From Farm to Chatbot: Evaluating AI Responses Defining the Modern Agrifood System

Sarah McCord
Texas A&M University
sarah.mccord@ag.tamu.edu

Dr. Rafael Landaverde
Texas A&M University
rafael.q@ag.tamu.edu

From Farm to Chatbot: Evaluating AI Responses Defining the Modern Agrifood System

Introduction

Like the general public, agricultural audiences have increased their consumption of information through the increasingly common use of chatbots (Maduri et al., 2021). A chatbot is a software application, generally powered by artificial intelligence (AI), designed to simulate everyday conversations between a human sender and receiver (Adamopoulou & Moussiades, 2020). These systems have traditionally been used in customer service; however, in recent years, they have gained popularity in educational contexts as tools for providing rapid and automated responses (Landaverde, 2025). Despite their potential, these innovations have been subject to criticism regarding the credibility and validity of the information they provide (Musheyev et al., 2024). Within agrifood systems, one of the primary risks associated with the implementation of chatbots is the potential dissemination of information that has not been scientifically verified, which could encourage erroneous practices or result in economic losses for users (Sánchez-Gómez et al., 2026). In response, this study seeks to examine the validity of responses generated by commonly used chatbots and to compare the similarity of outcomes across standardized queries.

Conceptual Framework

Framing the study, we employed Kamra et al.'s (2025) model for evaluating AI in educational contexts. The Chatbot Evaluation Framework integrates three primary assessment pillars: AI model, product, and user journey (Kamra et al., 2025). First, assessing chatbots with an AI model assessment focus, or an evaluative design, allows domain experts to assess the chatbot for reliable, accurate, and valid chatbot responses. Next, user journey assessment shifts the focus to the user, suggesting that an effective chatbot experience achieves high usability and a positive user experience (Kamra et al., 2025; Olla et al., 2025). Finally, AI product assessment gives evaluators a sense of chatbot performance, including response times, error handling, and more (Kamra et al., 2025). This approach gave us an iterative, flexible, and inclusive framework for evaluating chatbots in an agricultural context.

Methodology

To complete the comparison of chatbot responses for our study, we first developed a clear, yet specific prompt: *Describe the defining characteristics of agrifood systems with validity and accuracy in the modern 21st century*. This prompt was employed as it aligns with measurable characteristics suggested by Kamra et al.'s (2025) evaluative framework to assess both validity and accessibility of the chatbot's responses with relevance to the current state of the agrifood system. We used this exact prompt for every chatbot tested, including ChatGPT, Microsoft Copilot, and Google Gemini. Prompts were entered into each chatbot immediately after one another and implemented through a free account and Incognito browser to limit previous algorithm influence. Chatbot responses were saved and analyzed for comparison and thematic analysis (Braun & Clarke, 2006). We used a reflexive journal to note differences and jot connecting key points between chatbot responses.

Results

We present the results from our evaluation of chatbot responses from three AI platforms. Between the AI systems, ChatGPT, Microsoft Copilot, and Google Gemini, responses were highly similar in nature and nuance, yet were presented by each platform differently. For instance, each chatbot numerically listed specific characteristics of modern agrifood systems,

much as a direct response to the exact prompt we asked. The number of characteristics chatbots produced varied. ChaptGPT categorized agrifood systems into 11 themes, where CoPilot and Gemini described eight and five characteristics, respectively. Despite these differences in the number of outlined agrifood systems defining traits, responses were highly similar in the broader picture when analyzing each platform's output. All three chatbots described the modern agrifood system as a holistic system or web of collectives. ChatGPT and CoPilot had matching verbiage suggesting agrifood systems are "complex, interconnected networks that encompass" etc. To aid in visualizing the overlap in responses of the chatbots, we combined the outputs to develop a word cloud. This visualization showed that the most represented words in the AI-generated responses were terms such as global, supply, environmental, markets, farmers, consumption, security, circular, data, nutrition, economic, technological, transformation, sustainability, and more. These terms are representative of the agrifood system, recognizing its contribution starting at the production level, stretching to retail and global markets, and beyond. Thus, our finding was that each chatbot produced similar interpretations of modern agrifood systems, although some outputs were more detailed (ChatGPT), versus others which included direct citations or more scientific language (Gemini), using terminology such as bioeconomics, fragmentation, etc.

