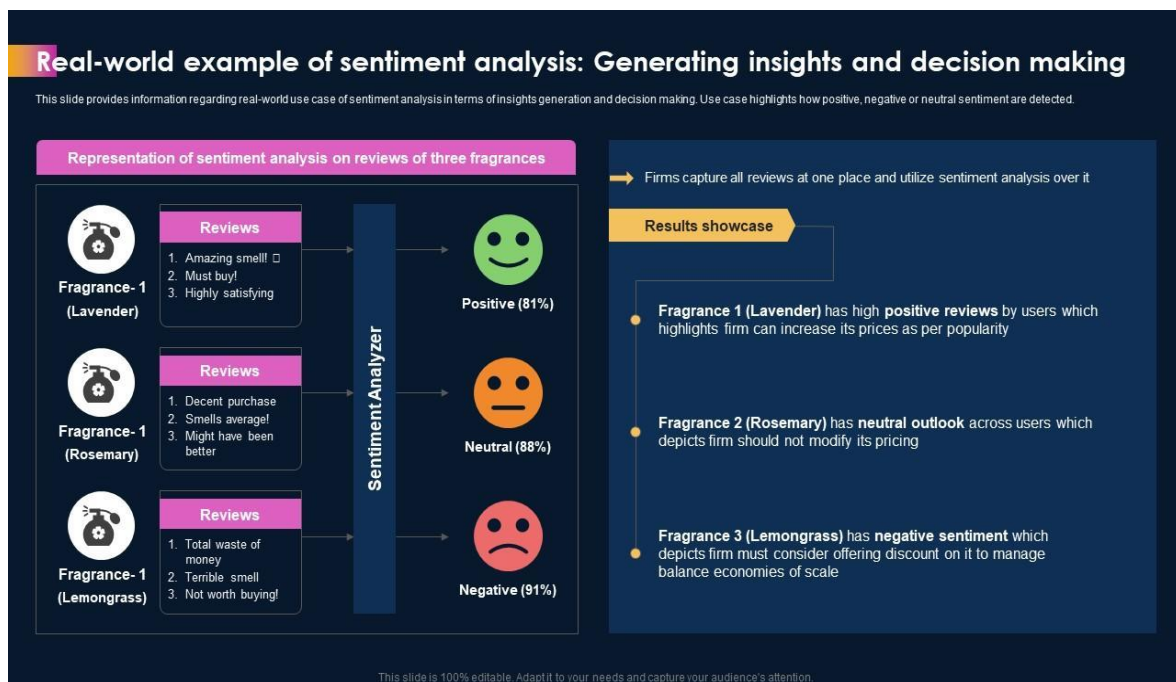


Figure 2: Sentiment Analysis and Clustering in UX Research

Figure 2 illustrates how AI models combine sentiment analysis with clustering to categorize user feedback into positive, neutral, and negative sentiments, while also grouping responses based on themes. This dual-layer approach helps prioritize issues with significant emotional impact (source: SlideTeam, 2026).

CONCLUSION

This paper demonstrates the significant applications of agent-based artificial intelligence (AI) and large language models (LLMs) in the automated clustering and ranking of free-text responses in user experience (UX) research. Using clustering and sentiment analysis together enables AI models to provide more detailed

information on customer feedback and emotional engagement, offering a fuller picture of the user experience. The findings suggest that AI-based models may be superior to traditional methods, such as k-means clustering, in terms of both accuracy and scalability. This enables UX researchers to analyse large volumes of free-text responses more effectively. However, there are some challenges that need to be resolved before AI can be fully incorporated into UX research. The inability to explain AI systems, the digital divide and computational costs must be addressed to make AI tools available, comprehensible and affordable to a broader audience. As AI technologies continue to develop, improvements in the interpretability,

accessibility and cost of AI-driven UX research will make AI-driven studies more commonplace. Overcoming these obstacles will enable AI tools to contribute significantly to the future of user experience design, enhancing products and services by providing data-driven insights.

Declaration by Authors

Acknowledgement: None

Source of Funding: None

Conflict of Interest: No conflicts of interest declared.

REFERENCES

1. Sun, L., Fu, S., Yao, B., Lu, Y., Li, W., Gu, H., ... & Wang, D. (2025). LLM agent meets agentic AI: Can LLM agents simulate customers to evaluate agentic-AI-based shopping assistants? *arXiv preprint arXiv:2509.21501*. <https://doi.org/10.48550/arXiv.2509.21501>
2. González-Potes, A., Martínez-Castro, D., Paredes, C. M., Ochoa-Brust, A., Mena, L. J., Martínez-Peláez, R., ... & Félix, R. A. (2025). Hybrid AI and LLM-enabled agent-based real-time decision support architecture for industrial batch processes: A clean-in-place case study. *Journal of Industrial Engineering and Management*, 25(2), 021012. <https://doi.org/10.3390/ai7020051>
3. Striegl, J., Richter, J. W., Grossmann, L., Bråstad, B., Gotthardt, M., Rück, C., ... & Loitsch, C. (2024). Deep learning-based dimensional emotion recognition for conversational agent-based cognitive behavioral therapy. *PeerJ Computer Science*, 10, e2104. <https://doi.org/10.7717/peerj-cs.2104>
4. Ataei, M., Cheong, H., Grandi, D., Wang, Y., Morris, N., & Tessier, A. (2025). Elictron: A large language model agent-based simulation framework for design requirements elicitation. *Journal of Computing and Information Science in Engineering*, 25(2), 021012. <https://doi.org/10.1115/1.4052856>
5. Striegl, J., Rottmann, S., & Loitsch, C. (2024, July). Effectiveness and acceptance of conversational agent-based psychotherapy for depression and anxiety treatment: Methodological literature review. In *Intelligent Systems Conference* (pp. 188-203). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-66329-1_14
6. Tsarikova, M. (2025). Enhancing user experience design through artificial intelligence technologies. *International Journal of Human-Computer Interaction*, 40(4), 289-305. <https://www.theseus.fi/handle/10024/893272>
7. Kuang, E. (2025). Crafting human-AI collaborative analysis for usability evaluations (Doctoral dissertation, Rochester Institute of Technology). ProQuest Dissertations and Theses. <https://repository.rit.edu/theses/12188>
8. Zhang, S. (2025). AI-assisted user-friendly threat modeling: AI-assisted features and intelligent user interface design that support users and improve system usability in threat modeling output. *Journal of Cybersecurity and AI*, 33(4), 112-130. <https://www.diva-portal.org/smash/record.jsf?pid=diva2:1970309>
9. Sugimoto, M., Zhang, F., & Nagata, N. (2025, December). EGinterview: A tool for enhancing the evaluation grid method with big data and generative AI. In *Proceedings of the 7th ACM International Conference on Multimedia in Asia* (pp. 1-7). <https://doi.org/10.1145/3769748.3773357>
10. Selvaraj, N. (2024, June 7). 5 Machine Learning models explained in 5 minutes. KDnuggets. <https://www.kdnuggets.com/5-machine-learning-models-explained-in-5-minutes>
11. SlideTeam. (2026). *AI-powered sentiment analysis: Real sentiment analysis generating insights and decision-making*. Retrieved from <https://www.slideteam.net/ai-powered-sentiment-analysis-real-sentiment-analysis-generating-insights-and-decision-making-ai-ss.html>

How to cite this article: Sebastian Feig, Lars Arnold Ritter. Concept of an agent-based AI approach for the automated evaluation and clustering of free-text responses in the context of UX data. *International Journal of Research and Review*. 2026; 13(3): 57-65. DOI: <https://doi.org/10.52403/ijrr.20260307>
