ASSOCIATION FOR COMMUNICATION EXCELLENCE

Journal of Applied Communications

Volume 110 | Issue 1 Article 4

Agricultural Program Evaluation Meets Artificial Intelligence: An Integrative Literature Review Brief

Julysa A. Benitez
Texas A&M University

Miguel Diaz-Manrique
Virginia Polytechnic Institute and State University

Karissa Palmer
Texas A&M AgriLife Extension

See next page for additional authors

Recommended Citation
Benitez, J. A., Diaz-Manrique, M., Palmer, K., Fuller, E. R., Landaverde, R., Leggette, H. R., & Wingenbach, G. (2026). Agricultural program evaluation meets artificial intelligence: An integrative literature review brief. *Journal of Applied Communications*, *110*(1), 1–10. https://doi.org/10.4148/jac.20854

Agricultural Program Evaluation Meets Artificial Intelligence: An Integrative Literature Review Brief

Abstract

The need for effective program evaluation in the agricultural industry is paramount to making efficient and well-informed decisions that meet the global demands for increasing agricultural productivity. Thus, the purpose of our study was to investigate the intersection of artificial intelligence (AI) and program evaluation, emphasizing its potential implications and applicability in agricultural contexts. Using integrative literature review methods, we reviewed existing literature to identify and analyze various AI tools and techniques that could be beneficial to program evaluators. Through our qualitative study, we identified eight emerging themes: Machine learning, Chat GPT, predictive analytics, data mining, generative AI, AI in evaluation training, cost-efficiency of AI, and trust in AI. Our findings indicate AI has increasing potential to serve as a tool for program evaluators across the program evaluation process. However, along with the need for more research to harness AI's full potential while minimizing any risks, maintaining human oversight and ethical considerations is essential. We recommend continued education and training initiatives to equip evaluators with the skills necessary to use AI in the evaluation process.

Keywords

agriculture, artificial intelligence, education, program evaluation

Authors

Julysa A. Benitez, Miguel Diaz-Manrique, Karissa Palmer, Emily R. Fuller, Rafael Landaverde, Holli R. Leggette, Gary Wingenbach

Introduction

In agricultural communications, programs often address complex global issues such as food security, population growth, climate change, and evolving policy demands (Thompson et al., 2021), making evaluation particularly challenging. These challenges indicate a need for further research on AI use in program evaluation and transferability to agricultural communication programs. Program evaluation uses systematic methods to address program operations and outcomes, aiming to measure impact and ensure effective resource use (Rossi et al., 2003). The evaluation process involves: 1) Identifying and describing programs and program stakeholders; 2) Drafting the evaluation purpose; 3) Writing evaluation questions and objectives; 4) Establishing methodology; 5) Planning data collection; 6) Collecting data; 7) Analyzing data; and 8) Reporting and using evaluation findings (Rossi et al., 2003). Strategies to increase evaluation efficiency include the integration of artificial intelligence (AI; Nielson, 2023; Raftree, 2023).

AI refers to computer systems that have characteristics of human intelligence and possess the ability to learn and perform any intellectual task (Marco et al., 2024). Within AI, there are many subsets and techniques, including machine learning (ML), predictive analytics, data mining, and other AI derivations (Gabsi, 2024). AI is emerging as a transformative technology in program evaluation, as demonstrated by growing interest in the topic, evidenced by numerous sessions on AI at the American Evaluation Association conference (Nielson, 2023; Patel, 2023; Raftree, 2023; Tilton et al., 2023). Thus, the purpose of our study was to explore existing literature on the use of AI tools in evaluation broadly and discuss potential adaptations to agricultural communications programming. To achieve our purpose, we sought to answer the following questions: (1) How can AI assist in the evaluation process? (2) How can AI applications currently apply in evaluation processes be adapted to agricultural communications? and (3) What areas of program evaluation should future research prioritize to advance the use of AI?