Conclusions

Applying our findings to the Chatbot Evaluation Framework developed by Kamra et al. (2025) and evaluative AI models suggested by others (Olla et al., 2025), we suggest that the three AI chatbots we employed to respond to the prompt: *Describe the defining characteristics of agrifood systems with validity and accuracy in the modern 21st century,* provided cohesive outputs. In terms of performance validity, the chatbot's responses appeared to be accurate based on our research team's expertise in food systems information. Chatbots, however, outsource data from across entire servers, forcing a lack of scientific or academically correct information readily backed up by citations. We suggest that while the performance and responses were accurate in terms of defining and painting a clear picture of agrifood systems, users should locate reputable information independently of the chatbot. As for the user journey and project assessment, open AI chatbots such as those employed here provide an incredibly convenient experience for their audiences. Outputs to our entered prompt took less than 40 seconds to generate completely, and optionally led to additional requests to generate a summary table or a more complex interaction of ideas. While ease of use and accessibility to chatbots becomes increasingly convenient to rely on, there are also extreme implications to utilizing AI as more than a digital tool to enhance creativity and individual knowledge gains.

Implications & Recommendations

Finally, we underscore the importance of leveraging AI to support independent thinking and information-seeking behaviors (Hasan, 2025; Landaverde, 2025). Our results suggest that a variety of chatbots, ChatGPT, Microsoft CoPilot, and Google Gemini, provide user-friendly, timely, and similarly accurate information when answering our prompt for defining agrifood systems. Despite the positives we annotate, users must also recognize that, as learners, AI exists to enhance knowledge, rather than complete cognitive processes, dismantling individual capabilities for deep thinking (Jose et al., 2025). Likewise, adverse climate and environmental impacts continue to arouse serious concerns regarding the legitimacy and trustworthiness of using AI (Xiao et al., 2025). We challenge: at what cost does the utility of AI have on society? We recommend that as educators and communicators, chatbots have incredible value to leverage technology and enhance digital research and teaching aids, but with the oversight of their users.

References

Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In I. Maglogiannis, L. Iliadis, & E. Pimenidis (Eds.), *Artificial Intelligence Applications and Innovations* (Vol. 584, pp. 373–383). Springer International Publishing. https://doi.org/10.1007/978-3-030-49186-4_31

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Hasan, Md. K. (2025). How AI quietly undermines the joy and effort of learning: A call for rebalancing education in the digital age. *Annals of Medicine & Surgery*, *87*(8), 4693–4694. https://doi.org/10.1097/MS9.0000000000003456

Jose, B., Cherian, J., Verghis, A. M., Varghise, S. M., S, M., & Joseph, S. (2025). The cognitive paradox of AI in education: Between enhancement and erosion. *Frontiers in Psychology*, *16*, 1550621. https://doi.org/10.3389/fpsyg.2025.1550621

Kamra, A., Jagannath, R., & Gupta, R. (2025, May 16). *Evaluating AI for learning: A framework*. Precision Development (PxD). https://precisiondev.org/evaluating-ai-for-learning-a-framework/

Landaverde, R. (2025, November 11). *Advancing agronomic outreach: The role of AI-powered chatbots in scientific information delivery* [Poster abstract]. CANVAS 2025, Salt Lake City, UT. https://scisoc.confex.com/scisoc/2025am/meetingapp.cgi/Paper/170048

Maduri, P. K., Dhiman, P., Shukla, M. R., Anand, S., & Singh, S. P. (2021, December 17-18). Farmers agriculture assistance chatbot. *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)* [Conference presentation], Greater Noida, India. 1884–1889. https://doi.org/10.1109/ICAC3N53548.2021.9725634

Musheyev, D., Pan, A., Gross, P., Kamyab, D., Kaplinsky, P., Spivak, M., Bragg, M. A., Loeb, S., & Kabarriti, A. E. (2024). Readability and information quality in cancer information from a free vs paid chatbot. *JAMA Network Open*, *7*(7), e2422275. https://doi.org/10.1001/jamanetworkopen.2024.22275

Olla, P., Wodwaski, N., & Long, T. (2025). Beyond the bot: A dual-phase framework for evaluating ai chatbot simulations in nursing education. *Nursing Reports*, *15*(8), 280. https://doi.org/10.3390/nursrep15080280

Sánchez-Gómez, J. S., Trujillo-Díaz, J., Mina, C. F. G., Martinez, C. D. A., Ruiz, F. V. G., Becerra-Fernández, M., & Díaz-Piraquive, F. N. (2026). Reco: An AI-powered chatbot system for optimizing reverse agri-food logistics and fighting food insecurity in Colombia. In H. Florez, L. Rabelo, & C. Diaz (Eds.), *Industrial Engineering and Operations Management* (Vol. 2557, pp. 193–209). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-98235-4_14

Xiao, T., Nerini, F. F., Matthews, H. D., Tavoni, M., & You, F. (2025). Environmental impact and net-zero pathways for sustainable artificial intelligence servers in the USA. *Nature Sustainability*, *8*(12), 1541–1553. https://doi.org/10.1038/s41893-025-01681-y