Methods

To answer the research questions, we conducted an integrative literature review (ILR). An ILR involves collecting both primary research studies and secondary documents (e.g., opinion pieces, discussion papers, policy documents; Lubbe et al., 2020). The flexibility of an IRL approach allowed us to include diverse source types, given the limited scholarly literature on AI in program evaluation. Lubbe et al. (2020) described five steps for conducting an ILR: (1) composing a research question, (2) sampling the literature, (2a) scoping search and search strategy development, (2b) screening, (2c) selections of full text documents, (3) critical appraisal of sample, (4a) data extraction, (4b) synthesis and thematic analysis, and (5) presentation of the findings and discussion.

Following the steps outlined by Lubbe et al. (2020), we first identified a gap in the literature regarding the use of AI in program evaluation, which shaped our purpose and research questions as previously presented. We established a sampling strategy using inclusion criteria for scholarly publications, websites, working papers, and blogs. Scholarly publications and working papers were found through Research Rabbit or Google Scholar, while opinion pieces, including popular press and policy documents, were accessed through a Google search engine. Although fundamental studies on AI exist, there was limited mention about the role of AI in program evaluation, as AI's relevance to evaluation has only increased recently (Chen et al., 2022). Hence, we decided to exclude publications before the year 2000.

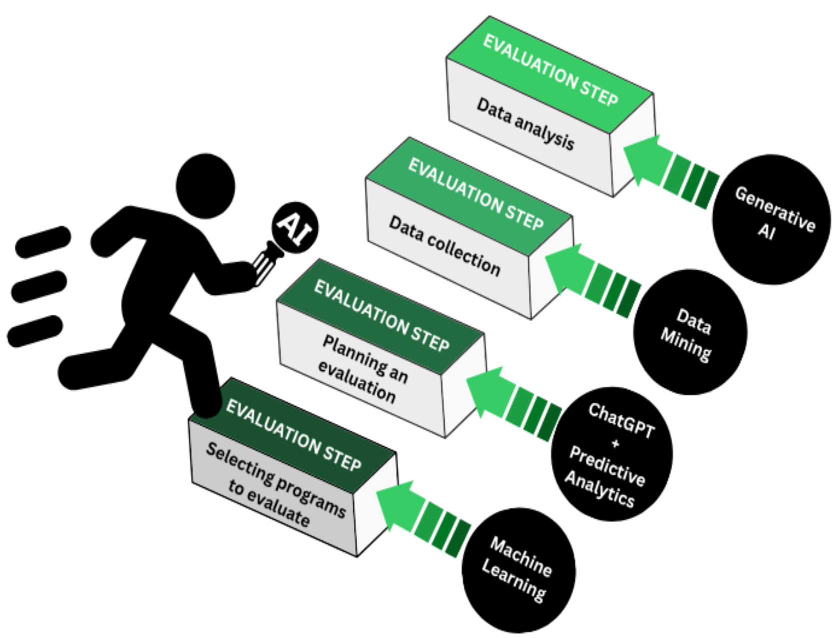
To guide our search, we developed a list of keywords including AI, program evaluation, integrative literature review, generative AI (GAI), machine learning, emerging technologies, digital technologies, natural language processing (NLP), data mining, evaluation services, evaluative knowledge, and evaluation industry. We screened 14 scholarly articles and six opinion pieces. After screening, nine scholarly articles and six opinion pieces were included in the data analysis. Through thematic analysis, we identified eight themes: ML, ChatGPT, predictive analysis, data mining, GAI, AI in program evaluation training, cost-efficiency of AI, and trust in AI. To ensure trustworthiness, we maintained an audit trail of literature, as well as a reflexive journal, and held regular peer debriefing sessions. Through peer debriefing, we compared our analyses and revisited themes until we reached an agreement on the final set of themes.

Results

With consideration to Rossi et al.'s (2003) evaluation process, Figure 1 illustrates the integration of AI tools throughout the steps, based on findings from the literature. We simplified the evaluation steps to support clearer representation.

Figure 1

Integration of AI Tools in the Program Evaluation Process



Note. This figure was developed based on the work of Beitel (2005), Grim (2023), Head et al. (2023), Jasper et al. (2019), Patel (2023), Tilton et al. (2023), and Zulaikha et al. (2020).

Machine Learning for Identifying Programs to Evaluate

ML can support evaluators by using topic modelling—the ability to insert specific topics and keywords—to select projects for evaluation. Rather than manually gathering information from projects to identify the subject matter of interest, topic modeling systematically provides the main ideas of each project corresponding to the desired focus, increasing the efficiency of selection of projects for evaluation (Jasper et al., 2019). In the context of agricultural communications, programs can cover many areas, making the screening process labor-intensive. ML enables evaluators to select projects more efficiently and rigorously identify trends in data that may not be obvious to human eyes (Jasper et al., 2019).

ChatGPT and Predictive Analytics for Planning an Evaluation

ChatGPT can briefly assist evaluators with brainstorming questions and drafting evaluation methods (Grim, 2023). According to Grim (2023), ChatGPT gave them a template for a mixed-methods program evaluation plan and highlighted the usefulness of this tool for evaluation field novices who seek guidance. Similarly, predictive analytics can support evaluators in developing an evaluation plan by using historical data to predict future outcomes (Zulaikha et al., 2020). According to Scharenbroch (2015), predictive analytics allows evaluators to better understand the program stakeholders, resulting in more intentional feedback for program improvements and making predictions on limitations, biases, and needed program refinements (Scharenbroch, 2015). Specific to agricultural communications programs, predictive analytics could guide evaluators in assessing the social impact of events, training, organizational improvements, and activities targeted to specific demographics (Taliep, 2015).

Data Mining for Data Collection

Data mining can be useful in program evaluation when information is clustered by district or demographics, such as education districts' data (Beitel, 2005). Data mining is supported by technologies such as NLP and Generative Language Models to process, organize, and analyze vast quantities of data in short periods of time (Beitel, 2005; Head et al., 2023; Patel, 2023; Tilton et al., 2023). With data mining, it is possible to reveal trends and causal relationships (Beitel, 2005; Head et al., 2023; Patel, 2023). A data mining advantage is the versatility in processing qualitative and quantitative data, common in complex evaluation contexts (Linden & Yarnold, 2016). As for agricultural communication programs, where impact may vary across regions or populations, data mining can be beneficial in identifying patterns and disparities and may offer valuable insights for more targeted communications.

Generative Artificial Intelligence for Data Analysis

GAI can analyze data and provide elaborate feedback almost immediately, which can be a timely task for an evaluator to complete (Patel, 2023). GAI is a technology able to create and analyze new data or information without any intervention by humans (Patel, 2023). In agricultural communications, programs frequently emerge to enhance the understanding of topics like crop production, livestock production, environmental issues affecting agriculture, sustainability, food security, etc. (Bartol, 2023). Yet, with GAI, evaluators can assess multiple programs in different areas more efficiently. For instance, the United Nations Educational, Scientific and Cultural Organization's evaluation report is limited to analyzing approximately 36 evaluations in a year on the most pressing topics (Hemmerich, 2023). With AI, evaluators can analyze over 1,000 evaluation reports from various United Nations agencies to gain insights into agency performance (Hemmerich, 2023).

AI in Program Evaluation Training

AI is emerging as a tool for training the next generation of program evaluators. With GAI tools, we can create customized evaluation scenarios and realistic simulations so that future evaluators can have learning-by-doing training in a controlled setting (Tilton et al., 2023). With simulations, evaluators could simultaneously apply theories, methodologies, and analyses to real-life-inspired simulated cases. Beyond training, NLP can support program evaluation literature translation, facilitating access to material that was previously limited to English speakers. In the case of international agricultural communications programs, AI-simulated scenarios and NLP can support evaluators in navigating traditions, social norms, and languages.

Further, to prepare evaluators to use these AI tools, Tilton et al. (2023) suggested opening a special chapter for AI in program evaluation curriculums. Ferretti (2023) and Head et al. (2023) add that the AI chapters should train evaluators on AI capabilities, limitations, potential risks, and how to generate appropriate prompts for optimal results.

Cost-efficiency of AI

Implementing AI is not necessarily related to decreasing evaluation cost rates. Head et al. (2023) noted access to high-quality AI tools represents a cost for evaluation providers, as software licenses, technical support, and data storage capacity are needed when using AI artifacts. By implementing AI, evaluation market dynamics will be influenced by price and quality of the process, affecting the public and social sectors, which are the highest procurers of program evaluation services (Nielsen, 2023). Despite these challenges, AI tools could still be beneficial in the evaluation process (Raftree, 2023), particularly in the agricultural communications sector, by facilitating the gathering, interpreting, and sharing of information from diverse voices.

Trust in AI

The complexity and lack of understanding of AI tools undermine their trustworthiness and have raised warnings about their reliability and safety (Cvetkovic et al., 2024; Habbal et al., 2024; Wong et al., 2023). According to Habbal et al. (2024) and Cvetkovic et al. (2024), users' perceptions of technology reliability, transparency, explainability, fairness, and accountability define a trustworthy relationship in AI models. As evaluators become more familiar with AI, their attitudes will be more positive if the technology is safe, offers competent outputs, and fulfills the expected tasks (Castelfranchi & Falcone, 2000; Glikson & Woolley, 2020; Hoff & Bashir, 2015; Kaushik et al., 2015; Tussyadiah et al., 2020).

Insecurities about job stability, vulnerability to cyberattacks, fears about harmful content, lack of accountability, and ethical considerations can influence evaluators' trust in AI (Fast & Horvitz, 2017). Personal characteristics such as age, gender, cultural background, technology proficiency, perceptions of usefulness, and individual attitudes may also play a role in evaluators' trust and predisposition to using AI (Hoff & Bashir, 2015). Therefore, evaluation agencies must design clear procedures and address all questions regarding when and how to use AI, as well as ownership and control over the information shared with AI tools (Cvetkovic et al., 2024). Further, transparent communication between evaluators and clients is needed to ensure AI tools are used ethically and efficiently (Cvetkovic et al., 2024).

Discussion, Conclusions, and Recommendations

ML was found useful in selecting programs to evaluate (Jasper et al., 2019). ChatGPT and predictive analytics have offered inspiration to evaluators when drafting evaluation plans (Grim, 2023; Scharenbroch, 2015). Data mining aids in program evaluation, data collection, and organization (Beitel, 2005; Head et al., 2023; Patel, 2023; Tilton et al., 2023) while GAI can streamline data analysis almost immediately (Patel, 2023).

Many of the AI tools currently being used in program evaluation can be adapted for agricultural communication programs, especially those that involve large datasets and diverse populations. Data mining tools and predictive analytics can help assess agricultural programs' effectiveness across diverse farming communities or climates (Beitel, 2005; Linden & Yarnold, 2016) while GAI can increase efficiency in the evaluation of multiple programs in different priority areas (Hemmerich, 2023). Additionally, the potential use of AI to train future evaluators through simulated scenarios (Tilton et al., 2023) along with the need to first train users about all considerations regarding AI technologies was highlighted.

For agricultural communicators, AI can facilitate the scanning of communication campaigns on social media, news stories, podcasts, and other agricultural communication outputs, allowing them to identify relevant topics that may require more immediate programming. For example, if the news is increasingly covering repeated food recalls linked to a specific product, evaluators could be prompted to prioritize food safety communication campaigns that otherwise would not have been a focus. Further, AI could help outline improved communication strategies and tailor messages to specific audience needs to improve program outcomes (Prestegaard-Wilson & Vitale, 2024).

The challenges of AI implementation in program evaluation include a lack of knowledge and training on AI technology use, cost, trust, and cultural sensitivity (Ferretti, 2023; Head et al., 2023; Patel, 2023). We recommend evaluators continue educating themselves on intellectual property rights, content reproduction, and legal issues using AI tools for evaluation reports. Evaluators should also work with AI experts to shape ethical frameworks and guidelines to support the better use, design, and development of AI tools (Patel, 2023). Agricultural communicators could collaborate with evaluators to leverage AI responsibly and ethically to translate evaluation findings into actionable information for all stakeholders. More comprehensive research is needed regarding the benefits, challenges, limitations, and potential negative implications of AI in the evaluation sector, specifically in the agricultural context.

References

Bartol, T. (2023). Smallholders and small-scale agriculture: Mapping and visualization of knowledge domains and research trends. *Cogent Social Sciences*, *9*(1). https://doi.org/10.1080/23311886.2022.2161778

Beitel, S. E. (2005). *Applying artificial intelligence data mining tools to the challenges of program evaluation* (Doctoral dissertation, University of Connecticut). ProQuest Dissertations Publishing. https://www.proquest.com/docview/305010432/abstract/EAD3A00DDE6A44CDPQ/1

Castelfranchi, C., & Falcone, R. (2000). Trust and control: A dialectic link. *Applied Artificial Intelligence*, *14*(8), 799–823. https://doi.org/10.1080/08839510050127560

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education. *Educational Technology & Society*, *25*(1), 28–47. https://www.jstor.org/stable/48647028

Cvetkovic, A., Savela, N., Latikka, R., & Oksanen, A. (2024). Do we trust artificially intelligent assistants at work? An experimental study. *Human Behavior and Emerging Technologies*, *2024*(1), 602237. https://doi.org/10.1155/2024/1602237

Fast, E., & Horvitz, E. (2017). *Long-term trends in the public perception of artificial intelligence*. Proceedings of the AAAI Conference on Artificial Intelligence, 31(1), Article 1. https://doi.org/10.1609/aaai.v31i1.10635

Ferretti, S. (2023). Hacking by the prompt: Innovative ways to utilize ChatGPT for evaluators. *New Directions for Evaluation*, *2023*(178–179), 73–84. https://doi.org/10.1002/ev.20557

Gabsi, A. E. H. (2024). Integrating artificial intelligence in industry 4.0: Insights, challenges, and future prospects–a literature review. *Annals of Operations Research*, 1–28. http://dx.doi.org/10.1007/s10479-024-06012-6

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Grim, E. (2023, May 12). *Leveraging language models like CHATGPT for evaluation*. https://web.archive.org/web/20230314140104/https://elizabethgrim.com/leveraging-language-models-like-chatgpt-for-evaluation/

Habbal, A., Ali, M. K., & Abuzaraida, M. A. (2024). Artificial intelligence trust, risk and security management (AI TRiSM): Frameworks, applications, challenges and future research directions. *Expert Systems with Applications*, 240, 122442. https://doi.org/10.1016/j.eswa.2023.122442

Head, C. B., Jasper, P., McConnachie, M., Raftree, L., & Higdon, G. (2023). Large language model applications for evaluation: Opportunities and ethical implications. *New Directions for Evaluation*, *2023*(178–179), 33–46. https://doi.org/10.1002/ev.20556

Hemmerich, K. (2023, October 8). *Using AI to power organizational learning from evaluations*. ReformWorks. https://www.reformworks.org/post/using-ai-to-power-organizational-learning-from-evaluations

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Jasper, P., Haldrup, S. V., & Mikhaylov, S. J. (2019, March). *Artificial intelligence and machine learning methods for programme evaluations in global affairs Canada*. Oxford Policy Management. https://www.opml.co.uk/sites/default/files/migrated_bolt_files/a3489-ai-and-ml-for-evaluations.pdf

Kaushik, A. K., Agrawal, A. K., & Rahman, Z. (2015). Tourist behaviour towards self-service hotel technology adoption: Trust and subjective norm as key antecedents. *Tourism Management Perspectives, 16*, 278–289. https://doi.org/10.1016/j.tmp.2015.09.002

Linden, A., & Yarnold, P. R. (2016). Combining machine learning and matching techniques to improve causal inference in program evaluation. *Journal of Evaluation in Clinical Practice*, *22*(6), 868–874. https://doi.org/10.1111/jep.12592

Lubbe, W., ten Ham-Baloyi, W., & Smit, K. (2020). The integrative literature review as a research method: A demonstration review of research on neurodevelopmental supportive care in preterm infants. *Journal of Neonatal Nursing*, *26*(6), 308–315. https://doi.org/10.1016/j.jnn.2020.04.006

Marco, G., Evertsson, E., Riley, D. J., Tyrchan, C., & Rathi, P. C. (2024). Augmenting DMTA using predictive AI modelling at AstraZeneca. *Drug Discovery Today*, *29*(4), 103945. https://doi.org/10.1016/j.drudis.2024.103945

Nielsen, S. B. (2023). Disrupting evaluation? Emerging technologies and their implications for the evaluation industry. *New Directions for Evaluation*, *2023*(178–179), 47–7. https://doi.org/10.1002/ev.20558

Patel, D. (2023). Revolutionizing program evaluation with generative AI: An evidence-based methodology. *International Journal for Multidisciplinary Research*, *5*(3), 1–11. https://www.ijfmr.com/papers/2023/3/4105.pdf

Prestegaard-Wilson, J., & Vitale, J. (2024). Generative artificial intelligence in extension: A new era of support for livestock producers. *Animal Frontiers*, *14*(6), 57–59. https://doi.org/10.1093/af/vfae024

Raftree, L. (2023, November 8). *What's next for emerging AI in evaluation? Takeaways from the 2023 AEA conference.* MERL Tech. https://merltech.org/emerging-ai-for-evaluation/

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A systematic approach.* SAGE Publications.

Scharenbroch, C. (2022, October 24). *Big data and predictive analytics in program evaluation. Evident Change*. https://web.archive.org/web/20230206225431/https://evidentchange.org/blog/big-data-and-predictive-analytics-in-program-evaluation/

Taliep, N. (2015). *Process evaluation of the development of a community-based participatory intervention promoting positive masculinity and peace and safety: Addressing interpersonal violence in a Western Cape community* (Doctoral dissertation, University of South Africa). https://ir.unisa.ac.za/handle/10500/20226

Thompson, N. M., Reeling, C.J., Fleckenstein, M.R., Prokopy, L.S., & Armstrong, S.D. (2021). Examining intensity of conservation practice and adoption: Evidence from cover crop use on U.S. Midwest farms. *Food Policy, 101*, 1–10. https://doi.org/10.1016/j.foodpol.2021.102054

Tilton, Z., LaVelle, J. M., Ford, T., & Montenegro, M. (2023). Artificial intelligence and the future of evaluation education: Possibilities and prototypes. *New Directions for Evaluation, 2023*(178–179), 97–109. https://doi.org/10.1002/ev.20564

Tussyadiah, I. P., Zach, F. J., & Wang, J. (2020). Do travelers trust intelligent service robots? *Annals of Tourism Research, 81*, 102886. https://doi.org/10.1016/j.annals.2020.102886

Wong, L.-W., Tan, G. W.-H., Ooi, K.-B., & Dwivedi, Y. (2023). The role of institutional and self in the formation of trust in artificial intelligence technologies. *Internet Research, 34*(2), 343–370. https://doi.org/10.1108/INTR-07-2021-0446

Zulaikha, S., Mohamed, H., Kurniawati, M., Rusgianto, S., & Rusmita, S. A. (2020). Customer predictive analytics using artificial intelligence. *The Singapore Economic Review*, 1–12. https://doi.org/10.1142/S0217590820480021